In memory of my friend Kolya K. (1933–2014)

“Independence is the best quality, the best word in all languages.”
J. Brodsky. From a personal letter.

The theory of filtrations of subalgebras, standardness, and independence

A. M. Vershik

Abstract. This survey is devoted to the combinatorial and metric theory of filtrations: decreasing sequences of \( \sigma \)-algebras in measure spaces or decreasing sequences of subalgebras of certain algebras. One of the key notions, that of standardness, plays the role of a generalization of the notion of the independence of a sequence of random variables. Questions are discussed on the possibility of classifying filtrations, on their invariants, and on various connections with problems in algebra, dynamics, and combinatorics.

Bibliography: 101 titles.

Keywords: filtrations, \( \sigma \)-algebras, independence, standardness, graded graphs, central measures.

Contents

1. Introduction 258
   1.1. A simple example and a difficult question 259
   1.2. Three languages of measure theory 261
   1.3. Where do filtrations appear? 261
   1.4. Finite and infinite in classification problems; standardness as a generalization of independence 264
   1.5. A summary of the paper 267
2. The definition of filtrations in measure spaces 268
   2.1. Measure theory: a Lebesgue space, basic facts 268
   2.2. Measurable partitions, filtrations 269
   2.3. Classification of measurable partitions 270
   2.4. Classification of finite filtrations 272
   2.5. Filtrations we consider and how one can define them 274
   2.6. The equivalence relation associated with a filtration, cocycles and associated measures 277

Partially supported by the Russian Science Foundation (grant no. 14-11-00581).

AMS 2010 Mathematics Subject Classification. Primary 05A05, 37A60, 37M99, 28A06, 60A10.

© 2017 Russian Academy of Sciences (DoM), London Mathematical Society, Turpion Ltd.
1. Introduction

This survey deals with an area in which measure theory is intertwined with combinatorics and asymptotic analysis. Its applications to dynamics, algebra, and other problems will be partly touched upon in this paper, as well as in subsequent publications. The survey contains many new problems from the dynamic theory of graphs and representation theory for groups and algebras.
1.1. A simple example and a difficult question. We begin with an elementary example that illustrates the notion of filtration and problems of combinatorial measure theory.

Consider the space of all one-sided sequences of zeros and ones:

\[ X = \{ \{x_n\}_{n=1}^{\infty} \}, \quad x_n = 0 \lor 1, \quad n = 1, 2, \ldots, \]

that is, the infinite product of the two-point space. We will regard \( X \) as a dyadic (Cantor-like) compactum in the weak topology, and also as a Borel space. In \( X \) we introduce the ‘tail’ equivalence relation and the tail filtration of \( \sigma \)-subalgebras of sets. Namely, two sequences \( \{x_k\}, \{x'_k\} \) are

- \( n \)-equivalent if \( x_{k+n} = x'_{k+n} \) for all \( k \geq 0 \), and
- equivalent with respect to the tail equivalence relation if they are \( n \)-equivalent for some \( n \).

Denote by \( \mathcal{A}_n, n = 0, 1, 2, \ldots \), the \( \sigma \)-algebra of Borel subsets of \( X \) that along with every point \( x \) contain all points \( n \)-equivalent to \( x \). In other words, \( \mathcal{A}_n, n = 0, 1, 2, \ldots \), is the \( \sigma \)-subalgebra of Borel sets determined by conditions on the coordinates with indices \( \geq n \). The decreasing sequence

\[ \mathcal{A}_0 \supset \mathcal{A}_1 \supset \mathcal{A}_2 \supset \cdots \]

is called the tail Borel filtration of the space \( X \) regarded as an infinite direct product,

\[ X = \prod_{n=1}^{\infty} [0; 1]. \]

If we endow \( X \) with an arbitrary Borel probability measure \( \mu \), then the same sequence of \( \sigma \)-subalgebras, now understood as \( \sigma \)-subalgebras of classes of sets coinciding \( \text{mod } 0 \) (with respect to the given measure \( \mu \)), provides an example of a filtration in a standard measure space.

For instance, take \( \mu \) to be the Bernoulli measure equal to the infinite product of the measures \( \theta = (1/2, 1/2) \); the resulting filtration is called the dyadic Bernoulli filtration. By the famous Kolmogorov zero–one (‘all-or-none’) law, the intersection \( \bigcap_n \mathcal{A}_n \) is the trivial \( \sigma \)-algebra \( \mathcal{G} \), which contains only two classes of sets, those of measure either zero or one.

For every positive integer \( n \) the space \( (X, \mu) \) can be represented as a direct product of measure spaces:

\[ (X, \mu) = \left( \prod_{1}^{n} \{(0; 1), \theta]\right) \times (X_n, \mu_n), \]

where \( (X_n, \mu_n) = \prod_{n+1}^{\infty} \{(0; 1), \theta\} \).

Now we state the main problem. Assume that we have a dyadic compactum \( X' \) with a Borel probability measure \( \mu' \) satisfying Kolmogorov’s zero–one law, and for every positive integer \( n \) there is an isomorphism of measure spaces

\[ (X', \mu') \simeq \left( \prod_{1}^{n} \{(0; 1), \theta]\right) \times (X'_n, \mu'_n), \]

where \( (X'_n, \mu'_n) \) are some measure spaces.
Can we assert that there exists an isomorphism \( T \) of measure spaces, \( T(X', \mu') = (X, \mu) \), such that \( T(X'_n, \mu') = (X_n, \mu) \) for all \( n \)?

In the language to be described below, this is the question of whether a finitely Bernoulli dyadic ergodic filtration is unique up to isomorphism.

It is similar to the following question from the theory of infinite tensor products of \( C^* \)-algebras and perhaps looks even more natural.

Consider a \( C^* \)-algebra \( \mathcal{A} \) and assume that there is a decreasing sequence \( \mathcal{A}_n \subset \mathcal{A} \) of \( C^* \)-subalgebras of \( \mathcal{A} \) with intersection equal to the space of constants, \( \bigcap_n \mathcal{A}_n = \{\text{Const}\} \), and such that for all \( n \)

\[
\mathcal{A} \simeq M_2(\mathbb{C})^\otimes n \otimes \mathcal{A}_n
\]

(here \( \simeq \) means an isomorphism of \( C^* \)-algebras).

Is it true that there exists an isomorphism

\[
\mathcal{A} \simeq \prod_1^\infty M_2(\mathbb{C})
\]

that sends \( \mathcal{A}_n \) to the subalgebra \( \prod_{n+1}^\infty M_2(\mathbb{C}) \subset \mathcal{A} \) for all \( n \)?

These problems were raised by the author about fifty years ago. In the original terms, the first of them was phrased as follows: is it true that every filtration such that its finite segments are isomorphic to finite segments of a Bernoulli filtration and the intersection of all its \( \sigma \)-algebras is trivial, is isomorphic to a Bernoulli filtration? The problem originated from ergodic theory, as we will discuss below. In [59] both questions were answered in the negative; moreover, it turned out that there is a continuum of measure spaces (respectively, \( C^* \)-algebras) with dyadic structures that are finitely isomorphic but pairwise non-isomorphic globally. This discovery launched an important field, which, however, still has not gained sufficient attention: the asymptotic theory of filtrations. For this survey we have selected the most important facts, both long known and new, and (most importantly) new problems in this area. The theory of filtrations has applications to measure theory, random processes, and dynamical systems, as well as to representation theory of groups and algebras. The main problem concerns the complexity of the asymptotic behaviour at infinity of monotone sequences of algebras and a possible classification of such sequences. The second question, about tensor products of \( C^* \)-algebras, will be considered in another paper; we have mentioned it here to emphasize the parallelism between problems from very different areas of mathematics. Both these problems belong to asymptotic algebraic analysis.

Conceptually new ideas, as compared with previous research on filtrations, appeared in connection with the theory of graded graphs (Bratteli diagrams); this is one of the central topics of our survey. Apparently, this connection was not noticed earlier, though the tail filtration is a most important object in the theory of AF-algebras. The author considered filtrations from the standpoint of the theory of graded graphs in his recent paper [82], but here, on the other hand, the focus is on the measure-theoretic aspect of the problem and the use of notions and techniques related to graphs.
1.2. Three languages of measure theory. Let us return to the measure-theoretic statement of the first problem. If we consider the algebra $L^\infty(X, \mu)$ of all classes of bounded measurable functions coinciding mod 0 on the space $(X, \mu)$, we obtain a filtration of subalgebras:

$$L^\infty(X, \mu) \equiv \mathcal{A}_0 \supset \mathcal{A}_1 \supset \mathcal{A}_2 \supset \cdots,$$

where $\mathcal{A}_n$ is the subalgebra of $L^\infty(X, \mu)$ consisting of all functions that depend on coordinates with indices $\geq n$ ($n = 0, 1, 2, \ldots$).

Now let us describe this example in the language of partitions. Let $\xi_n$ be the partition of $(X, \mu)$ into the classes of sequences in which the coordinates with indices greater than $n$ coincide; then $\mathcal{A}_n$ is the $\sigma$-algebra of sets measurable with respect to the partition $\xi_n$ (that is, sets composed of elements of this partition). And $\mathcal{A}_n$ is the space of functions measurable with respect to $\xi_n$. The fact that the sequence of partitions is decreasing means that almost every element of $\xi_n$ is a union of some elements of $\xi_{n-1}$, $n = 1, 2, \ldots$ (The partial order in the space of measurable partitions is discussed in §2.2.)

In these terms, $\mathcal{A}_n$ is the algebra $L^\infty(X/\xi_n, \mu_n)$, where $X/\xi_n$ is the quotient of $(X, \mu)$ by $\xi_n$ and $\mu_{\xi_n}$ is the quotient measure on $X/\xi_n$.

Thus, there is a functorial equivalence of the three languages of the theory of filtrations described above: the language of filtrations of $\sigma$-subalgebras of a measure space, the language of filtrations of subspaces of measurable functions, and the language of filtrations of measurable partitions (precise definitions will be given below).

By filtrations we will mean either infinite decreasing sequences of subalgebras of a commutative algebra with involution (for instance, $L^\infty(X, \mu)$), or (equivalently) infinite decreasing sequences of $\sigma$-subalgebras of measurable sets in a standard measure space $(X, \mu)$, or, finally (in yet another equivalent geometric language which we will use most frequently), infinite decreasing sequences of measurable partitions.

The three languages are equivalent; the difference between them is terminological. The first context (filtrations of subalgebras of an algebra with involution) is the most general one, and it admits an important generalization achieved by abandoning the commutativity of the ambient algebra and passing to AF-algebras; as mentioned above, this generalization is not touched upon here, although it is closely related to the content of the paper.

We will also consider decreasing sequences of $\sigma$-subalgebras of sets in a standard Borel space, that is, Borel filtrations. In this case no measure on the space is given, and the problem arises of describing all measures that agree with the given filtration. This statement covers almost all problems related to invariant measures in diverse contexts.

1.3. Where do filtrations appear? Filtrations arise in the theory of random processes, potential theory and the theory of Markov processes, ergodic theory, and topological and metric dynamics.

The most interesting applications relate to the theory of graded graphs (Bratteli diagrams), the theory of AF-algebras, and asymptotic combinatorics. Here are several general examples.
A. The ‘past’ of random processes. Consider an arbitrary random process, for instance, a process \( \{x_m\}_m \) with real values and discrete time, \( m \leq 0 \) (it is convenient to index the random variables by negative integers).

The \( \sigma \)-algebra \( \mathfrak{A}_n \) consists of all measurable sets that can be described by the ‘past’ of the process, that is, by random variables with indices \( m < -n \). In this context \( \{\mathfrak{A}_n\}_n \) is the past filtration of the random process. In a similar way we can define the ‘future’ filtration of a random process. A case of importance in ergodic theory is that of a stationary filtration:

\[
\{\xi_n\}_{n=0}^{\infty} \simeq \{\xi_n / \xi_1\}_{n=1}^{\infty},
\]

where \( \simeq \) means metric isomorphism, that is, the existence of a measure-preserving automorphism that sends the first filtration to the second one. In this case the process \( \{x_m\}_m \) described above is stationary, that is, the corresponding probability measure is invariant under the left shift \( \hat{T}\{x_m\}_m = \{x_{m-1}\}_m, m < 0 \). In other words, the shift executes the passage to the quotient by the first partition, and the quotient filtration is isomorphic to the original one. There are several papers on the theory of stationary filtrations; for applications of the standardness criterion, see §5. However, one should keep in mind that far from all invariants of a stationary filtration are endomorphism invariants. A remarkable example of such an invariant is Ornstein’s ‘very weak Bernoulli’ (VWB) condition.

B. Filtrations in approximation theory. Consider an ergodic automorphism of a measure space \( (X, \mu) \): \( T: X \to X, T\mu = \mu \). By the classical Rokhlin theorem, it can be approximated in the uniform topology by periodic automorphisms. Periodic approximations can be amended in such a way that the partitions into their trajectories decrease, that is, form a filtration of partitions into the trajectories of periodic approximations.

It is this construction that leads to the so-called adic realization of an automorphism [67], [68]. Many important metric invariants of the automorphism can be extracted from the properties of this filtration; this will also be discussed below.

This example can be extended to amenable groups and amenable actions of arbitrary groups (see [37]). Consider a dynamical system in a space \( X \) with an invariant measure \( \mu \) in which ‘time’ is a countable amenable group \( G \). In other words, the group \( G \) is represented by automorphisms of the space \( (X, \mu) \), that is, measurable measure-preserving transformations:

\[
T: G \to \text{Aut}(X, \mu),
\]

\[
g \mapsto T_g : X \to X,
\]

\[
\mu \mapsto \mu.
\]

A fundamental theorem of ergodic theory says that if the group \( G \) is amenable, that is, has an invariant mean, then the orbit partition of \( G \) is tame (hyperfinite), that is, there exists a (non-unique) filtration that determines a decreasing sequence of partitions which tends to the partition into the trajectories of the group action. There is a particularly natural connection between the theory of filtrations and the theory of actions of locally finite groups: in this case, \( G = \bigcup_n G_n \) is an inductive
limit of finite groups, and for every action of \( G \) there is a canonically defined sequence of partitions into the trajectories of the finite subgroups \( G_n \).

Some invariants of filtrations constructed in this way, for example, entropy (see below), are also invariants of the action.

The adic realization of a group action is a kind of alternative to the generally accepted symbolic realization of this action by a group of shifts in a space of functions on the group. It can be regarded as a generalization of the odometer to an arbitrary graded graph. This realization determines an approximation of the group action by finite groups, and thus it is a far-reaching generalization of Rokhlin’s lemma on periodic approximation.

C. Invariant and central measures. Assume that we have an \( \mathbb{N} \)-graded locally finite graph \( \Gamma = \bigsqcup_n \Gamma_n, \Gamma_0 = \{ \emptyset \} \), and let \( T(\Gamma) \) be the space of infinite paths in \( \Gamma \) starting from the vertex \( \emptyset \). The space \( T(\Gamma) \), being an inverse spectrum of finite spaces (of finite paths), is a Cantor-like topological space, and hence is endowed with a Borel structure. In \( T(\Gamma) \) the tail filtration is defined: the \( n \)-th \( \sigma \)-algebra consists of all sets that do not change if we take an arbitrary path from such a set and modify its initial segment up to level \( n \). If one endows the space of paths with a Borel measure, then the tail filtration turns into the filtration from Example A above. But here a new problem arises, related to measures with given cotransition probabilities, namely, the problem of describing the central measures (see §7). This problem appeared in the theory of \( C^* \)-algebras (AF-algebras) and locally finite groups (the list of traces and characters), in the theory of Markov processes and invariant Markov measures (‘entrance and exit boundaries’ in the sense of Dynkin), and in the theory of invariant measures for dynamical systems. This circle of problems has equivalent combinatorial and geometric formulations that are interesting in their own right. We note that this class of examples of filtrations has important additional specific features which we will use in what follows.

One of the results of this paper is as follows (§4.1).

**Theorem 1.** Every ergodic locally finite filtration is isomorphic to a Markov filtration.

This means that one can study the most interesting class of filtrations by considering only one-sided Markov chains (in general, inhomogeneous in time and with an arbitrary state space). In particular, this provides a fruitful reduction of the theory of AF-algebras to the probabilistic theory of Markov chains, as well as the reverse transition from the theory of Markov chains to the theory of Bratteli diagrams with their algebraic underpinning. Actually, this theorem is a refinement of the theorem on the existence of an adic realization for actions of amenable groups. It is also appropriate to mention that theorems on tail filtrations of path spaces of graded graphs are closely related to filtrations of commutants of finite-dimensional subalgebras of AF-algebras.

D. Statistical physics and inverse limits of simplices. Finally, the most popular reason for considering filtrations is provided by the modern formalism of statistical physics, both classical and quantum. Consider a lattice, or even an
arbitrary countable graph $\Gamma$, and represent it as a union of finite subsets (volumes):

$$\Gamma = \bigcup_n \Gamma_n.$$ 

Consider the space of configurations (for example, subsets) on $\Gamma$, that is, the space of functions taking values in some alphabet (for example, the alphabet $(0,1)$). Endow it with the structure of a Borel space by taking a Borel base consisting of the cylinders, that is, the sets of configurations determined by conditions on a finite part of a configuration (its restriction to some volume $\Gamma_n$) and arbitrary outside this volume. Then we have a filtration in this Borel space: its $n$th $\sigma$-subalgebra consists of the cylinder sets determined by conditions on the restriction of a configuration to the volume $\Gamma_n$. In other words, two configurations lie in the same element of the partition if they coincide outside $\Gamma_n$ for some $n$. If we have a probability measure (statistic) on the space of configurations, then we obtain a filtration of this measure space. A popular scheme for studying phase transitions and other properties of statistics, adopted in recent years after the work of Dobrushin and others [5], [36], is as follows: one defines not a measure, but only a system of conditional probabilities of the restriction of a configuration to an arbitrary volume, under the condition that its restriction outside this volume is fixed, and then one studies the measures that have these conditional measures. This is a method of defining measures in infinite-dimensional spaces alternative to the widely accepted Kolmogorov method via joint distributions. This scheme for studying filtrations is briefly described below (see §3, especially §§3.5 and 3.6).

We emphasize that the above list of areas related to filtrations does not pretend to be exhaustive, and this survey, as well as the list of references, is far from being complete. Especially interesting and little studied connections are those with combinatorics; possible material on this topic can be drawn from [54], [55]. We mean that an asymptotic combinatorial problem always involves an inductive family of problems with a parameter (the number of objects, the collection of dimensions, and so on), but often one can also define a filtration, whose properties are hidden quite deeply, and without an analysis of these properties one cannot fully reveal the true nature of the problem. An example of such a property is standardness, to be discussed below.

1.4. Finite and infinite in classification problems; standardness as a generalization of independence. In mathematics and physics there are many classification problems involving ‘finite’ and ‘infinite’ invariants, and finding the latter ones is usually the core of a classification problem. The problem under study, that of metric classification of filtrations, seems to be typical for the class of problems involving analytic, algebraic, and combinatorial components, and moreover, there is a non-trivial relation between finite and infinite invariants. Perhaps the ideas described here will be useful also in other problems.

In what follows, classification and types of filtrations are understood in the sense of measure-preserving transformations, and equivalence is understood with respect to the group of such transformations. Instead of ‘equivalent objects’ (filtrations), we will more often say ‘isomorphic objects’.
Every infinite filtration of algebras (or similar objects)

\[ \{A_n\}_{n=0}^{\infty} = A_0 \supset A_1 \supset \cdots \]

can be regarded as an infinite collection of its finite fragments, that is, finite filtrations:

\[ \{A_k^n\}_{k=0}^n, \quad n = 0, 1, 2, \ldots . \]

Assume that we can give a classification, with respect to some equivalence relation, of finite filtrations of given length \( n \) for all finite \( n \). More precisely, assume that we can construct a topological or Borel moduli space \( \mathcal{M}_n, n = 0, 1, 2, \ldots \), that is, a space of complete invariants of finite filtrations of a given length. We will say that two infinite filtrations are \textit{finitely isomorphic if for every} \( N \) \textit{their fragments of length} \( N \) \textit{are equivalent}, that is, determine the same point in the space \( \mathcal{M}_N, N = 0, 1, 2, \ldots . \).

Of course, the fact that infinite filtrations are finitely isomorphic does not in general imply that they are isomorphic. In other words, even if the moduli space of infinite filtrations can be well defined, it is not in general the union of the moduli spaces of finite fragments of filtrations. We will consider this phenomenon in more detail below.

The main topic of this paper and the new results presented here are in one way or another related to the problem of classifying infinite filtrations and singling out the classes most important for applications. There exist infinite filtrations whose isomorphism type is uniquely determined by the finite isomorphism type. For example, if the group of symmetries of a finite fragment of a filtration is trivial, then the (isomorphism) type of this fragment already determines the type of the whole filtration. Thus, a meaningful theory should deal only with the class of filtrations for which the group of symmetries of every infinite ‘tail’ is infinite. Partly because of this, we consider only locally finite filtrations (for the definition, see §2): for them, the corresponding groups are infinite, and this feature of them is already reminiscent of the Markov property. In the class of locally finite filtrations, finite isomorphism does not impose substantial restrictions on the ‘infinite properties’ of filtrations, in particular, it does not even determine whether the filtration is ergodic or not.

Attempts to give a classification of infinite sequences should be preceded by building a classification of finite filtrations (that is, finite decreasing sequences of subalgebras), and in the categories we are interested in, it either is known or can easily be obtained. This classification was initiated by Rokhlin (in the case of one \( \sigma \)-algebra or one partition), and one can use this pattern to construct invariants of finite filtrations (see, for example, [21]). We retell the classification of arbitrary finite decreasing sequences of partitions in the convenient language of graphs and trees. In these cases the invariants, that is, the moduli spaces, are quite manageable: they are spaces of finite measured trees. This is the content of a part of §2.

But what else should we add to finite isomorphism to obtain full isomorphism? In other words, what ‘infinite invariants’ are not determined by finite ones? And do we really always need them? Are there classes of filtrations for which finite isomorphism already implies isomorphism? We note that the analysis of finite invariants does not allow one to deduce whether or not they suffice to determine the type of a filtration.
This is one of the main questions, which is rather typical for any discussions about finite and infinite.

The answer to the above question is positive: the desired class is the class of so-called standard filtrations. In the case of homogeneous (for example, dyadic) filtrations, it consists, as follows immediately from the standardness criterion (see [59] and § 4 below), of Bernoulli filtrations, that is, past filtrations of arbitrary sequences of independent random variables.

For general locally finite filtrations, even the statement of the problem is not so obvious and appears as follows. Let us try to generalize the problem of classifying filtrations, that is, to endow filtrations with a certain structure in such a way that the behaviour of finite invariants of a filtration in the augmented problem allows one to determine whether the usual finite invariants uniquely determine the type of this filtration.

Filtrations for which this condition is satisfied will be said to be standard: this class generalizes the notion of independence (= the Bernoulli property).

But how do this generalized problem and these generalized finite invariants look? The answer is that one should construct a classification of finite fragments of filtrations together with some additional data; for instance, these data can be fixed measurable functions defined on the base space, but the most appropriate choice is to take a metric on the measure space. Then the measured trees which are invariants of finite fragments of filtrations will also be endowed with a metric, and these new terms allow one to state the criterion of asymptotic behaviour of invariants, called the standardness criterion; the validity of its condition is exactly equivalent to the fact that the finite invariants completely determine the type of a filtration. Moreover, the choice of a specific metric does not affect the validity of this condition, it only has to be admissible (see [92], [93]). This is precisely the main idea, which we describe in § 5. Therefore, the moduli space is exactly the space of finite invariants, and the standardness criterion is the condition meaning that the invariants of the conditional filtrations (on the elements of the nth partition) converge in measure as n tends to infinity. Invariants of finite filtrations whose partitions all have finite elements are measures on spaces of trees endowed with a measure and a metric. Below we also state this criterion in terms similar to those of martingale theorems.

At first sight, it seems that one can continue this procedure of selecting good classes, extending the problem of classifying filtrations, and obtaining further classes subject to classification. But this is not so: outside the standardness class there is essentially one class, indivisible and very interesting, that of totally non-standard filtrations, for which there is still no complete classification (and perhaps it does not exist), but there are non-trivial invariants, both resembling old invariants like scale and (secondary) entropy, and possibly new ones such as ‘higher zero–one laws’, which provide far-reaching generalizations of Kolmogorov’s laws. Perhaps the terminology may change, because standardness can be understood both in the stringent sense (as minimality) and in the milder sense (as proper standardness); the future will reveal a convenient terminology.

It is interesting to compare this with Ornstein’s theory of the very weak Bernoulli property (VWB), d- and f-metrics, and so on; indeed, the VWB property is similar to standardness, the difference being only in the choice of metrics on the space.
of conditional measures, as well as in the fact that ergodic theory deals only with stationary filtrations.

We show that finite classification of filtrations is a tame problem, that is, the corresponding moduli space is a reasonable Borel space. But the general problem of metric classification of arbitrary filtrations is ‘wild’ in the accepted sense and seems to be just as difficult as the problem of classifying automorphisms of spaces and similar problems.

1.5. A summary of the paper. This is mainly a survey paper which partially summarizes the research on filtrations from the introduction (in 1969) of the notion of non-standardness (more precisely, the discovery of examples of non-Bernoulli filtrations finitely isomorphic to Bernoulli ones) and the standardness criterion. We omitted several important topics directly related to the theory of filtrations, for example, the notion of ‘cosy’ filtrations in the sense of Tsirelson, recent papers by Laurent, and so on. Our main purpose is, first, to extend basic results on standardness to inhomogeneous and non-finitely Bernoulli filtrations and, second, to combine the theory of filtrations with the theory of graded graphs (Bratteli diagrams), and thus with the theory of AF-algebras, the combinatorics of graphs, and Markov chains. Hence, our main interest is in locally finite and locally representable filtrations, that is, filtrations in which fibres (elements of partitions) are finite for every partition and the number of different types of conditional measures is also finite. A detailed introduction, including Rokhlin’s theory of a single partition, is given in §2. Connections with the theory of graded graphs and Markov chains are described in §§3 and 4. It should be said that the existence of these connections enriches both the theory of filtrations and the theory of graphs, Markov chains, and algebras. Examples that have appeared in recent years brilliantly demonstrate this. There have appeared many interesting and complicated graphs originating from the theory of adic (that is, ‘graphic’) approximations of dynamical systems (the graphs of ordered and unordered pairs, graphs of words, and so on) which provide examples of non-standard filtrations; on the other hand, graph theory is a source of new problems related to algebras and new realizations of filtrations. The theorem on a Markov realization of a locally finite filtration is proved in §4. The key notion of the theory is standardness. For homogeneous and finitely Bernoulli ergodic filtrations, this notion is the same as independence, that is, a standard filtration is the ‘past’ of a Bernoulli scheme. For general locally finite filtrations, this notion is more complicated, and we give its definition and a criterion for testing it. In particular, we state the standardness criterion in terms of random processes, namely, as a strengthening of the martingale convergence theorem. This is a theorem on diagonal convergence in a certain sense. There is a scale of strengthenings which we interpret as a scale of zero–one laws: at one end of the scale we have Kolmogorov’s zero–one law, and at the other end standardness or (in the homogeneous case) the Bernoulli property. These questions are discussed in §5. It should be said at once that in the case of general filtrations one can define different degrees of ‘non-standardness’, or different degrees of closeness to standard filtrations. Most interesting is non-standardness in the homogeneous (for example, dyadic) case. In §6 we analyze several key examples, in particular, a scheme of random walks in random scenery. But this work is still in progress, so this section is of a preliminary
nature. Finally, §7 deals with filtrations in Borel spaces and the problem of enumerating all measures with a given equipment. This subject has broad applications, and we describe its connections with the theory of random walks on groups and the theory of traces of algebras and characters of groups.

I considered it useful to supplement the survey with several historical remarks on the interesting history of the theory of filtrations.

2. The definition of filtrations in measure spaces

2.1. Measure theory: a Lebesgue space, basic facts. We consider a certain category of measure spaces, the category of Lebesgue spaces in the sense of Rokhlin [45], or standard measure spaces. A Lebesgue space is a space endowed with a separable \( \sigma \)-algebra of sets that is complete with respect to some (and hence every) basis, and a \( \sigma \)-additive probability measure is given on the \( \sigma \)-algebra. Separability of a \( \sigma \)-algebra means the existence of a countable basis of measurable sets that separates the points of the space, and the definition of a basis includes the condition that an arbitrary measurable set \( A \) should be contained in some set \( B \) in the Borel hull of the basis with \( \mu(B \setminus A) = 0 \) (the weaker condition \( \mu(B \triangle A) = 0 \) is not sufficient, but we do not dwell on this here; see [45]).

All objects and notions, such as spaces, bases, completeness, sets, functions, partitions, maps, and so on, should be understood \( \text{mod } 0 \), that is, up to modifications on sets of measure zero. A Lebesgue space is a class of Lebesgue spaces coinciding \( \text{mod } 0 \), a measurable function is a class of functions coinciding \( \text{mod } 0 \), and so on.

This convention makes it necessary to check that all notions, definitions, statements, and proofs are correct (well defined) with respect to passage \( \text{mod } 0 \), a necessity usually ignored by all authors, though experienced authors do not get into trouble because of this neglect.

Morphisms in the category of Lebesgue spaces are measurable (in the obvious sense) measure-preserving maps, or more precisely, classes of maps coinciding \( \text{mod } 0 \) (in the image and inverse image). In more detail, a measurable map of a Lebesgue space to an arbitrary ‘measurable’ space, that is, a space endowed with a \( \sigma \)-algebra (for instance, to a standard Borel space with countable base: the interval \( [0,1] \)), is a map such that the inverse image of a measurable set is measurable. If the image space is also endowed with a measure and the measure of an image is equal to the measure of the inverse image, then the map is said to be measure-preserving. If the same is true for the inverse map, then it is called an isomorphism (an automorphism in the case where the image and the inverse image coincide) of measure spaces.

Note that the image of a Lebesgue space under a measurable map is a Lebesgue space (cf. the corresponding theorem for compact sets).

Two objects corresponding to each other under an isomorphism and the inverse isomorphism are said to be isomorphic (and sometimes, equivalent).

The following classification theorem due to Rokhlin refined an earlier theorem of von Neumann [96] which was of the same type, but was based on a more complicated axiomatics of measure spaces.

*There exists a unique, up to isomorphism, Lebesgue space with a continuous (that is, having no atoms) probability measure.* The interval \([0,1]\) endowed with Lebesgue
measure is a universal example of a Lebesgue space with a continuous measure. Any separable compact space with a continuous Borel probability measure is another such example. The most convenient universal example of a Lebesgue space is the countable product of two-point spaces, that is, the dyadic compactum, with an arbitrary continuous Borel measure. Hence, in the category of Lebesgue spaces, the most interesting objects (with a continuous measure) are isomorphic.

Note that the construction of the theory of Lebesgue spaces in the sense of Rokhlin in [45] essentially reproduces the theory of metric compacta and, even more specifically, of the dyadic compactum, with necessary modifications taking into account the specifics of measure theory.

Passing to a classification of general Lebesgue spaces, that is, those with an arbitrary measure, consists in adding at most countably many atoms of positive measure.

Therefore, a complete (metric) invariant of a general Lebesgue space is an ordered sequence of non-negative numbers

$$m_1 \geq m_2 \geq m_3 \geq \cdots \geq 0, \quad \sum_i m_i \leq 1,$$

namely, the sequence of sizes of all atoms, that is, the points of positive measure.

The zero sequence corresponds to a space with a continuous measure, and a sequence with unit sum, $\sum_i m_i = 1$, to a space with an atomic measure.

2.2. Measurable partitions, filtrations. A measurable partition $\xi$ (see Fig. 1) of a Lebesgue space $(X, \mu)$ is the partition of $X$ into the inverse images of points under a measurable map $f: X \to R$, where $R$ is an arbitrary standard Borel space, for instance (without loss of generality), the interval $[0, 1]$. The fact that a set $C$ is an element (block) of a partition $\xi$ is denoted as follows: $C \in \xi$. Partitions are also regarded mod 0.

![Figure 1. A measurable partition.](image)

The following is a fundamental fact which goes back to works on probability theory from the 1930s (Doob and others), but was rigorously proved by Rokhlin at the initiative of Kolmogorov (some hints about the need for a rigorous study of conditional measures are contained in his famous book [33]).

**Theorem 2** (Rokhlin [45]). Each measurable partition $\xi$ of a Lebesgue space $(X, \mu)$ can be equipped with a unique mod 0 canonical system of measures $\{\mu^C\}, C \in \xi$ (in
a more popular language, a system of conditional measures), on almost all elements of $\xi$ in such a way that:

1) almost all (with respect to the measure $\mu$) elements $(C, \mu_C)$ endowed with the conditional measures are Lebesgue spaces, and the quotient space $X/\xi$ endowed with the quotient measure $\mu_\xi$ is also a Lebesgue space;

2) for every measurable set $A \subset X$ the function $\mu_C(A \cap C)$ is measurable (as a function on $(X/\xi, \mu_\xi)$), and

$$
\mu(A) = \int_{X/\xi} \mu_C(A \cap C) \, d\mu_\xi(C)
$$

(Fubini, iterated integration formula).

Recall that a measurable partition $\xi$ gives rise to a unique $\sigma$-algebra $\mathcal{A}_\xi$ of sets measurable with respect to $\xi$, that is, sets composed of elements of $\xi$. In Lebesgue spaces the converse is also true: every $\sigma$-subalgebra $\mathcal{A}$ of measurable sets in a space $(X, \mu)$ gives rise to a unique (mod 0) measurable partition $\xi$, and $\mathcal{A}_\xi = \mathcal{A}$.

In the spaces $L^1$ and $L^2$ we have the operator of conditional expectation corresponding to a measurable partition $\xi$, that is, the projection $P_\xi$ (in $L^2$, the orthogonal projection) to the subspace of functions constant on elements of $\xi$; this is the operator of averaging over the elements of $\xi$ endowed with the conditional measures.

From an alternative point of view, the theorem on conditional measures is simply a theorem on an integral representation of the operator of conditional expectation:

$$(P_\xi f)(C) = \int_C f(x) \, d\mu_C(x).$$

The left-hand side can be understood as a function on the quotient space, but also as a function on $X$ belonging to the subalgebra of functions constant on elements of $\xi$.

The projection to this subalgebra is well defined, and the proof of the existence of an integral representation usually proceeds as follows: first one endows the space $X$ with a compact topology, or even a metric, then one considers the projection to the space of continuous functions, and finally, one approximates functions in $L^1(X, \mu)$ by continuous functions.

Introducing a partial order on the classes of partitions coinciding mod 0 turns the collection of partitions into a lattice, where $\xi \succ \xi'$ means that elements of $\xi$ are finer than elements of $\xi'$, that is, an element of $\xi'$ is composed of elements of $\xi$; it is this order that agrees with the inclusion order on the spaces of measurable functions associated with partitions. The largest partition is the partition into singletons mod 0, and the smallest is the trivial partition (with only one non-empty component). Note that this order is opposite to that adopted in combinatorics, where the largest partition is the trivial one and the smallest is the partition into singletons.

### 2.3. Classification of measurable partitions.

The purpose of the upcoming sections is to describe metric invariants (that is, give a classification) of partitions, and then of finite filtrations, that is, finite decreasing sequences of measurable partitions. The case of one partition is investigated in the well-known paper [45],
and we cite the corresponding result (with a slightly modified statement) only for convenience of further generalizations. We remark that this is perhaps the main original result of [45].

**Theorem 3** (Rokhlin [45]). A metric invariant of a measurable partition \( \xi \) in a continuous Lebesgue space \((X, \mu)\) is a metric invariant of the map

\[
V_\xi : (X/\xi, \mu_\xi) \to \Sigma,
\]

where \((X/\xi, \mu_\xi)\) is the quotient space, \(\mu_\xi\) is the quotient of the measure \(\mu\) under the projection \(X \to X/\xi\), and the map \(V_\xi\) sends every element \(C \in \xi\) to the ordered sequence of measures of the atoms \(m_1(C) \geq m_2(C) \geq \cdots\), regarded as a point of the simplex \(\Sigma\) of all ordered series of non-negative numbers with sum at most 1.

Thus, two measurable partitions of Lebesgue spaces are isomorphic, that is, correspond to each other under an isomorphism of measure spaces, if and only if the corresponding maps \(V\) are metrically isomorphic.

But we can go further and ask about the metric invariants of the map \(V_\xi\). The answer is as follows: these invariants are the \(V_\xi\)-image of the measure \(\mu_\xi\), which is a Borel measure on the simplex \(\Sigma\), and the so-called multiplicities (we will not need them, so we refer the interested reader to the paper [70], where the invariants of measurable maps are described in detail, in the generalization of Rokhlin’s paper [46] on classification of functions).

Speaking rather loosely, we can say that a measurable partition is a random measure space whose realizations are all regarded as disjoint subsets in the same space.

Before turning to the case of several partitions, we describe a number of classes and examples of partitions.

1. **Partitions that have independent complements.** Assume that almost all elements \(C\) of a partition \(\xi\) are isomorphic as spaces endowed with the conditional measures \(\mu^C\), that is, the \(V_\xi\)-image of the measure \(\mu_\xi\) is the \(\delta\)-measure at the point of \(\Sigma\) corresponding to the (common) type of the conditional measures \(\mu^C\). In this and only this case the partition \(\xi\) has an **independent complement** \(\xi^-\), and the space \((X, \mu)\) can be decomposed into a direct product of measure spaces:

\[
(X, \mu) = (X/\xi, \mu_\xi) \times (X/\xi^-, \mu_{\xi^-}),
\]

where the partition \(\xi^-\) is the unique (up to isomorphism, but not geometrically) independent complement of \(\xi\). Almost every element of \(\xi\) has a non-empty intersection with almost every element of \(\xi^-\), and the independence condition is satisfied:

\[
\mu(A \cap B) = \mu_\xi(A) \mu_{\xi^-}(B)
\]

for any sets \(A\) and \(B\) measurable with respect to \(\xi\) and \(\xi^-\), respectively.

**Corollary 1** (Rokhlin [45]). A Lebesgue space with a continuous measure has a partition with continuous conditional measures that is unique up to isomorphism.

This is the partition of the square \([0, 1]^2\) endowed with Lebesgue measure into vertical intervals. A (non-unique) independent complement is the partition into horizontal intervals.
2. *Finite partitions.* Partitions with finitely many elements (blocks) are said to be finite; the complete invariant of finite partitions, in the case where \((X, \mu)\) is a continuous measure space, is the collection of measures of the elements; in this case, \(V_\xi\mu = \delta_0\), where \(0 \in \Sigma\) is the zero sequence.

3. *Atomic partitions.* Let \(\xi\) be a partition with almost all elements finite, more precisely, almost all measures \(\mu^C\) are atomic measures with finitely many atoms; such partitions are said to be discrete (though it would be more appropriate to call them 'cofinite'). A measurable partition \(\xi\) of a Lebesgue space \((X, \mu)\) with a continuous measure will be said to be *atomic* if the number of metric types of conditional measures is finite; in other words, if the \(V_\xi\)-image of \(\mu^{\xi}\) is a measure on \(\Sigma\) with finitely many atoms.

4. A *semihomogeneous partition* is an atomic partition with finite blocks and uniform conditional measures on almost every block. The invariant of a semihomogeneous partition is the collection of positive integers equal to the numbers of points in the blocks and the measures of the sets of all blocks with a given number of points. A homogeneous (dyadic, triadic, and so on) partition is a partition with the same number of points in all blocks; in this case, the \(V_\xi\)-image of \(\mu^{\xi}\) is the \(\delta\)-measure at the point 

\[
(1/n, \ldots (n) \ldots, 1/n) \in \Sigma.
\]

2.4. *Classification of finite filtrations.* The study and classification of (finite or infinite) collections of \(\sigma\)-subalgebras, or, which is equivalent, measurable partitions, in Lebesgue spaces with continuous measures is perhaps the most general and difficult geometric problem of measure theory and probability theory. Already the case of two measurable partitions in general position is of great interest and involves combinatorial and algebraic difficulties; we will consider it elsewhere.

In this section we first give a classification of finite monotone sequences of measurable partitions, that is, finite filtrations. It is quite simple. However, the main subject of the paper is the analysis of infinite filtrations.

The study of finite filtrations in a space with a continuous measure should be preceded by the study of such filtrations in the following finite measure space:

\[
(C, m), \quad C = \{c_1, c_2, \ldots, c_k\}, \quad m(c_i) = m_i, \quad i = 1, \ldots, k.
\]

A filtration of length \(n\) in a finite space is a hierarchy: \(\xi_0\) is the partition of \(C\) into singletons; \(\xi_1\) is a partition of \(C\); \(\xi_2\) combines some blocks of \(\xi_1\) into larger blocks, and so on; the last partition \(\xi_{n-1}\) consists of a single block, which is the whole space \(C\). It is convenient to think of a hierarchy as a tree of rank \(n\), that is, with \(n\) levels, endowed with a measure (see Fig. 2).

For \(n = 1\), the tree has two levels: several vertices at the zero level and a single vertex at the first level. The vertices at the zero level are linearly ordered according to the measures of the points.

In the general case the vertices at the \(r\)th level are the elements of the partition \(\xi_r\), \(r = 0, 1, \ldots, n - 1\), and an edge joining a vertex at level \(r - 1\) with a vertex at level \(r\) means an inclusion between the corresponding elements of partitions. The vertices at the zero level, that is, the points of the original space, are ‘leaves’ of the tree. They have the measures \(m_i\), \(i = 1, \ldots, k\); let us introduce a partial linear order on the leaves belonging to the same element of the partition \(\xi_1\) according to the size of
their measures (up to an arbitrary order on leaves of equal measure). Given these
measures, one can recover, by a simple summation, the conditional measures on
the vertices at the first level, that is, on the elements of the partition \( \xi_1 \), and then
continue in the same way for the elements of \( \xi_2, \ldots, \xi_{n-1} \) at the subsequent levels.

**Definition 1.** Denote by \( \text{Tree}_n \) the space of finite single-root trees of rank \( n \) (that
is, trees with \( n \) levels labeled by the numbers \( 0, 1, \ldots, n-1 \) and a unique vertex
at the last level as a root) endowed with a strictly positive probability measure on
the leaves, that is, on the vertices at the zero level, and hence on the vertices at
all levels; the values of the measure induce a linear order on the vertices up to an
arbitrary order of vertices with equal measure.

One may say that \( \text{Tree}_n \) is the space of finite filtrations of length \( n \) on finite
measure spaces.

In what follows, we will equip these trees with another important structure,
a metric, and the resulting category of trees endowed with a measure and a metric
will be a powerful tool in solving our problems.

Now we are ready to describe a complete system of invariants of a finite filtration
of an arbitrary Lebesgue space. We consider only filtrations in which all measurable
partitions are atomic:

\[
\{\xi_0 \succ \xi_1 \succ \xi_2 \succ \cdots \succ \xi_n\} = \{\xi_k\}_{k=1}^n,
\]

where \( \xi_0 = \varepsilon \) is the partition into singletons mod 0.

The fact that such a sequence is decreasing means that almost every element \( C_n \)
of the last partition \( \xi_n \) is finite and composed of finitely many elements of the
partition \( \xi_{n-1} \), that is,

\[
C_n = \bigcup C_{n-1}^i, \quad C_{n-1}^i \in \xi_{n-1},
\]

the elements \( C_{n-1}^i \) are unions of elements of \( \xi_{n-2} \), and so on.

In other words, almost every element \( C_n \in \xi_n \) supports a hierarchy, that is,
a finite filtration with a measure. Hence, we have a homomorphism

\[
t_n: X/\xi_n \rightarrow \text{Tree}_n,
\]
which sends almost every element $C_n \in \xi_n$, endowed with the conditional measures of its points, to the measured tree corresponding to the finite filtration on this element induced by the filtration on the whole space. Denote the $t_n$-image of the measure $\mu_{\xi_n}$ on $X/\xi_n$ by $\omega_\tau$.

**Theorem 4.** The complete system of isomorphism invariants of a finite filtration $\tau \equiv \{\xi_k\}_{k=0}^n$ of atomic partitions of a Lebesgue space with a continuous measure coincides with the system of invariants of the homomorphism $t_n$ to the space $\text{Tree}_n$ of measured trees, which consists of the measure $\omega_\tau$ on $\text{Tree}_n$ obtained as the $t_n$-image of the measure $\mu_{\xi_n}$ on $X/\xi_n$ and the multiplicity function. If the filtration has finitely many types of conditional measures, then the complete invariant is simply the atomic measure $\omega_\tau$ on the set $\text{Tree}_{n+1}$ of trees of rank $n + 1$ (all multiplicities are continual, where the latter term refers to the cardinality of the continuum).

In the case $n = 1$ (that is, the case of one atomic partition $\xi_1$), this is a measure on ordered probability vectors, in accordance with Theorem 3.

The proof that these are invariants is obvious, since the construction used only invariants of partitions, namely, systems of conditional measures. The completeness follows from the fact that two non-isomorphic filtrations must differ at some set of positive measure that is measurable with respect to the partition $\xi_n$, and hence the images of the corresponding measures on $\text{Tree}_{n+1}$ do not coincide. Repeating what has been said in the case of one partition, one may think of the invariant of a finite filtration as a random tree endowed with a measure.

It is not difficult to extend this procedure to classification of finite filtrations with arbitrary measurable (rather than only atomic) partitions: one should pass from finite trees to countable, or even continual, trees and general measures on them.

Earlier, classification of finite filtrations was considered in different terms (see, for example, [21]); the language of trees or hierarchies proposed above was not used. The classification itself is not very interesting; it could be useful if one could apply it to classification of infinite filtrations. But one cannot simply construct the inductive limit of the spaces $\text{Tree}_n$ with respect to the natural embeddings and consider measures on the resulting space as invariants of filtrations, and a direct attempt to pass to the limit is of no use for classifying infinite filtrations. The reason is that when we pass to the limit, the bijectivity of the correspondence between classes of filtrations and finite invariants is lost. In fact, the information connected with finite filtrations that is useful for classifying infinite filtrations is somewhat different: finite fragments of a filtration should be considered not by themselves, as we did above, but together with some functions or metrics, as we will do below.

### 2.5. Filtrations we consider and how one can define them.

#### 2.5.1. Classes and properties of filtrations.

First of all, we mention the most interesting class of filtrations (to be considered in more detail in another paper). This is the class of filtrations in which all partitions have purely continuous conditional measures. It is of great importance for the theory of random processes and the theory of $C^*$-algebras, but it is not related to combinatorial problems.
The theory of filtrations of subalgebras

The key notion is that of ergodicity of filtrations, also called the Kolmogorov property, regularity, and other terms.

A filtration

$$\tau = \{A_n\} \simeq \{\xi_n\} \simeq \{L^\infty(\xi_n)\}$$

is said to be **ergodic** if

- the intersection of the $\sigma$-algebras $A_n$ is the trivial $\sigma$-algebra $\mathcal{N}$

$$\bigcap_n A_n = \mathcal{N}, \quad \text{or}$$

- the measurable intersection of the partitions $\xi_n$ is the trivial partition

$$\bigwedge_n \xi_n = \nu, \quad \text{or}$$

- the intersection of the algebras of functions constant on elements of the partitions $\xi_n$ over all $n$ consists of the constants

$$\bigcap_n L^\infty(X/\xi_n) = \{\text{Const}\}.$$

In our further considerations we deal with filtrations in which all partitions have atomic conditional measures without continuous components (filtrations with continuous conditional measures we leave for another paper).

We distinguish the following classes: **locally finite filtrations**, **homogeneous filtrations**, and **semihomogeneous filtrations**.

**Definition 2.** (i) A filtration $\tau = \{\xi_n\}_{n=0}^\infty$ of a Lebesgue space $(X, \mu)$ with a continuous measure $\mu$ is said to be **locally finite** if all the partitions $\xi_n$, $n = 0, 1, 2, \ldots$, are atomic, that is, almost all conditional measures on the elements of $\xi_n$ are atomic, and the number of types of conditional measures is finite, in particular, for every $n$ the number of atoms in almost all elements of $\xi_n$ is finite (but possibly depends on $n$).

(ii) A locally finite filtration is said to be **semihomogeneous** if the conditional measures of almost all elements of all partitions $\xi_n$, $n = 0, 1, 2, \ldots$, are uniform distributions (that is, the measures of all the points of a given element are equal).\(^1\)

(iii) A locally finite filtration is said to be **homogeneous** if for every $n$ the conditional measures of almost all elements of $\xi_n$ are uniform and equal. If the number of elements in $\xi_n$ is $r_n$, then the filtration is said to be $\{r_n\}_n$-adic; for $r_n \equiv 2$ it is dyadic, and for $r_n \equiv 3$ it is triadic.

Relaxing the finiteness condition in the definition of a locally finite filtration (for example, allowing elements of partitions to have countably many atoms or an arbitrary number of types of conditional measures) does not raise serious new difficulties and essentially new phenomena in the theory of filtrations, and thus we restrict ourselves only to locally finite filtrations.

We consider also the following special classes.

\(^1\)But the conditional measures of elements of the quotient partitions $\xi_n/\xi_{n-1}$ are not necessarily uniform.
Definition 3. We say that a filtration $\tau$ is:

(a) a Bernoulli filtration if it is the tail filtration for an arbitrary sequence of independent random variables, in other words, if there exists a sequence of independent partitions $\eta_n, n = 1, 2, \ldots$, in a Lebesgue space with a continuous measure such that

$$\xi_n = \bigvee_{k=n}^{\infty} \eta_k, \quad n = 0, 1, 2, \ldots$$

(b) more generally, a Markov filtration if it is the tail filtration of a one-sided Markov chain with discrete time and arbitrary state space.

Let us turn our attention to the following important class of Markov filtrations which generalizes random processes with independent increments.

Consider a group $G$ and a finitely supported probability measure $m$ on $G$. The Markov chain $y_n, n = 0, 1, 2, \ldots$, corresponding to the (say, left) random walk on $G$ with transition probability $m$ generates a Markov filtration. The ergodicity of this filtration is equivalent to the triviality of the Poisson–Furstenberg boundary. And the triviality criterion for the boundary is that the entropy vanishes (see [26]).

More generally, assume that the group $G$ acts on a Lebesgue space $(Z, \nu)$ with a continuous (or even $\sigma$-finite) invariant measure $\nu$:

$$T_g \nu = \nu \quad \forall g.$$ 

This gives rise to a one-sided Markov process $\{z_n\}_n$ with the state space $(Z, \nu)$ and transition probabilities

$$\text{Prob}(z|u) = m(g), \quad \text{where } z = T_g u,$$

and to the corresponding filtration, which will be considered in the upcoming sections. If $m$ is the uniform measure on the generators of the group, then the filtration is homogeneous. Note that the state space of this process is continual, but the filtration is locally finite.

As we will see in the next section, in this case the filtration can be defined by a sequence of finite partitions

$$\eta_n, \quad n = 1, 2, \ldots, \quad \bigwedge_n \eta_n = \epsilon$$

(where $\epsilon$ is the partition into singletons). The filtration $\{\xi_n\}_n$ is defined as follows:

$$\xi_n = \bigvee_{k=n+1}^{\infty} \eta_k, \quad n = 0, 1, 2, \ldots$$

2.5.2. The basis method of defining filtrations. The most efficient method of defining filtrations, in both measure spaces and Borel spaces, is to introduce first an increasing sequence of finite partitions (a basis), and then define a filtration as the sequence of products of ‘tails’. We will call it the basis method.

Consider a sequence of finite partitions (a basis)

$$\{\eta_n\}_{n=1}^{\infty}, \quad \bigwedge_n \eta_n = \epsilon$$
(where $\epsilon$ is the partition into singletons), of a measure space or Borel space. A filtration $\{\xi_n\}_n$ is defined as follows:

$$\xi_n = \bigvee_{k=n+1}^{\infty} \eta_k, \quad n = 0, 1, 2, \ldots.$$ 

The basis method of defining filtrations arises naturally in connection with random processes with time $\mathbb{Z}_+$ and in other probabilistic situations (for example, $\eta_n$ is the partition corresponding to the $\sigma$-algebra of sets determined by the state of the process at time $n$). We will use it systematically in §3 when defining tail filtrations on path spaces of graphs.

We emphasize that geometrically one and the same filtration can be defined via different bases, and the problem of (metric or Borel) isomorphism of filtrations defined by the basis method is quite non-trivial. Moreover, the conditional measures of the partitions $\xi_n$ (that is, the equipments; see §3.3) in the case of measure spaces can be defined only by passing to the limit.

The relationship between two sequences of partitions, the increasing one (a basis)

$$\{\zeta_n\}_n : \quad \zeta_n = \bigvee_{k=1}^{n} \eta_k, \quad n = 1, 2, \ldots,$$

and the decreasing one (a filtration)

$$\{\xi_n\}_n : \quad \xi_n = \bigvee_{k=n+1}^{\infty} \eta_k, \quad n = 0, 1, 2, \ldots,$$

which look like dual objects, is by no means symmetric: the theory of filtrations is much deeper than the theory of increasing sequences, and the most important difference is in the asymptotic behaviour. Namely, passing to the limit in the theory of filtrations is much more refined than in the case of increasing sequences. In §8 we show an example from Kolmogorov’s paper with a commentary by Rokhlin, which demonstrates the absence of continuity of the product when passing to the limit in a filtration. This fact distinguishes filtrations of subalgebras of functions from filtrations of subspaces of a Hilbert space, where there is a complete symmetry between increasing and decreasing sequences. The reason, of course, is that the lattice of subalgebras (partitions), in contrast to the lattice of subspaces, has no canonical involution similar to passing to the orthogonal complement, which sends increasing sequences to decreasing ones and vice versa.

2.6. The equivalence relation associated with a filtration, cocycles and associated measures. Along with the canonical intersection of $\sigma$-algebras, which turns out to be trivial for ergodic filtrations, there is another set-theoretic intersection, and the corresponding well-defined equivalence relation on the measure space itself. Assume that we have a locally finite filtration (or even just a filtration with discrete elements) $\tau = \{\xi_n\}_n$. Consider the monotone limit of the measurable partitions $\xi_n$; denote it by $\bigcap_n \xi_n = \xi_{\tau}$ (as opposed to the measurable intersection $\bigwedge \xi_n$).
Proposition 1. The partition $\bigcap_n \xi_n$ is well defined as the monotone limit of the decreasing sequence of partitions $\{\xi_n\}_n$. It is not measurable, since in the ergodic case there are no non-constant measurable functions constant on elements of this partition. Nevertheless, there is a canonical object corresponding to this partition in a function space (for example, in $L^2(X, \mu)$):

$$\mathcal{C} = \left\{ f \in L^2(X, \mu) : \exists n, \int_C f(x) \, d\mu_C(x) = 0 \text{ for a.e. } C \in \xi_n \right\}. \quad (2.1)$$

Clearly, the integral in (2.1) vanishes also for all $m > n$. The linear space $\mathcal{C} \subset L^2(X, \mu)$ is not closed, and its closure is the orthogonal complement of the intersection $\bigcap_n L^\infty(X/\xi_n, \mu_{\xi_n})$ (in the ergodic case, the closure is the space of all functions with zero integral).

Proof. 1. If we modify all the partitions $\xi_n$, $n = 1, 2, \ldots$, on a set of measure zero, then $\xi_\tau$ will also change only on a set of measure zero (it is important here that the partitions $\xi_n$ decrease and that their elements are finite).

2. The space $\mathcal{C}$ is uniquely determined by the filtration and does not change under automorphisms of this filtration. $\square$

The space $\mathcal{C}$ can be extended to an invariant functional analog of the intersection $\xi_\tau$ itself rather than the filtration. To this end, in the definition of $\mathcal{C}$ one should take functions $f$ for which the zero integral condition is satisfied for almost all elements $C$ of arbitrary measurable partitions greater than the intersection: $\xi \succ \xi_\tau$. This condition no longer changes when a filtration is replaced by another filtration having the same intersection of partitions. This space is a functional analog of a tame partition.\footnote{This observation was made by the author in [88], [83]. Clearly, classification of tame partitions is equivalent to classification of non-closed subspaces of the above form with respect to automorphisms, that is, unitary real multiplicative operators.}

The most important property of the intersection of the partitions is that its (countable) elements support a projective conditional measure, namely, for almost all pairs of points $x, y$ of its elements, the ratio of the conditional measures $\frac{\mu_C(x)}{\mu_C(y)}$, where $C$ is an element of an arbitrary measurable partition $\xi \succ \bar{\xi}$, does not depend on the choice of $\xi$. This ratio is a well-defined cocycle, which will be discussed in more detail in the next section (see §3.5). In the case where $\xi$ is the orbit partition for a transformation with a quasi-invariant measure (or for an amenable group), this cocycle is called the Radon–Nikodym cocycle. See also [94].

The paradoxical role of the cocycle defined on the intersection of measurable partitions (on the tail equivalence relation) and constructed from a given measure $\mu$ on $X$ (in the next section, a measure is constructed on paths in a graph) is that it can simultaneously be a cocycle for other measures singular with respect to $\mu$. Thus, it turns out to be related to another notion mod 0 and, consequently, to other classes of functions coinciding mod 0. In other terms this paradox can be stated as follows: let $\tau = \{\xi_n\}_n$ be a filtration; then the conditional measures of the partitions $\xi_n$ for all $n$ do not uniquely determine the measure $\mu$, namely, there can also exist other associated measures that are singular with respect to $\mu$ but have the same conditional measures with respect to the partitions $\xi_n$.\footnote{This observation was made by the author in [88], [83]. Clearly, classification of tame partitions is equivalent to classification of non-closed subspaces of the above form with respect to automorphisms, that is, unitary real multiplicative operators.}
The theory of filtrations of subalgebras

The notion of cocycle and the problem of describing all measures with a given cocycle emerged in the theory of Markov processes (works of Dynkin [9]–[11]), in ergodic theory (Schmidt [52]), and in the theory of graded graphs (papers of the author and Kerov (see [86]) and their followers). The problem of constructing all measures with a given cocycle is a far-reaching generalization of the traditional problem of describing all invariant measures for group actions. For more details, see [82] and §6.

3. Tail filtrations in graded graphs and Markov chains

In this section we introduce a new circle of notions, which brings to the theory of filtrations a substantially different point of view on basic notions and provides it with an huge supply of examples. We use \( \mathbb{N} \)-graded graphs and multigraphs (Bratteli diagrams); the theory of such graphs and their connections with \( C^\ast \)-algebras were intensively studied in the 1970s to 80s (by O. Bratteli, G. Elliott, E. Effros, D. Handelman, and others; see [18], [86] and the references therein). New connections of this theory with dynamics and ergodic theory were discovered by the author in the 1970s and became a source of new problems in both representation theory and dynamics; in part, they became the subject of asymptotic representation theory developed in the same years by the author together with Kerov and their followers.

In this survey and several previous papers by the author, attention is drawn, apparently for the first time, to the role of filtrations of \( \sigma \)-algebras and filtrations of subalgebras of \( C^\ast \)-algebras in the analysis of the structure of the algebras themselves. But here we are primarily interested in the measure-theoretic structure of filtrations appearing in connection with graded graphs.

3.1. Path spaces of graded graphs (Bratteli diagrams) and Markov compacta. Consider a locally finite, infinite, \( \mathbb{N} \)-graded graph \( \Gamma \) (in other words, a Bratteli diagram). The set of vertices of order \( n \), \( n = 0, 1, 2, \ldots \), will be denoted by \( \Gamma_n \) and called the \( n \)th level of the graph, and

\[
\Gamma = \coprod_{n \in \mathbb{N}} \Gamma_n,
\]

where \( \Gamma_0 \) consists of the unique initial vertex \( \{\emptyset\} \). We assume that every vertex has at least one successor, and every vertex except the initial one has at least one predecessor. In what follows, we also assume that edges of \( \Gamma \) are simple.\(^3\)

A graded graph \( \Gamma \) gives rise canonically to a locally semisimple algebra \( A(\Gamma) \) over \( \mathbb{C} \), however, here we consider neither this algebra nor the relationship of the notions introduced below with this algebra and its representations.

A path is a (finite or infinite) sequence of adjacent edges of the graph (for graphs without multiple edges, this is the same as a finite or infinite sequence of adjacent vertices). The number of paths from the initial vertex to a given vertex is finite, which means exactly that the graph is locally finite. The space of all infinite paths

\(^3\)Considering multigraphs, that is, graphs with multiple edges, in general gives nothing new for our purposes, since the cotransition probabilities (equipments) introduced below replace and generalize the notion of multiplicities of edges. However, in what follows we use the language of multigraphs to simplify some statements.
in a graph $\Gamma$ is denoted by $T(\Gamma)$; in a natural sense, it is an inverse limit of the spaces $T_n(\Gamma)$ of finite paths from the initial vertex to the vertices at the $n$th level. Therefore, $T(\Gamma)$ is a Cantor-like compactum.

Cylinder sets in the space $T(\Gamma)$ are sets determined by conditions on initial segments of paths up to level $n$; these sets are closed and open and define a base of a topology. One can naturally define the notion of the tail equivalence relation $\tau_{\Gamma}$ on $T(\Gamma)$: two infinite paths are equivalent if they eventually coincide; we also say that these two paths lie in the same block of the tail partition.

The tail filtration (for the moment, a Borel one, without any measure)

$$\Xi(\Gamma) = \{A_0 \supset A_1 \supset \cdots\}$$

is the decreasing sequence of $\sigma$-algebras of Borel sets, where the $\sigma$-algebra $\mathcal{A}_n$, $n \in \mathbb{N}$, consists of all Borel subsets of $T(\Gamma)$ that together with every path $\gamma$ contain all paths coinciding with $\gamma$ up to level $n$. In a clear sense the $\sigma$-algebra $\mathcal{A}_n$ is complementary to the finite $\sigma$-algebra of cylinder sets of order $n$.

3.2. Trees of paths. Let us make a general remark on the path space of a graded graph: the paths joining a given vertex with the initial one form a tree whose root is the given vertex, and each path (a leaf of the tree) is endowed with a measure.

The connection between the tree constructed from the paths of a graph and that constructed from a measurable partition is obvious.

Definition 4. Every vertex $v$ at level $n$ in a graded graph determines in a natural way a tree of rank $n$, which is composed of the paths leading from the initial vertex to $v$. These trees agree in a natural sense: the tree corresponding to a vertex $u$ is embedded as a subtree in the tree corresponding to a vertex $v$ if $v$ succeeds $u$. Leaves of the tree correspond to paths in the graph.

Proposition 2. The measured trees constructed in §2.4 from finite fragments of an infinite filtration $\tau$ are exactly the trees constructed from the vertices of the graph $\Gamma(\tau)$. (The possibility of combining them into an inductive limit that determines the type of the filtration up to finite isomorphism was discussed above. However, this is not needed, since the type of the filtration is determined by the minimal model itself.)

Hence, an infinite path in a graded graph gives rise to an inductive limit of a sequence of nested trees corresponding to successive vertices of this path. The sequence of the roots of these trees has no limit, and the increasing sequence of nested sets of leaves of the trees constitutes the set of all infinite paths in the graph that are cofinal with the given path; the number of such paths, as well as the number of vertices at every level of the trees, goes to infinity. The measures on the leaves have no limit, but there arises a cocycle of ratios of measures. Therefore, we have obtained a rootless tree whose levels are all infinite. It is easier to view this tree as an infinite sequence of hierarchies on the infinite set of its leaves supporting the cocycle of ratios of conditional measures.

Of course, the construction depends on the choice of an infinite path. Here one should take into account whether the filtration is ergodic or not, but we will not need to study this problem.
At the same time, it is clear that this construction cannot be used to advance the classification of finite filtrations in §2.4 towards a classification of infinite filtrations, since it takes into account only finite invariants of a filtration, which do not suffice for this purpose. However, these data suffice for classifying filtrations up to finite isomorphism, which means the existence of a complete system of Borel invariants for such a classification.

3.3. An equipment of a multigraph and cotransition probabilities. We introduce an additional structure on a (multi)graph, a system of cotransition probabilities, or an equipment,

$$\Lambda = \{ \lambda = \lambda^u_v; u \in \Gamma_n, v \in \Gamma_{n+1}, (u, v) \in \text{edge}(\Gamma_n, \Gamma_{n+1}), n = 0, 1, 2, \ldots \},$$

by associating with each vertex $v \in \Gamma_n$ a probability vector whose component $\lambda^u_v$ is the probability of an edge $u \prec v$ that comes to $v$ from the previous level; thus, $\sum_{u: u \prec v} \lambda^u_v = 1$ and $\lambda^u_v \geq 0$. In the case of a multigraph, probabilities are assigned to each edge in the set of edges joining vertices $u, v$ such that $u \prec v$.

Definition 5. An equipped (multi)graph is a pair $(\Gamma, \Lambda)$ consisting of a (multi)graph $\Gamma$ and a system $\Lambda$ of cotransition probabilities on the edges of $\Gamma$. An equipment allows one to define the probability of a path leading from $\emptyset$ to a given vertex as the product of the cotransition probabilities over all edges constituting the path.

The most important special case of an equipment (that is, of a system $\Lambda$ of cotransition probabilities), which is called the canonical or central equipment and is studied in combinatorics, representation theory, and algebraic situations, is as follows:

$$\lambda^u_v = \frac{\dim(u)}{\sum_{u: u \prec v} \dim(u)},$$

where $\dim(u)$ is the number of paths from the initial vertex $\emptyset$ to a vertex $u$ (in representation-theoretic terms, this is the dimension of the representation space of $A(\Gamma)$ corresponding to the vertex $u$).

One can easily see that for the central equipment, the measure on the set of paths leading from $\emptyset$ to a given vertex is uniform, and the cotransition probability of reaching a vertex $v$ from a preceding vertex $u$ is proportional to the fraction of the paths from $\emptyset$ to $v$ that pass through $u$.

The canonicity of this system of cotransition probabilities lies in the fact that it is determined only by the graph itself. The corresponding Markov measures on the path space $T(\Gamma)$ are called central measures; up to now, the study of Bratteli diagrams was restricted to only these measures. In terms of the theory of $C^*$-algebras, central measures are traces on the algebra $A(\Gamma)$, and ergodic central measures are indecomposable traces. For more details on the case of central measures, see [77], §7 below, and the extensive bibliographies in the papers from our (incomplete) list of references.

Measures on the path space of a graph that agree with the canonical equipment are said to be central, and in the theory of stationary Markov chains they are called measures of maximal entropy.
3.4. The Markov compactum of paths. The term ‘cotransition probabilities’ is borrowed from the theory of Markov chains [9]: if we regard a graph, rotated by 90° (see Fig. 3), as the set of states of a Markov chain starting from the state $\emptyset$ at the time $t = 0$, and the numbers of the levels as moments of time, then $\Lambda = \{\lambda^u_v\}$ should be viewed as the system of cotransition probabilities for the Markov chain:
\[
\text{Prob}\{x_t = u \mid x_{t+1} = v\} = \lambda^u_v.
\]

We recall the definition of a Markov compactum in the generality we need.

**Definition 6.** Consider a sequence $\{X_n\}_{n=1}^{\infty}$ of finite sets, and associate with it a sequence of ‘multi-edges’, that is, a collection of matrices $M_n$ of size $(|X_n| =: d_n) \times (d_n + 1)$ in which an element $m_{u,v}$ with $u \in X_n$ and $v \in X_{n+1}$ is a non-negative integer equal to the number of edges joining $u$ and $v$; to each such edge we assign a non-zero number.

A path (trajectory) is a sequence of adjacent edges. If all the multiplicities $m_{u,v}$ are equal to zero or one, then a path is a sequence of adjacent vertices.

The space of all paths (trajectories), endowed with the natural topology of an inverse spectrum of sets of finite paths, is called a Markov (topological) compactum. A Markov compactum is said to be stationary if the state spaces and the matrices do not depend on $n$:
\[
X_n \equiv X_1, \quad M_n \equiv M_1, \quad n = 1, 2, \ldots;
\]
in this case we can consider not only the one-sided, but also the two-sided Markov compactum.

A Markov compactum with an equipment (or with cotransitions) is a Markov compactum in the sense of the above definition in which the entries $m_{u,v}$ of the matrix $M_n$ are vectors whose coordinates are positive numbers not exceeding 1 corresponding to the probabilities of edges going from $u \in X_n$ to $v \in X_{n+1}$, so that the sum of all the probabilities along a row (over all edges going to $v$) is equal to 1.

Sometimes a Markov compactum (either equipped or not) will be called a Markov chain. We are going to consider Markov measures on a (usually, non-stationary) Markov compactum.

One can see from the definition that the path space of a graded (multi)graph is a Markov compactum and that an equipment of the graph as defined above is
an equipment of the Markov compactum. In what follows, we do not distinguish between these two languages: the language of Markov compacta and that of path spaces in graded (multi)graphs. To begin with, these two (seemingly quite different) areas of mathematics, the theory of Markov processes and the theory of approximately finite-dimensional algebras (together with its combinatorial ‘insides’, the analysis of graded multigraphs), are in essence different versions of the same theory, and the multigraph corresponding to a Markov chain is exactly a Bratteli diagram. However, the settings of the problems are different, being algebraic in one case and probabilistic in the other, and the traditions of drawing graphs are different, too.

It may seem that the difference between these two seemingly dissimilar theories could be removed by rotating the figure by $90^\circ$. However, it is important to overcome this difference conceptually and to understand to what extent both areas, the theory of graded graphs and $\text{AF}$-algebras on the one hand, and the theory of Markov chains on the other hand, can be enriched by exchanging their dissimilar problems and intrinsic methods.

For example, I long ago raised perhaps paradoxical problems about the $K$-functor of a Markov chain and about ergodic $K$-theory, or about Shannon-type theorems for $\text{AF}$-algebras (see [87]).

We remark that the analysis of asymptotic properties of the sequence of matrices $M_n$ introduced above is also of great independent interest from the technical point of view. One can say that the whole theory of Markov compacta of paths in graded graphs we are considering here is a part of the asymptotic theory of infinite products of Markov matrices, which is important in itself and is connected with the theory of invariant measures, phase transitions, and other topics.

3.5. A cocycle of an equipped graph. The notion of an equipment is related to the notion of a cocycle on the path space of a graph. Consider a function $c(\cdot, \cdot, \cdot)$ on the set of pairs of cofinal (that is, eventually coinciding) paths on the graph with values in the set of non-negative real numbers. Assume that

$$c(t, s)c(s, t) = 1, \quad c(t, s) = c(t, z)c(z, s), \quad c(t, t) = 1$$

for all pairs of cofinal paths $(s, t)$. Sometimes instead of ‘2-cocycle’ one says simply ‘cocycle on an equivalence relation’. The definition makes sense for an arbitrary equivalence relation. A cocycle is said to be trivial, or cohomologous to 1, if there exists a non-negative function $g$ on paths such that

$$c(t, s) = \frac{g(t)}{g(s)}.$$ 

A simple example: for a measurable partition of a space, the function

$$c(t, s) = \frac{\mu^C(t)}{\mu^C(s)} \quad (3.1)$$

(the ratio of the conditional measures of points lying in the same element of the partition) is a trivial cocycle.
Lemma 1. For every locally finite filtration in a Lebesgue space with a continuous measure there is a well-defined cocycle on the tail equivalence relation.

Proof. Consider two paths $s, t$ lying in the same element of the partition $\xi_n$, and define $c(t, s)$ to be the ratio of their conditional measures (see (3.1)). As follows from the definition, conditional measures have the transitivity property, that is, the ratio does not change if we consider the conditional measures of a partition $\xi_m$ with $m > n$. □

Assume that we have an equipment of a graded graph $\Gamma$ (or a Markov compactum). The cocycle corresponding to the equipment is defined as follows: for two paths $t = (t_1, t_2, \ldots, t_n, \ldots), s = (s_1, s_2, \ldots, s_n, \ldots)$ with $t_i = s_i$ for $i \geq n + 1$, we have

$$c(t, s) = \prod_{i=1}^{n-1} \frac{\lambda_{t_{i+1}}}{\lambda_{s_{i+1}}}.$$  

Such a cocycle will be called a Markov cocycle. Of course, there exist non-Markov cocycles.

Every Borel probability measure defined on the path space of a graph (the space of trajectories of a Markov chain) determines a system of cotransition probabilities; namely, it is the system of conditional measures of the measurable partitions $\xi_n$. One says that a measure agrees with a given system $\Lambda$ of cotransition probabilities (with a given cocycle) if the collection of the corresponding cotransition probabilities for all vertices coincides with $\Lambda$.

Lemma 2. If the tail filtration on the path space of an equipped graph is ergodic with respect to a given measure (or simply the measure is ergodic), then the cocycle constructed from the equipment is non-trivial.

Indeed, the triviality of the cocycle would mean the existence of measurable functions constant on tail equivalence classes, which contradicts the ergodicity.

The canonical equipment gives rise to the cocycle identically equal to 1, and measures that agree with this cocycle are central, or invariant, measures.

The Markov property of central measures on an arbitrary graded graph follows automatically from their definition, as was observed in old papers by the author and Kerov [86], [30].

Recall that a system of cotransition probabilities does not in general allow one to uniquely determine a system of transition probabilities

$$\text{Prob}\{x_{t+1} = v \mid x_t = u\};$$

in other words, this system does not in general uniquely determine the Markov chain. On the other hand, the transition probabilities uniquely determine the Markov chain if an initial distribution is given. An analog of the initial distribution for the cotransition probabilities is the final distribution, that is, a measure on the boundary (see below), which is the collection of all ergodic measures with given cotransitions.
A measure on the path space of a graph will be said to be \textit{ergodic} if the tail $\sigma$-algebra (that is, the intersection of all the $\sigma$-algebras of the tail filtration) is trivial mod 0.

\textit{One of our purposes is to enumerate all ergodic Markov measures with a given system of cotransition probabilities. The list of such measures will be called the absolute boundary, or the absolute, of an equipped Markov compactum, or the absolute of a graded multigraph. This is a topological boundary, and, as we will see, the Choquet boundary of a certain simplex (a projective limit of finite-dimensional simplices).}

\textbf{3.6. Examples of graded graphs and Markov compacta.} The list of graded graphs and multigraphs includes a series of classical examples: the Pascal graph (already known to ancient Chinese mathematicians), which is the lattice $\mathbb{Z}_2^+$ graded along the diagonal (see Fig. 4), and its multidimensional generalizations $\mathbb{Z}_d^+$, $d > 2$, the Euler graph, the Young graph (see Fig. 5), and, more generally, Hasse diagrams of lattices, primarily distributive lattices. Among the more recent examples, we can mention the graph of unordered pairs, or the tower of dyadic measures ([69], [82]; see §7.3 below), the graph of ordered pairs [91], and so on, which appeared in connection with the theory of filtrations and the adic realization of ergodic automorphisms [67], [37], [89].

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{pascal_graph.png}
\caption{The Pascal graph.}
\end{figure}

A graph is said to be of dyadic type if the number of paths from the initial vertex to every vertex at the $n$th level is equal to $2^n$ for all $n = 1, 2, \ldots$. This class includes the graphs of ordered and unordered pairs, the minimal dyadic graph, and the minimal polyadic (Glimm) graph (see Fig. 6).

We emphasize once more that the path space in any graph can be regarded as the space of trajectories of a topological (in general, non-stationary) Markov chain. This space (of paths or trajectories), regarded as a Borel space, supports the tail filtration. The graph (chain) can be endowed with an equipment. And if we have a Borel measure (not necessarily Markov) on the space of paths (trajectories), then we obtain a \textit{locally finite} filtration on the resulting measure space. The local finiteness of this filtration follows from the local finiteness of the graph.
4. The Markov realization of locally finite filtrations

We have come to one of the main theorems of the paper, which connects the theory of filtrations with the combinatorics of Markov chains and graded graphs.

4.1. The theorem on a Markov realization.

Theorem 5. Every locally finite filtration is isomorphic to the tail filtration of a Markov chain, or to the tail filtration of the path space of a graded graph.\(^4\)

Proof. Let

\[ \tau = \{\mathcal{A}\}_{n=0}^{\infty} \simeq \{\xi_n\}_{n=0}^{\infty} \]

\(^4\)Here we more often associate the theory of filtrations with Markov chains rather than with graded graphs, because the term ‘Markov property’ and the language of Markov chains is better known and more widely used than the equivoluminous language of graded graphs. One of the purposes of this survey is to popularize the equivalence of these languages and to report the new things brought into the field by graded graphs.
be a locally finite filtration in a Lebesgue space \((X, \mathcal{A}_0, \mu)\) with a continuous measure \(\mu\). Starting from \(\tau\), we will construct a Markov chain with finite state spaces that has tail filtration isomorphic to \(\tau\).

The finite partitions introduced below are labeled, that is, their elements are assigned labels (for example, positive integers). The (ordered) product of two or more labeled partitions is again a partition of the same type, whose elements are labeled by the ordered tuples of labels of the factors. In what follows, it is convenient to assume that for finite partitions \(\eta \succ \eta'\) the set of labels of the larger partition \(\eta\) includes the set of labels of the smaller partition \(\eta'\).

Denote by \(\phi = \{\xi_n\}_n\) the sequence of measurable partitions corresponding to a filtration \(\mathcal{F}\) of \(\sigma\)-algebras. Choose a basis of the \(\sigma\)-algebra \(\mathcal{A}_0\), that is, an arbitrary increasing sequence \(\{\eta_n\}_{n=1}^{\infty}\) of finite labeled partitions that tends to the partition into singletons of the space \((X, \mu)\).

We start an inductive construction with the base case, namely, with the definition of the first state space of the future Markov chain. For this, we first choose an arbitrary partition complementary to \(\xi_1\), that is, a labeled measurable partition \(\zeta_1\) such that almost every element of \(\zeta_1\) intersects almost every element of \(\xi_1\) in at most one point (this can be done, since the elements of \(\xi_1\) are finite, and hence \(\zeta_1\) can be chosen to be finite).

Consider the product of partitions \(\sigma_1 = \eta_1 \lor \zeta_1\); it is a finite partition. We declare the set of its elements, that is, the quotient space \(Y_1 = X/\sigma_1\) regarded as a measure space, to be the set of labeled states of the Markov chain at the time \(t = 1\).

The induction step is a construction of the set \(Y_n\) of states at the time \(t = n\) provided that we have already constructed a finite Markov chain of length \(n - 1\) as the quotient of the space \((X, \mu)\) by the finite partition \(\bigvee_{i=1}^{n-1} \sigma_i\). We construct a partition \(\sigma_n\), consider the quotient space \(X/\xi_{n-1}\), and define on this space a finite labeled partition \(\zeta_n\) complementary (in the sense described above) to the partition \(\xi_n/\xi_{n-1}\) in the space \(X/\xi_{n-1}\).

Note that if the intersection of an element of the partition \(\zeta_n\) with an element \(C\) of the partition \(\xi_n/\xi_{n-1}\) is non-empty, then this intersection consists of a single point, and hence every point in \(C\) obtains a uniquely defined label, an element \(D \in \zeta_n\). The same is true if (as below) we replace the partition \(\zeta_n\) by a finite subpartition.

Denote by \(\pi_n\) the canonical projection to the quotient space:

\[
\pi_n : X \to X/\xi_n.
\]

We define a finite partition \(\sigma_n\) of the space \(X/\xi_{n-1}\) as a subpartition of \(\zeta_n\) \((\sigma_n \succ \zeta_n)\) by the following rule:

\[(*)\] by definition, two points \(y_1\) and \(y_2\) of an element \(D\) of the partition \(\zeta_n\) lie in the same element of \(\sigma_n\) if and only if the elements \(C_1 = \pi_n^{-1}(y_1)\) and \(C_2 = \pi_n^{-1}(y_2)\) endowed with the conditional measures of \(\xi_{n-1}\) are isomorphic as finite measured trees with points labeled by the elements of the partitions \(\sigma_i, i = 1, 2, \ldots, n - 1;\)

\[\text{5}\]The choice of a partition \(\zeta_1\) complementary to a given partition whose elements are finite or countable is a standard procedure developed in [45]: it proceeds by successively constructing sets of maximal measure for which the intersection with almost all elements of the given partition consists of at most one point. This is the main step in the classification of measurable partitions.
the partition $\eta_n$ is obtained by restricting each of the finite partitions $\sigma_i$ and $\eta_i$, $i = 1, \ldots, n - 1$, to the elements $C_1$ and $C_2$.

Note that, as explained above, an isomorphism between two elements $C_1$ and $C_2$, if it exists, is unique.

Now we can define the $n$th state space $Y_n$ as the quotient $\{X/\xi_1\}/\sigma_n$, and the path space of the $n$-step Markov chain as the quotient $\{X/\xi_1\}/\bigvee_{i=1}^n \sigma_i$; denote these spaces by $(\mathcal{M}_n, \nu_n)$.

The Markov property follows immediately from the construction: an element of the partition contains points with equal conditional measures.

As also follows from the construction, these spaces constitute, in a natural sense, an inverse spectrum of spaces, that is, $(\mathcal{M}_n, \nu_n) \leftarrow (\mathcal{M}_{n+1}, \nu_{n+1})$, and the projection preserves the Markov structure (a new space is being added), as well as the measure. Therefore, we can consider the inverse limit of these spaces, which is already the space of infinite trajectories of a one-sided Markov chain; denote this space, endowed with the limiting measure, by $(\mathcal{M}_\infty, \nu_\infty)$.

Thus, we have constructed a homomorphism of the Lebesgue space with a filtration $(\mathcal{M}_\infty, \nu_\infty, \mathfrak{F})$ to the space $(\mathcal{M}_\infty, \nu_\infty, \mathfrak{S})$, where $\mathfrak{S}$ is the tail filtration of the Markov chain.

By construction, $\sigma_n \succ \eta_n$ for all $n$, and consequently,

$$\bigvee_n \sigma_n \succ \bigvee_n \eta_n = \epsilon,$$

that is, the limiting homomorphism $X \to \mathcal{M}$ is an isomorphism of the measure spaces $(X, \mu)$ and $(\mathcal{M}, \nu)$, where $\nu$ is the image of the measure (the quotient measure). Denote by $\theta_n$ the $n$th coordinate partition of the space $\mathcal{M}$, that is, the partition into the classes of trajectories that coincide from the $(n + 1)$th position. The tail filtration $\mathfrak{S}$ gives rise to the sequence of partitions $\{\theta_n\}_n$.

Let us show that this isomorphism sends the filtration $\mathfrak{F}$ to $\mathfrak{S}$. For this, it suffices to observe that our construction (for the same reason as above) sends the space $X/\xi_1$ isomorphically to the quotient of $(\mathcal{M}, \nu)$ by the partition into the classes with respect to the first coordinate, and more generally, the space $X/\xi_n$ to the analogous quotient with respect to the $n$th coordinate. Hence, the bases of the corresponding partitions are sent to each other isomorphically, and, together with our choice of the partitions $\zeta_n$ (complementary to $\xi_n$), which are sent to the partitions into the classes of sequences that coincide from the $n$th coordinate, this yields the desired isomorphism of filtrations. Theorem 5 is proved. \(\Box\)

4.2. Remarks on Markov models. 1) We recall that in measure theory there are several theorems on Markov realizations of objects similar to the theorem proved above. For example: every transformation with invariant or quasi-invariant measure has an adic realization on the path space of a graded graph with a Markov measure.$^6$

$^6$An adic transformation [67], [68] (a generalized odometer) is defined for graphs in which for each vertex there is a linear order on the edges going to this vertex; using these orders, one can define in a natural way a lexicographic order on the sets of cofinal paths; then the adic automorphism sends a path to the next path with respect to this order if such a path is defined, which is the case on a set of full measure.
The theory of filtrations of subalgebras became a new source of examples in ergodic theory. Even the first example proposed by the author, that of the Pascal automorphism, still intrigues researchers; it is not yet known whether its spectrum is continuous, and many other properties are also unknown (see [75], [80], [41], [24]).

2) Another example is the theory of tame (hyperfinite) partitions (see [63]). Every discrete tame partition (in the category of spaces with a quasi-invariant measure) is isomorphic to the tail partition on the space of trajectories of a stationary Markov chain with finite state spaces and a measure quasi-invariant with respect to the shift.

3) Here is a corollary of Theorem 5.

**Corollary 2.** For an arbitrary locally finite filtration, the cocycle corresponding to the tail equivalence relation (see the definitions in §3.5) is isomorphic to a cocycle of a Markov chain with finite state spaces.

In particular, every hyperfinite (tame) partition with countable elements in a Lebesgue space is isomorphic to the tail partition of a Markov chain. In other words, every hyperfinite equivalence relation is isomorphic to the tail equivalence of a Markov chain; this is also a corollary of Theorem 5 on the Markov property of a locally finite filtration, a fact known for a long time (see [60]).

4) Of course, the model constructed in Theorem 5 is by no means unique: there are many graded graphs whose tail filtrations are isomorphic. In spite of this, the language of graded graphs and their tail filtrations on spaces of paths, or on spaces of trajectories of Markov chains, is extremely convenient. Note, however, that when passing from the original filtration (even a Markov one) to a new realization, we can lose some additional (non-invariant) properties, for example, the stationarity of the approximation: the original Markov chain could be defined as a stationary one, but the approximation constructed in the proof of Theorem 5 is rarely stationary. As we will see from examples in what follows, this loss is compensated by the simplicity of the analysis of the constructed approximation. The proof of Theorem 5 contains a method of constructing a graph in whose path space the given filtration can be realized. We obtain a class of rather difficult problems on the construction of graded equipped graphs with tail filtrations on path spaces isomorphic to a given filtration. Let us consider an important example.

5) Consider a dyadic filtration, for which all the partitions

$$\xi_1, \xi_2/\xi_1, \ldots, \xi_n/\xi_{n-1}, \ldots$$

have two-point elements with the measures $(1/2, 1/2)$. Our construction proves that every such filtration is isomorphic to the tail filtration of a dyadic graded graph endowed with a central measure on its path space; in this graph, all vertices at all levels except the initial vertex have two predecessors with equal cotransitions.

Therefore, the problem of metric classification of dyadic filtrations is reduced to the essentially combinatorial problem of classifying dyadic graded graphs (more precisely, path spaces of such graphs) endowed with central measures. Rephrasing results obtained in the 1970s to 90s, as well as recent facts and theorems, in these essentially combinatorial terms allows one to take a fresh look at the problem. The standardness criterion for dyadic filtrations can be rephrased as a criterion saying
when a dyadic graph endowed with a central measure can be reduced to the simplest Glimm graph (see Fig. 6). Specific Markov models of filtrations will be considered in more detail in §6, which is devoted to non-standard filtrations.

5. Standardness, criteria, finite isomorphism

5.1. The minimal model of a filtration. In this section, we associate with each infinite locally finite filtration \( \tau = \{ \xi_k \}_{k=1}^{\infty} \) of an arbitrary Lebesgue space \((X, \mu)\) with a continuous measure \(\mu\) its canonical realization as the tail filtration of an equipped graded multigraph, or equivalently, of a Markov chain. In contrast to the result of Theorem 5, the filtration constructed will not, in general, be isomorphic to the original filtration, but only finitely isomorphic.

As we will see, this construction leads to a special class of equipped graded multigraphs, which we call minimal. The model itself will be called the canonical minimal model of a filtration.

Note that a finite isomorphism does not preserve even the ergodicity of a filtration (this is obvious from examples of dyadic filtrations), so we add the ergodicity requirement whenever we compare a finite isomorphism and a true one.

Given an ergodic filtration \( \tau = \{ \xi_k \}_{k=1}^{\infty} \) on a Lebesgue space \((X, \mu)\), we will construct an equipped graded multigraph \( \Gamma(\tau) \) and a measure \(\nu\) on the path space \(T(\Gamma(\tau))\) that agrees with the equipment.

The construction, as in Theorem 5, is inductive: we successively build a multigraph (by levels), an equipment of this multigraph, and measures on its levels that will determine a measure on the whole path space.

We start from the vertex \(\emptyset\) at the zero level and associate with it the whole space, or more precisely, the quotient of \((X, \mu)\) by the trivial partition. Consider the first partition \(\xi_1\) and the finite (by the local finiteness of the filtration) partition \(\delta_1\) of the quotient space \(X/\xi_1\) into sets \(A_1, A_2, \ldots, A_{k_1}\) with the following property: any two points \(C, D \in X/\xi_1\), regarded as elements of \(\xi_1\), have isomorphic conditional measures if and only if they lie in the same element of \(\delta_1\). With the elements of the partition \(\delta_1\) of the space \(X/\xi_1\), that is, with the sets \(A_i, i = 1, 2, \ldots, k_1\), we associate the vertices at the first level of the graph \(\Gamma\). In other words, the vertices at the first level \(\Gamma_1\) correspond to the types of the conditional measures of the partition \(\xi_1\).

The number of edges going from each vertex to the vertex \(\emptyset\) is equal to the number of points in the corresponding element of the partition \(\xi_1\), and the cotransition probabilities are exactly the conditional measures of the points in this element. Finally, the measures of the vertices themselves are equal to the measures \(\mu_{\xi_1}\) of the sets \(A_i, i = 1, \ldots, k_1\).

The process continues in exactly the same way. It is easiest to say that if we already have the level \(\Gamma_m\), then the construction of the vertices, edges, and cotransition probabilities of the next level \(\Gamma_{m+1}\) proceeds as above, with the original filtration replaced by the filtration \(\{\xi_{m+s}/\xi_m\}_{s=1}^{\infty}\).

\footnote{Another natural term for minimal filtrations is ‘finitely determined filtrations’, since they are completely determined by their finite invariants (the same term ‘finitely determined’ should have been used for standard filtrations; see the next section 5.2).}
The sequence of partitions $\delta_1, \delta_2, \ldots$ of the respective spaces $X/\xi_1, X/\xi_2, \ldots$ corresponds to the partition of the vertices of the graph $\Gamma(\tau)$ into levels, and the elements of each partition $\delta_n$ correspond to the vertices at the $n$th level.

Thus, from the invariants of the filtration $\tau$ we have constructed a multigraph $\Gamma(\tau)$ with an equipment and a Markov measure on cylinder sets that agrees with this equipment and thus defines a Markov measure $\mu_{\Gamma(\tau)}$ on the space $T(\Gamma_{\tau})$ of infinite paths in $\Gamma(\tau)$. The resulting equipped locally finite graded multigraph $\Gamma(\tau)$ (we will call this multigraph minimal, and the corresponding tail filtration $\tau$ we call the minimal model of the original filtration $\tau$) has the following characteristic property, which obviously follows from the construction.

**Proposition 3** (theorem-definition of a minimal graph). An equipped multigraph is a minimal model of a locally finite filtration on a Lebesgue space with a continuous measure if and only if for any two distinct vertices $u, v$ at an arbitrary level, the trees with roots at $u$ and $v$ of paths leading from $\emptyset$ to these vertices are not isomorphic as measured trees.

The filtration $\tau$ is finitely isomorphic to $\tau$ (this is obvious from the construction), but, in general, not isomorphic to it.

Clearly, the above construction can be applied to the tail filtration of an arbitrary equipped multigraph to obtain a new multigraph with an equipment and a measure. If the equipment of the original graph is canonical (that is, the filtration is semi-homogeneous), then its standard model will also be a semihomogeneous filtration, since the image of a canonical equipment is obviously also canonical. Hence, our construction defines, in particular, an operation on graded graphs, by associating with an arbitrary multigraph a minimal multigraph. Of course, the minimal model of a minimal multigraph is the multigraph itself. In other words, this operation is a projection in the space of all graded multigraphs onto the space of minimal multigraphs.

We emphasize that in the basis method of defining a filtration, the basis of the algebra of all measurable sets in the path space of the constructed multigraph $\Gamma(\tau)$ is the algebra of cylinder sets, that is, sets of paths determined by some conditions on finite segments of paths. But when we map the tail filtration of one multigraph onto the tail filtration of another one, the inverse image of the $\sigma$-algebra of all measurable sets of the second multigraph does not in general coincide with the $\sigma$-algebra of all measurable sets of the first one in the space of the original filtration. In other words, we obtain homomorphisms, and not isomorphisms, of the spaces $X/\xi_n$, in contrast to Theorem 5 in the previous section, where the construction was aimed at obtaining an isomorphism. Namely, in the proof of Theorem 5 we had to subdivide the spaces $X/\xi_n$ and introduce in the graphs additional vertices corresponding to the same conditional measures on the path space of the tree of the original graph. In the new construction we identify these vertices and glue them together.

**Example 1.** The standard model of every ergodic dyadic filtration is the tail filtration of the Glimm multigraph of beads (see Fig. 6); this graph is minimal, and its path space is the dyadic compactum $\{0,1\}^\infty$ with the usual tail equivalence relation. Note that in this example there is a unique central measure, the Bernoulli
measure with probabilities $(1/2, 1/2)$. For brevity in what follows, filtrations finitely isomorphic to Bernoulli filtrations will be said to be finitely Bernoulli.

**Example 2.** The Pascal graph is not minimal, since symmetric vertices in the graph have isomorphic trees (see Fig. 7, showing the quotient graph). Moreover, if we introduce the ergodic central measure on the path space (this is the Lebesgue measure on paths regarded as sequences of zeros and ones), then the set of paths passing through one of the two first vertices shows that this is just the situation mentioned in the remark in the first paragraph of §5.1. Therefore, the filtration on the path space endowed with this measure is not standard. A simple example of an inhomogeneous non-standard Markov chain is given at the end of §5.5. An example of two measured trees that are isomorphic as trees but non-isomorphic as measured trees is shown in Fig. 8.

![Figure 7. A minimal graph: half of the Pascal graph.](image)

![Figure 8. Non-isomorphism of words with respect to the group of symmetries of a tree.](image)

Let us summarize the main conclusions.

**Proposition 4.** 1) Every locally finite filtration $\tau$ is finitely isomorphic to its minimal model $\overline{\tau}$, that is, to the tail filtration on a minimal multigraph $\Gamma(\tau)$, but is not in general isomorphic to it.

2) The minimal model, regarded as an equipped graph with a measure on the path space, is a complete finite-isomorphism invariant of filtrations, that is, two filtrations are finitely isomorphic if and only if their minimal models coincide.
3) The minimal model of an \( \{r_n\}_n \)-adic and, more generally, finitely Bernoulli filtration is a Bernoulli filtration. Therefore, an example of a minimal ergodic filtration is provided by any Bernoulli filtration.

**Proof.** Indeed, each finite fragment of the filtration has been mapped isomorphically with preservation of measure onto a finite fragment of the graph. And since we have used only conditional measures of all partitions and measures of sets with a fixed type of conditional measures, finitely isomorphic filtrations lead to the same standard models. Conversely, if two filtrations \( \tau = \{\xi_n\} \) and \( \tau' = \{\xi'_n\} \) are not finitely isomorphic, then for some \( n \) the partitions \( \xi_n \) and \( \xi'_n \) are not isomorphic, and hence their invariants contained in the minimal model will differ.

It is clear from the above that for every Markov chain with finite state spaces there exists a minimal Markov chain generating (as the tail filtration) a filtration finitely isomorphic to the given one. \( \square \)

We see that the problem of finite classification of locally finite filtrations is equivalent to the problem of constructing the space of minimal equipped graded graphs with a measure on the path space that agrees with the equipment, that is, the space of complete invariants of such filtrations (for more on the space of filtrations, see §5.6 below).

Unfortunately, passing to minimal models and minimal filtrations does not simplify the study of filtrations, since in a natural sense any filtration can be approximated arbitrarily closely by minimal ones. However, minimality brings us nearer to the fundamental notion of standardness.

5.2. Minimality of graphs, standardness, and the general standardness criterion. The following important definition will often be used in what follows.

**Definition 7.** A locally finite ergodic filtration is said to be standard if it is isomorphic to a minimal filtration.\(^8\)

For example, a \( p \)-adic filtration is standard if it is isomorphic to a Bernoulli filtration; in other words, in the class of filtrations finitely isomorphic to Bernoulli ones, only the Bernoulli filtrations themselves and those isomorphic to them are standard.

The difficulty is that a filtration isomorphic to a minimal one (that is, a standard filtration) is not necessarily minimal, in other words, is not necessarily the tail filtration of a minimal multigraph (or a minimal Markov chain). Thus, the minimality property for filtrations is not invariant under isomorphism, and the ‘invariant hull’ of this property is exactly ‘standardness’.

Hence, the problem arises of how to describe standard filtrations in invariant (intrinsic) terms, that is, how to check that a given filtration is isomorphic to a standard one.

In the next theorem we give an invariant characterization of a standard filtration (the ‘standardness criterion’).

Let \( \tau = \{\xi_n\}_n \) be an ergodic locally finite filtration, and let \( \eta_n \) be the finite partition of the space \( X/\xi_n \) into the maximal classes of elements \( C \in \xi_n \) on which the restrictions of the finite filtration \( \{\xi_k\}_{k=1}^{n-1} \) are isomorphic (equivalently, the corresponding measured trees are isomorphic). Fix an arbitrary isomorphism of

---

\(^8\)See footnote 7 about the term ‘finitely determined filtration’.
these elements with some typical element $\overline{C}$ regarded as a measured tree. Then obviously every element $D_i$, $i = 1, \ldots, r_n$, of the partition $\eta_n$ is a direct product:

$$D_i \simeq \overline{C} \times (D_i/\xi_n), \quad D_i \in \eta_n.$$ 

Thus, we have defined an independent complement $\xi_{n,i}^-$ of the restriction of the partition $\xi_n$ to $D_i$; this is a finite partition of $D_i$.

**Theorem 6.** A locally finite ergodic filtration $\tau = \{\xi_n\}_n$ of a Lebesgue space $(X, \mu)$ with a continuous measure is standard if and only if, in the above notation, for every $\varepsilon > 0$ and every measurable set $A$ there exists a number $n$ and a set $A'$ with $\mu(A \triangle A') < \varepsilon$ such that

$$A' = \bigcup_{i=1}^{r_n} A'_i$$

and $A'_i$ is measurable with respect to the independent complement $\xi_{n,i}^-$. 

The condition of the theorem means that every measurable subset of $X$, for sufficiently large $n$ and up to a modification on a set of small measure, splits into a union of sets measurable with respect to the maximal possible independent complement of the restrictions of the partition $\xi_n$. An important fact here is the maximality of the restrictions of $\xi_n$ for which such a complement exists. However, the choice of these independent complements is subject to additional considerations.

For filtrations finitely isomorphic to Bernoulli ones (for example, for dyadic filtrations), the partitions $\xi_n$ themselves have independent complements, so for such filtrations the criterion consists in verifying the possibility of constructing coherent independent complements constituting a basis of the space.

In this sense the condition of Theorem 6 simply reduces the situation to the case of a homogeneous filtration, where it is *tautological*.

**Proof of Theorem 6.** The ‘if’ part is equivalent to the assertion that the partitions into cylinders in the path space of the graph in the definition of a standard model (see Definition 7) constitute a basis of the $\sigma$-algebra of measurable sets. The ‘only if’ part follows from the fact that the basis of cylinders is exactly the basis of independent complements of the indicated (maximal) restrictions of the partitions of the filtration. $\square$

The condition of Theorem 6 shows that the standardness of a filtration is in a sense the property of its decomposability into Bernoulli components. For finitely Bernoulli filtrations, this is exactly the Bernoulli property, that is, independence. Hence, the notion of standardness is a generalization of the notion of independence.

In the next section we will formulate a concrete standardness criterion which uses and generalizes the criterion for finitely Bernoulli filtrations proposed by the author in 1970 [59].

It follows from the conditions of the theorem that a possible obstacle to standardness is the existence of a special group of symmetries preserving the measure in the path space of the graph.
5.3. The combinatorial standardness criterion. The general standardness criterion given above (Theorem 6) requires checking the measurability of sets with respect to a system of independent complements, but does not show how one can do this. In this section we give a more constructive, in fact combinatorial, method for verifying standardness.

5.3.1. The combinatorial criterion in terms of partitions or metrics. We start with the standardness criterion for finitely Bernoulli filtrations; as we will see, the general case will be obtained by ‘mixing’ them. The criterion to be stated differs from the standardness criterion proposed by the author in the 1970s for dyadic sequences only by a more universal statement. Note that in the form given below this criterion holds also for filtrations with continuous conditional measures.

For a Lebesgue space \((X, \mu)\) with a continuous measure let \(\tau = \{\xi_n\}\) be a locally finite filtration that is finitely isomorphic to a Bernoulli filtration, that is, for every \(n\) the quotient partition \(\xi_n/\xi_{n-1}\), \(n = 1, 2, \ldots\), is isomorphic to a partition whose elements are all finite, and all the conditional measures coincide and are equal to the distribution determined by a finite probability vector \(p^{(n)}\), \(n = 1, 2, \ldots\), possibly different for different \(n\). The leading special case is the dyadic case:

\[ p^{(n)} \equiv (1/2, 1/2) \quad \text{for all } n. \]

In our case an element \(C\) of the partition \(\xi_n\), regarded as a finite set endowed with the conditional probability measure, has the structure of a direct product of finite spaces endowed with the measure \(\prod_{k=1}^{n} p_k\), but it is more convenient to represent it as a homogeneous tree of rank \(n\) with the product measure on leaves.

Assume that on the space \((X, \mu)\) we have a bounded measurable function \(f\) with real values, or a metric, or a semimetric \(\rho\) (understood here to mean that \(\rho\) satisfies the conditions of a metric except that it may equal 0 on a pair of distinct points); then on the space \(X/\xi_n\) we can introduce a distance \(d^n_f(\cdot, \cdot)\) \((d^n_\rho)\), which should be interpreted as the distance between elements \(C_1, C_2\) of the partition \(\xi_n\) regarded as measured trees. This distance is a version of the Kantorovich metric for probability measures in a metric space, but in our case there is an additional condition.

Namely, we consider so-called couplings \(\psi\), that is, measures on the direct product \(C_1 \times C_2\) with the given marginal projections \(\mu^{C_1}\) and \(\mu^{C_2}\) equal to the conditional measures. The additional condition is that these couplings are not simply measures, but also couplings of trees, that is, \(\psi\) should be concentrated on some tree, as a subset of \(C_1 \times C_2\) whose projections on both coordinates are \(C_1\) and \(C_2\) as trees. An example of a coupling of trees is the graph of a measure-preserving isomorphism of one tree onto the other.

The measures \(\psi(\cdot, \cdot)\) are called couplings, or transportation plans; the set of all couplings will be denoted by \(\Psi_{C_1, C_2}\).\(^9\)

Then the distance \(d^n_f\) \((d^n_\rho)\) is given by the formula

\[
d^n_f(C_1, C_2) = \min_{\psi \in \Psi_{C_1, C_2}} \sum_{x \in C_1, y \in C_2} |f(x) - f(y)| \psi(x, y)
\]

\(^9\)The metric or semimetric \(d^n_\rho\) is exactly the Kantorovich metric on the set of probability measures of the (semi)metric space \((Y, \rho)\); in our case, we take the semimetric \(\rho(x, y) = |f(x) - f(y)|\) on \(Y = C_1 \cup C_2\).
(in case of metrics, one should replace $|f(x) - f(y)|$ by $\rho(x, y)$).

**Theorem 7.** An ergodic filtration finitely isomorphic to a Bernoulli one is standard, that is, isomorphic to a Bernoulli filtration, if and only if the following condition holds:

- for every measurable function $f$ on the space $(X, \mu)$, or
- for every admissible metric $\rho(\cdot, \cdot)$ on $(X, \mu)$, or
- for every cut semimetric

the equality

$$\lim_{n \to \infty} \int_{X/\xi_n \times X/\xi_n} d^n_\rho(C_1, C_2) \, d\mu_{\xi_n} \, d\mu_{\xi_n} = 0$$

is valid.

For what follows, it is convenient to give an equivalent $\varepsilon$-form of the condition in the theorem:

$$\forall \varepsilon > 0, \; \exists N, \; \forall n > N, \; \exists D_n \subset X/\xi_n : \mu_{\xi_n}(D_n) > 1 - \varepsilon \quad \text{and} \quad \sup_{C_1, C_2 \in D_n} d^n_\rho(C_1, C_2) < \varepsilon.$$

The condition in the theorem is especially clear for dyadic filtrations of cut semimetrics with respect to partitions into $k$ subsets: in this case a substantial simplification of the condition is that we may take couplings to be the graphs of isomorphisms of trees; that is, we may consider the Hamming metric on the set of vertices of the cube $k^{2^n}$, and then (instead of couplings $\psi$) take the distance between the orbits of the action of the group of automorphisms of the dyadic tree on this cube. In other words, the standardness criterion reduces to the fact that the space of orbits of the action of the group of automorphisms of the tree on the cube $k^{2^n}$ collapses in this metric to a single point (see [59], [69]).

The proof of the criterion for homogeneous, that is, $\{r_n\}$-adic, filtrations exactly coincides with the proof for dyadic ($r_n \equiv 2$) filtrations and was presented in [65], [69]. Several other expositions appeared later (for example, see [7], [13], [14]). The scheme of the proof remains exactly the same also for finitely Bernoulli filtrations, as well as for continuous filtrations (in which all conditional measures of all partitions $\xi_n/\xi_{n-1}, \; n = 1, 2, \ldots$, are continuous).

We will explain only the main idea of the proof of Theorem 7 (for details, see [69]). First of all, the ‘only if’ part follows from the fact that, by definition, for a Bernoulli filtration there exists a sequence $\{\eta_n\}_n$ of independent partitions constituting a basis of the space such that

$$\xi_n = \bigvee_{k=n}^{\infty} \eta_k, \quad n = 0, 1, 2, \ldots.$$

Namely, it follows that for all finite partitions $\bigvee_{k=1}^n \eta_k$ the condition in the theorem follows from their independence of the corresponding partitions $\xi_n$. On the other hand, $\{\eta_n\}_n$ is a basis, and hence the condition is satisfied for any finite partitions.

---

10That is, a semimetric $\rho(x, y)$ equal to 1 if $x, y$ lie in different elements of a finite partition of the space $(X, \mu)$, and to 0 if $x, y$ lie in the same element.
The proof of the ‘if’ part is more laborious, one should prove that the condition of the theorem implies the existence of a sequence \( \{\eta_n\} \) of independent partitions associated with \( \{\xi_n\} \) (as described above) that is a basis of the measure space. This can be done again by induction, and the main lemma is that the condition of the theorem can be extended to the quotient filtrations \( \{\xi_{n+k}/\xi_k\}_{n=0}^{\infty}, k = 1, 2, \ldots \), and hence one can construct a basis \( \{\eta_n\} \) by successively passing to quotient spaces and approximating some basis chosen in advance. This scheme was first used in the proof of the lacunary isomorphism theorem for dyadic sequences [58].

5.3.2. Comments on the criterion. The statement of Theorem 7 involves the notion of an admissible metric (see [71]). Recall that an admissible metric in a Lebesgue space is a metric defined on a set of full measure as a measurable function of two variables such that the \( \sigma \)-algebra generated by all balls of positive radius in this metric generates the whole \( \sigma \)-algebra of measurable sets (see [92]). If a function of two variables satisfies the conditions of a metric (non-negativity, symmetry, triangle inequality) as conditions mod 0 on classes of metrics, as well as the above condition on the \( \sigma \)-algebra of balls, then the mod 0 class of this function contains an admissible metric (see [61]).

Theorem 8. Let \( \rho \) be an arbitrary admissible metric on a Lebesgue space with a continuous measure. An ergodic filtration finitely isomorphic to a Bernoulli filtration is standard, that is, isomorphic to a Bernoulli filtration, if and only if the following condition holds: for any \( \varepsilon > 0 \) there exists a positive integer \( N \) such that for every \( n > N \) and some set
\[
A_n \subset X/\xi_n, \quad \mu_{\xi_n}(A_n) > 1 - \varepsilon,
\]
the Kantorovich distance between the conditional measures on arbitrary elements \( C_1, C_2 \subset A_n \) of the partition \( \xi_n \) in the space \( (X, \rho) \) regarded as a metric space does not exceed \( \varepsilon \). This condition is or is not satisfied independently of the choice of \( \rho \).

Note that if in the condition of the last theorem we take \( \rho \) to be the cut semimetric corresponding to a partition, then we obtain the condition of the theorem for functions with finitely many values. Hence, the conclusion of the theorem for cut semimetrics follows from the conclusion for functions. On the other hand, every semimetric is a limit of sums of cut semimetrics, hence the conclusion of the theorem for functions implies the conclusion for a dense family of semimetrics. Finally, it is not difficult to check that the fact that the elements of the partition endowed with the conditional measures approach one another in the introduced metric remains true when passing to the limit along a sequence of metrics.\(^{11}\)

\(^{11}\)We note, by the way, that the idea of varying metrics and the notion of an admissible metric for a fixed measure in a space was proposed by the author in the 1980s and has been repeatedly used, for example, in the definition of scaled entropy (see [71], [82]).

The conclusion of the theorem can be equivalently stated as the convergence of distances to zero in measure, and similarly.

\(^{11}\)More precisely, the Kantorovich metric on the space of probability measures of a given Borel space is continuous with respect to the weak topology in the space of admissible metrics (regarded as functions of two variables). The same considerations imply that the conclusion of the theorem is independent of the choice of a metric.
5.4. The standardness criterion for an arbitrary filtration. How should one modify the condition of Theorem 6 to obtain a standardness criterion for an arbitrary locally finite filtration? In fact, the condition for the general case reduces to a mixture of conditions on each finitely Bernoulli component of the filtration.

On the spaces \((X/\xi_n, \mu_{\xi_n})\) we will need the finite partitions \(\delta_n\) introduced when defining minimality in §5.1. Recall that two elements \(C_1, C_2\) of the partition \(\xi_n\) lie in the same element of \(\delta_n\) if the finite filtration \(\{\xi_k\}_{k=1}^{n-1}\) induces isomorphic filtrations on \(C_1\) and \(C_2\), or equivalently, if these elements \(\tilde{C}_1\) and \(\tilde{C}_2\) are isomorphic as trees endowed with the conditional measures. It is essential that the partition \(\delta_n\) is determined only by the conditional measures. For homogeneous filtrations this is the trivial partition of the space \(X/\xi_n\), and for minimal graphs this is the partition into singletons at each level.

Then the statement in the \(\varepsilon\)-language is as follows.

**Theorem 9.** Let \(\tau = \{\xi_n\}_{n=0}^{\infty}\) be an arbitrary ergodic locally finite filtration in a Lebesgue space \((X, \mu)\) with a continuous measure. The filtration \(\tau\) is isomorphic to a standard filtration if and only if the following condition holds: for any \(\varepsilon > 0\) there exists an \(N\) such that for every \(n > N\) the partition of \(X/\xi_n\) into the types of elements of \(\xi_n\) isomorphic with respect to the previous fragment of the filtration,

\[
X/\xi_n = \bigcup_{i=1}^{k_n} A^n_i, \quad \{A^n_i\}_i = \delta_n,
\]

has the property that

there exists a \(D \subset X/\xi_n\) with \(\mu_{\xi_n}(D) > 1 - \varepsilon\)

such that

\[
sup_{C_1, C_2 \in A^n_i \cap D} d_f(C_1, C_2) < \varepsilon,
\]

in other words, inside each element of the partition \(\delta_n\), the distances between most elements of the partition \(\xi_n\) are small with respect to the metric \(d_f\) or \(d_\rho\).

The sense in which standardness is a generalization of independence should be clear from this statement.

In fact, it differs from Theorem 8 (that is, the theorem for finitely Bernoulli filtrations) only in that the pairwise distances should be small only for most pairs of isomorphic elements, and not for all such pairs. Therefore, the proof of the ‘if’ part, which is the main part of the theorem, does not differ from the previous case. However, checking the condition is by no means easy.

It makes sense to state the ‘only if’ part of the criterion as a separate proposition.

**Proposition 5.** 1) The filtrations satisfying the standardness criterion form an invariant class, that is, a filtration isomorphic to a filtration satisfying the standardness criterion also satisfies this criterion.

2) A minimal filtration satisfies the standardness criterion.

**Proof.** The first assertion follows directly from the statement of the criterion.

The proof of the second assertion splits into two cases. In the first (obvious) case, assume that a minimal multigraph is a graph, that is, has no multiple edges. Then the partition into the types of elements of the partitions \(\xi_n\) coincides with
the partition into cylinder sets, and the $\sigma$-algebra generated by all complements coincides with the whole $\sigma$-algebra. In the second case the partition into types reduces to Bernoulli components, for which the sufficiency of the criterion has been proved. $\square$

The standardness criterion is of interest in the case of graphs essentially different from minimal ones. For canonical equipments, that is, central measures, the criterion can be somewhat simplified, and it was stated in this simplified form in our previous papers; but for filtrations that are not semihomogeneous, the simplified statement is not equivalent to the general one given above, and it is in this sense that the term ‘standardness’ is used. We will return to this question in connection with the so-called internal metric (see [82], [79], [78], [81], and §7).

5.5. The martingale form of the standardness criterion and intermediate conditions. It is natural to ask how, given a Markov, or even arbitrary, one-sided random process, one can determine whether or not its tail filtration is standard. The problem is to state the standardness criterion in terms of the process.

The most natural form of the answer to this question, as well as to other questions related to the interpretation of weakenings of the independence condition in the theory of random processes, looks like a strengthening of martingale convergence theorems. The classical theorem for a random process $\{x_n\}_{n \leq 0}$ in the form we require is as follows.

If the tail filtration of the process (that is, the intersection of the past $\sigma$-algebras) is trivial,

$$\bigcap_n \mathcal{A}_n = \mathcal{N},$$

then for any real functional $F$ in $k + 1$ coordinates of this process, the conditional distributions of $F$ given a fixed past (starting from $-n - 1 < k$) converge almost everywhere to the unconditional distribution of this functional:

$$\lim_n \text{Prob}(F(x_0, x_{-1}, \ldots, x_{-k}) \geq t \mid x_{-n-1}, \ldots) = \text{Prob}(F(x_0, x_{-1}, \ldots, x_{-k}) \geq t), \quad t \in \mathbb{R};$$

in other words, as $n$ tends to infinity, the conditional measures on $k + 1$ coordinates of the process converge weakly almost everywhere (or in measure) to the corresponding unconditional measures. Here it is essential that $k$ is fixed and $n \to \infty$. Can $k$ be allowed to go to infinity along with $n$? Literally, this does not make sense, since the limit of the conditional measures for growing $k$ is not defined, but one may speak about the conditional measures approaching one another as $n$ goes to infinity. For this, one should define some metric in the space of conditional measures for a given $k$ and require, for instance, that the expectation of the distances between the conditional measures converges to zero under different conditions. Here the choice of a metric plays a crucial role, and the classes of processes satisfying such requirements depend substantially on this choice. But in any case, the validity of conditions of this type shows that the process is kind of similar to a Bernoulli process.
Let us state the standardness criterion in terms of convergence of a sequence of conditional measures (that is, measures on spaces of measured trees); the criterion can be regarded as a ‘martingale’ definition of standard random processes.

**Theorem 10** (Martingale standardness criterion). Consider an arbitrary (not necessarily Markov) random process \( \{x_n\}_{n<0} \) with values in \([0,1]\) (or in an arbitrary Borel space), and consider its tail filtration \( \mathcal{A}_n, n = 0, 1, 2, \ldots \), where \( \mathcal{A}_n \) is the \( \sigma \)-algebra generated by the random variables \( x_k, k \leq -n, n = 0, 1, 2, \ldots \). Assume that the filtration is ergodic, that is, the intersection of the \( \sigma \)-algebras satisfies the zero–one law:

\[
\bigcap_{n=0}^{\infty} \mathcal{A}_n = \mathcal{N}.
\]

Take either

- an arbitrary cylinder function \( f \) on the space of trajectories of the process, or
- an arbitrary metric \( \rho \) on the space of trajectories of the process.

Fix two trajectories \( C_1 \) and \( C_2 \) of the process for time from \(-\infty\) to \(-n\) with isomorphic conditional measures \( \mu^{C_1} \) and \( \mu^{C_2} \) on the finite segments

\[
(x_0, x_{-1}, x_{-2}, \ldots, x_{-n}|C_1) \quad \text{and} \quad (x_0, x_{-1}, x_{-2}, \ldots, x_{-n}|C_2),
\]

that is, two points of the quotient space \( X/\xi_n \) lying in the same element of the partition \( \delta_n \) (see the standardness criterion, Theorems 7 and 8), and consider the trees of trajectories corresponding to the elements \( C_1 \) and \( C_2 \) endowed with the conditional measures on the leaves. Then the filtration is standard if and only if

\[
\lim_{n \to \infty} \int_{X/\xi_n} \int_{X/\xi_n} df(\mu^{C_1}, \mu^{C_2}) d\mu_{\xi_n} d\mu_{\xi_n} = 0
\]

(in the case of a metric, one should replace \( df \) by \( d\rho \)) for all measurable functions \( f \) on the space of trajectories of the process (it suffices to check this only for cylinder functions with finitely many values), or for all admissible metrics (actually, it suffices to check this only for one metric).

This statement is a verbatim reproduction of the standardness criterion, but it leads to an important conclusion: between the usual martingale convergence theorem, which holds for all processes determining an ergodic filtration (that is, regular processes), on the one hand, and the standardness criterion, which has less applicability, on the other hand, there is a natural spectrum of intermediate conditions on processes. They are determined by the relation between the moment \( n < 0 \) that fixes a condition in the past and the moment \( k(n) > n \) up to which we compare the conditional measures on trees of rank \( |k(n)| \). If \( k(n) \) does not depend on \( n \) and \( n \to \infty \), then this is the usual martingale convergence theorem, but if \( k(n) = n \), then this is the standardness criterion. Hence, standardness is the greatest possible strengthening of the zero–one law. The author does not know anything about the intermediate cases, which can be called ‘higher zero–one laws’, but one can expect that processes for which the conditions hold for different intermediate growth rates of the sequence \( k(n) \) have different properties.

The most interesting part of the theory is the study of filtrations that are finitely Bernoulli but not isomorphic to Bernoulli filtrations. Deviations from the conditions

300 A. M. Vershik
of the standardness criterion can be measured by various invariants, for instance, the entropy of filtrations. We will return to this problem in §6.

The simplest example of an inhomogeneous finitely Bernoulli non-standard filtration is the Markov chain with two states and transition matrix

\[
\begin{pmatrix}
p & q \\
q & p
\end{pmatrix}, \quad p \neq q, \quad p + q = 1, \quad p, q > 0
\]

(see [82]); the cylinder set \( \{ x = \{ x_n \} : x_1 = 0 \} \) does not satisfy the standardness criterion. This filtration is a double cover of the Bernoulli filtration with probabilities \((p, q)\).

Conversely, the Fibonacci Markov chain with transition matrix

\[
\begin{pmatrix}
\lambda^2 & \lambda \\
1 & 0
\end{pmatrix}, \quad \lambda^2 + \lambda = 1,
\]

which is not locally Bernoulli, determines a standard non-homogeneous tail filtration, since the partitions \( \delta_n, n = 1, 2, \ldots, \) are the partitions into singletons of each level of the graph.

5.6. The spaces of filtrations, Markov compacta, and graded multigraphs. The theorem on the Markov property of filtrations allows one to take the space of equipped graded multigraphs, or equivalently, the space of Markov compacta with Markov measures, as the space of all locally finite filtrations in a Lebesgue space. In other words, we give the following definition.

**Definition 8.** The space of locally finite filtrations in a Lebesgue space with a continuous measure is identified with the space of equipped graded multigraphs (respectively, with the space of Markov chains with cotransition probabilities) endowed with a fixed continuous measure that agrees with the equipment (respectively, with the cotransitions).

With this definition, a filtration is identified with the tail filtration of the path space of a multigraph (respectively, with the space of trajectories of a Markov chain). We emphasize that an important role in the convention to identify the collection of filtrations in measure spaces and the collection of equipped graded multigraphs is played by the measure on the path space that agrees with the equipment: the equipped graph alone does not determine, for instance, whether the filtration is ergodic; this is determined by the measure.

For definiteness, we will speak about multigraphs; reproducing all statements in terms of Markov chains brings nothing new.

As we have seen, an equipped multigraph can be defined by a sequence of finite matrices whose non-zero entries are the probabilities of cotransitions of adjacent vertices. Each such sequence determines an equipped graded multigraph \( \Gamma \). The path space of \( \Gamma \) will be denoted by \( T(\Gamma) \). Moreover, given a multigraph \( \Gamma \), we introduce the simplex \( \Sigma_{\Gamma} \) of all probability measures on the space \( T(\Gamma) \) that agree with the equipment.
Denote by $\mathcal{F}$ the space of all pairs $(\Gamma, \Sigma_\Gamma)$, where $\Gamma$ runs over the set of all infinite equipped graded locally finite multigraphs and $\Sigma_\Gamma$ is the corresponding simplex of measures. This space can also be viewed as the space of all locally finite filtrations of a continuous Lebesgue space, and simultaneously as the space of all locally finite Markov compacta with Markov measures. We endow the space with the weak topology, that is, the topology of an inverse spectrum.

We may consider the following important subspaces of this space.

1. The subspace of ergodic multigraphs, that is, multigraphs for which there exists at least one continuous ergodic measure on the path space that agrees with the equipment. This class does not contain, for example, trees and other decomposable multigraphs. It is of most interest for the needs of the theory of filtrations and ergodic theory. The combinatorial structure of such multigraphs must be completely described. In fact, this is a question about hyperfinite equivalence relations on a standard Borel space (the path space) with a fixed cocycle of measures (cf. [6], [78], [52]).

2. The subspace of multigraphs with canonical equipments (the tail filtration in this case is homogeneous, that is, the conditional measures on the elements of $\xi_n$, $n = 1, 2, \ldots$, are uniform). The geometry of the projective limit of simplices and the limiting simplex $\mathcal{F}_0$ of central measures are quite special; in the dyadic case, the limiting simplex is the simplex of unordered pairs, or the so-called 'tower of dyadic measures' (see Fig. 11 in §7.3). In the language of Markov chains (in the stationary case), central measures are measures of maximal entropy. The canonical equipment is determined by the very structure of the graph, which is why this space is of special interest.

3. The space $\mathcal{F}^{\text{st}}$ of stationary filtrations (that is, filtrations invariant under the shift, which is the passage to the quotient by the first partition: $\tau \equiv \{\xi_n\}_{n=0}^\infty \simeq \{\xi_n/\xi_1\}_{n=1}^\infty \equiv \tau'$), or equivalently, the space of stationary multigraphs (that is, multigraphs with constant cotransition matrices) with an additional symmetry, and so on. The analysis of this space is of special importance for ergodic theory (of one automorphism).

4. Moreover, in all cases one can avoid considering measures on path spaces and pass to the Borel theory of graphs and filtrations. This means that we study only the space $\mathcal{G}$, without fixing the second component in pairs $(\Gamma, \Sigma_\Gamma)$, that is, the simplex of measures $\Sigma_\Gamma$, and consider only cotransitions. (See §7.)

It is more important to analyze the space $\mathcal{F}$ from the point of view of the problem of finite and general isomorphism. Let us select in $\mathcal{F}$ filtrations corresponding to minimal multigraphs (minimal Markov chains) with an equipment and a measure on the path space. By the previous analysis, this is the subspace of $\mathcal{F}$ consisting of the minimal locally finite filtrations. By Theorem 5, this is the space of complete invariants of filtrations up to finite isomorphism: no two different standard filtrations are finitely isomorphic, and any filtration is finitely isomorphic to a standard filtration.

Our definitions immediately imply the following assertions.

**Proposition 6.** 1) The space of all equipped graded multigraphs (respectively, the space of all Markov compacta with cotransitions) is fibred over the space of minimal equipped graphs (respectively, over the space of minimal Markov compacta with
Hence, the space of all filtrations is fibred over the space of minimal filtrations. The space of minimal equipped graphs is dense in the space of filtrations.\(^{12}\)

2) The set of standard ergodic filtrations is dense in the space of all filtrations.

Basically, the space of all minimal multigraphs (Markov chains) is the space of all locally finite types of filtrations.

Therefore, in contrast to the problem of finite isomorphism, the problem of isomorphism of locally finite filtrations is wild: passing to the quotient of \(\mathfrak{F}\) by the classes of isomorphic filtrations does not result in a space with a separable Borel structure. Moreover, the latter remains true even if we consider the problem of isomorphism of ergodic dyadic filtrations, that is, in one (in fact, an arbitrary non-trivial) class of finite isomorphism. The situation is similar to many isomorphism problems in algebra and analysis, for example, the problem of isomorphism of actions of infinite groups with an invariant measure, or the problem of classifying irreducible unitary representations of the infinite symmetric group, and so on. This question arises in the theory of dynamical systems, representation theory, asymptotic combinatorics, and other topics. The wildness of the classification problem follows from the fact that the classification includes a typical wild problem; we will return to this in another paper.

On the one hand, we consider the isomorphism problem below for some classes of filtrations close to standard ones, and on the other hand, we propose types of invariants for general filtrations substantially different from standard ones.

6. Non-standardness: examples and sketch of a general theory

The most difficult problems in the theory of filtrations involve non-standard filtrations. But, as follows from the standardness criterion, non-standardness for inhomogeneous filtrations can have different causes. Recall that a minimal filtration is uniquely determined by its finite type. The standardness of a filtration means that it does not differ much from its finite type; roughly speaking, its homogeneous parts must be standard, that is, Bernoulli.

Thus, the nature of true non-standardness shows itself already in homogeneous (and, moreover, even dyadic) filtrations. As we have already mentioned, the group of symmetries of the filtration itself and its quotients by the partitions \(\xi_n\) with arbitrarily large \(n\) must be infinite, otherwise the filtration cannot have a complicated structure at infinity.

We continue constructing Markov chains and graphs for non-standard dyadic filtrations in this section. Actually, these chains with finite state spaces, that is, path spaces of graded multigraphs, should be regarded as an approximation, or more precisely, a simple model of continual, as well as more complicated, Markov chains. But perhaps the most instructive thing in these constructions is that the graphs obtained from these approximations are interesting in themselves and, possibly, of even more importance than the approximation itself. This fact should be emphasized, because these graded graphs, regarded as Bratteli diagrams, represent

\(^{12}\)At the same time, in the space, say, of ergodic dyadic filtrations there is only one minimal ergodic filtration and only one minimal graph. In general one can introduce a more complicated topology in which the set of minimal filtrations is closed.
important $C^*$-algebras which seem to have never been studied in the literature. For example, the algebras generated by the graph of words for the Abelian group $\mathbb{Z}^d$ (see below) look mysterious.

The method of proving non-standardness contained in the criterion becomes clearer if one endows it with a combinatorial meaning. This is exactly what we do below in several examples. In fact, the problem is reduced to the analysis of a measure on the orbits of an action of the group of automorphisms of a tree on some subsets of configurations. The most transparent case is the action of this group on the set of vertices of the cube $2^n$ with the uniform measure or with an arbitrary product measure with equal factors. The standardness criterion requires that the measure concentrates (with respect to the Hamming metric) in a neighbourhood of one orbit, and we need advanced methods to prove this fact. Unfortunately, such clarity has not yet been achieved in the examples given below.

6.1. Graphs of random walks over orbits of groups, graphs of words. The construction of a Markov approximation (whose existence follows from Theorem 5) as a Markov chain with finite state spaces is a rich source of new graded graphs. On the other hand, these constructions allow one to study filtrations given a priori as Markov filtrations with continual state spaces. Specializing the proof of Theorem 5, we will give an explicit construction of a Markov chain with finite state spaces and obtain a ‘graphic’ interpretation of the filtration. We mean models of random walks over the orbits of an action of a countable group in a space with a continuous invariant measure and the tail filtration of the corresponding Markov chain. This subject is sometimes called the theory of random walks in random scenery (RWRS), and it is a further step in the theory of random walks on groups, a step currently the subject of hundreds of papers. It should be said that this relatively new field deals with problems of a different nature: for instance, problems related to boundaries of walks, entropy, and so on (see, for example, the circle of problems in [26]) are replaced by problems about the relation between the action of the group and properties of filtrations (see [77]).

The question of when the corresponding filtrations (even dyadic ones) are standard has not been sufficiently studied. It is known only that the filtration of the Markov process $(T, T^{-1})$, that is, the simple random walk over the orbits of an automorphism $T$, is not standard if the entropy is positive, $h(T) > 0$ (see [22]).

For automorphisms $T$ with a discrete spectrum, this filtration is presumably standard. For the rotation $T_\lambda$ of the circle, the author proved this (unpublished) for well-approximable $\lambda$, and then W. Parry proved it in full generality (unpublished).

But these are only the first examples of random walks over orbits of actions; walks that are non-simple (that is, walks not only over generators) offer many more possibilities. One can hope that the machinery of graphs will provide new tools for investigating such problems.

We begin the construction with the following general example. Let $G$ be a group and let $S$ with $|S| = s < \infty$ be a set of generators of $G$. Consider the set $k^G$, where

$$k = \{0, 1, \ldots, k - 1\},$$
and let $m$ be a Bernoulli measure on $k^G$, that is, the direct product $\prod_G p$, where
\[
\{p_g\}, \quad g \in S, \quad \sum_g p_g = 1, \quad p_g \geq 0,
\]
is a probability vector of dimension $s$. Assume that there is a free action of $G$ on a Lebesgue space $(X, \mu)$ with a continuous invariant measure. Consider the (one-sided) Markov process $(y_n)_{n \leq 0}$ with the state space $X$ and with transition probabilities (since we consider negative moments of time) equal to the cotransition probabilities:
\[
\text{Prob}(y|x) = \begin{cases} p_g & \text{if } y = T_g x, \ g \in S, \\ 0 & \text{otherwise.} \end{cases}
\]
The measure $\mu$ is the invariant initial measure of the Markov process, and we also have the Markov measure $M$ defined on the space $\mathcal{X}$ of all trajectories of the process.

Consider on the space $(\mathcal{X}, M)$ the tail filtration $\tau = \{\mathcal{A}\}_{n=0}^{\infty}$ determined by this stationary Markov process. Obviously, it is locally finite (since $S$ is finite), but the Markov process has a continual state space. However, by Theorem 5 the corresponding filtration can be defined as the tail filtration in the path space of some graded graph, or as the tail filtration of a Markov chain with finite state spaces. Of course, this Markov chain will no longer be stationary, since, for instance, the cardinality of the state spaces, that is, the number of vertices on levels of the graph, is growing. But the filtration itself certainly remains stationary, only the approximation is not shift-invariant. This chain is not unique, of course, and, as already mentioned, in general there is no canonical choice of such a chain.

We will give only a typical example of the construction of such a graph.

Let $B_n$ be the set of all elements of the group $G$ that can be written as words of length at most $n$ in the generators $S$; here $B_1 = S$. Observe that
\[
gB_n-1 \subset B_n \quad \text{if } g \in S.
\]
We introduce the following series of graded graphs $\Gamma_{G,S} \equiv \Gamma$.

The set $\Gamma_k$ of vertices at the $k$th level of the graph $\Gamma$, $k \geq 1$, consists of all functions on the ball $B_k$ with values in $k$; an edge joins a vertex $v \in \Gamma_k$ with a vertex $u \in \Gamma_{k-1}$ if $u$, regarded as a function on $B_{k-1}$, is the restriction of $v$, regarded as a function on $B_k$, to $B_{k-1} \subset B_k$. An equipment on the edge $(u, v)$ is defined in this case as the product of the numbers $p_g$ which corresponds to the product of the generators completing the word $u$ to the word $v$.

Thus, we have constructed an equipped graded graph $\Gamma_{G,S}$. We will call it the graph of words of the given group with generators $(G, S)$. This construction admits useful generalizations, for example, one can consider the graph of words which corresponds to the sequence of balls $B_{n_1}$, where $\{n_k\}$ is a sequence of positive integers going to infinity.

We define a homomorphism of the space $\mathcal{X}$ onto the path space $T(\Gamma_{G,S})$ of the constructed graph as follows: a trajectory of the process is a sequence of configurations $\{\phi_n\}_{n<0}$ on the group. Consider the map
\[
\Phi: \{\phi_n\}_n \to \{\phi_n|B_n\}_n.
\]
Theorem 11. 1) The image of the space $\mathcal{X}$ under the homomorphism $\Phi$ is the path space of the graph of words $\Gamma_{G,S}$.

2) The homomorphism $\Phi$ agrees with the equipment, that is, the parameters of the Markov chain (the transition probabilities) coincide with the cotransition probabilities on the path space. Hence $\Phi$ maps the measure $M$ to a measure on the path space that agrees with the equipment.

3) The homomorphism $\Phi$ of the space $\mathcal{X}$ onto the path space $T(\Gamma_{G,S})$ defines a homomorphism of the space of configurations, that is, trajectories of the Markov chain, onto the path space of the graph, and the filtration is homomorphically mapped with preservation of measure onto the tail filtration of the path space.

Proof. Follows immediately from our definition of the graph and its structure. $\square$

The answer to the important question of when this homomorphism is an isomorphism depends on the group, on the choice of a sequence $\{n_k\}_k$, and on the transition probabilities, more precisely, on whether the shifts of the balls in the random walk cover the whole group. In the general case one should increase the size of the balls, which dictates the choice of the sequence $\{n_k\}_k$. We will return to this question later, but now we consider concrete examples in the upcoming sections.

It is useful to observe that this construction can also be used for semigroups; for instance, one can take $S$ to be a set of generators without inverses. The resulting graph corresponds to the filtration of the process constructed from the group, but it is simpler, and it sometimes suffices to analyze this simpler graph in order to study the process itself.

Below we give examples of applying this scheme.

6.2. Walks over orbits of a free group. Historically the first example of a non-standard dyadic filtration [59] involved the free group $G = F_2$ with two generators $(a, b)$ and the space $X = 2^{F_2}$ endowed with a Bernoulli measure (the infinite product of the measures $(1/2, 1/2)$). We considered the random walk over the trajectories (orbits) of the Bernoulli action of the free semigroup with two generators, that is, the Markov process with the transitions $y(g) \to y(ag)$ and $y(g) \to y(bg)$ having the probabilities $(1/2, 1/2)$.

We cite the paper [59]: “The constructed Markov process has the following paradoxical property, which can be loosely stated as follows. Assume that the process of development of mathematics is the same as the constructed Markov process. Even though each year, starting from $-\infty$, the Abstract Journal of Mathematics\textsuperscript{13} publishes a volume containing all essentially new results (that is, results independent of previous results) discovered in the given year, reading all these volumes still does not enable one to fully recover the picture of mathematics for any given year. Of course, it is understood that in the year $-\infty$ there were no discoveries.”

Let us add that the above is true for any way of writing the volumes, that is, for any choice of an independent complement of the previous volumes. In other words, there are essential events (that is, events of positive measure) that occurred in the given year which are not measurable with respect to any system of independent complements of the partitions of this dyadic filtration.

\textsuperscript{13}Which, unfortunately, has not survived to modern times.
The construction of the graph described above for the general case determines an approximation of this process by finite Markov chains. One of the corresponding multigraphs looks as follows: the number of vertices at level \( n \) is equal to \( 2^{2^n} \), which corresponds to the set of all functions with values 0, 1 on all words of length \( n \). Every function-vertex \( v \) is joined by an edge with two functions-vertices \( u \) and \( u' \) at the previous level if \( u \) (respectively, \( u' \)) is the function on words of length \( n - 1 \) obtained from \( v \) by considering it on the indices beginning with \((0,*,*,\ldots,*))\) (respectively, \((1,*,*,\ldots,*))\). But since functions are arbitrary, this means that the set of vertices at the \( n \)th level is the set of all pairs of vertices at the \((n - 1)\)th level. In other words, the graph of words in this case is the \textit{graph of ordered pairs} \cite{82,91} (see Fig. 9).  

From here it is not difficult to deduce that the corresponding filtration is non-standard (see below) and to obtain a number of other properties. 

Exactly the same numerical analysis was used in \cite{59} to prove non-standardness of filtrations arising in the study of actions of locally finite groups, for example, a Bernoulli action of the group of roots of unity of orders of the form \( 2^n \), or a direct sum of groups of order 2. The main argument proving the non-standardness in this case was that the orders of the orbits of the groups of automorphisms of the dyadic tree in these examples were substantially smaller than the order sufficient for the measure to be concentrated near one orbit. This argument was used by the author to introduce the notion of the entropy of a dyadic filtration. After the author’s paper \cite{59}, in which non-standard dyadic filtrations and the standardness criterion first appeared together with the notion of entropy of dyadic filtrations (see \cite{69}), A.M. Stepin observed that in examples connected with actions of locally finite groups (like the infinite sum of groups of order 2), one can also use the entropy of the action as an invariant of a trajectory dyadic sequence. In \cite{64}, \cite{65}, \cite{69} an estimate of the growth of the orders of groups is given for which the entropy of the action coincides with the entropy of the filtration. Further research \cite{22} showed that this estimate is sharp. However, the entropy of filtrations in the sense of \cite{62}, \cite{69} is of a much more general nature. 

We also add that the tree of paths \( T_u \) with root at a given vertex \( u \) at level \( n \) together with the sequence of zeros and ones on the leaves of the tree indicating at which of the two first-level vertices \( a \) or \( b \) the given path ends, is exactly what in \cite{59} was called the \textit{universal projection} of the cylinder set \( \xi_a \) or \( \xi_b \) of vertices at the first level on the initial fragment of the filtration, that is, on the first \( n \) partitions. The fact that the universal projection does not stabilize, that is, that for different vertices \( u \) at level \( n \) the trees do not approach one another in measure, means, according to the standardness criterion, exactly that the filtration is non-standard.

In the conclusion of this example, we note that although (by the theorem on a Markov realization) filtrations generated by different RWRS actions can be represented as filtrations of Markov chains with finitely many states, an explicit form of this realization is known only in a small number of examples. This is also true for adic realizations of the corresponding groups and semigroups. Such a realization is not even known for all Abelian groups, which is discussed in the next section.

\footnote{The possibility of using this graph in the study of random walks on free semigroups and the corresponding filtrations was not mentioned in \cite{91}, and this connection will be considered in another paper.}
Figure 9. The graph of ordered pairs.
6.3. Random walks over orbits of free Abelian groups. Walks over orbits of infinite Abelian groups provide an example of a filtration that is simpler in one respect and more complex in another. We will also consider a model example.

If $G = \mathbb{Z}^d$ and $S = (e_1, e_2, \ldots, e_d)$, where $e_i$ are the unit coordinate vectors, then the corresponding graph looks as follows: the set of vertices at level $n$ is the set of all functions with values 0 or 1 on the integer simplex

$$\Sigma^n_d = \left\{\{r_1, r_2, \ldots, r_d\} : \sum_i r_i = n, r_i \in \mathbb{Z}_+\right\},$$

and the edges join a vertex-function $u$ with vertices-functions at the previous level that are restrictions of $u$ to subsimplices obtained by removing one of the faces.

Consider the case when $d = 1$, $G = \mathbb{Z}$, $S = \{-1, +1\}$, and the corresponding graph is the graph of words for $\mathbb{Z}$. The vertices at the $n$th level are the words of length $n$,

$$(\epsilon_1, \epsilon_2, \ldots, \epsilon_n), \quad \epsilon_i \in 0, 1,$$

and the corresponding edges lead to the two subwords

$$(\epsilon_1, \epsilon_2, \ldots, \epsilon_{n-1}), \quad (\epsilon_2, \ldots, \epsilon_n), \quad n = 2, 3, \ldots.$$

The graph of words for $\mathbb{Z}$ is an Abelian analog of the graph of ordered pairs. This is a dyadic (multi)graph, and the $n$th level has $2^n$ vertices, each having two predecessor vertices and four successor vertices (see Fig. 10). The natural central measure is the uniform measure on the vertices at each level.

![Figure 10. The graph of binary words for the group $\mathbb{Z}$.](image)

The adic shift on this graph, somewhat resembling the $(T, T^{-1})$-transformation (where $T$ is a Bernoulli automorphism), is quite interesting. In the well-known paper [27], it is proved that the former transformation is not Bernoulli and even not weakly Bernoulli. But here we are interested in the properties of the filtration; the relation of these properties to ergodic characteristics of transformations has not yet been sufficiently studied. The non-standardness of the filtration for $(T, T^{-1})$ was conjectured by the author and first proved in [22] (see also [84]). An extension of the results in [27] to the groups $\mathbb{Z}^d$ was obtained in [4]. The non-standardness of the filtrations for all free Abelian groups with $d \geq 1$ and for nilpotent groups
was proved in [84], along with the fact that the scaled entropy of these filtrations depends on \( d \). We will return to these graphs and their generalizations to other groups in subsequent papers.

6.4. Lacunarity. The theorem on lacunary isomorphism [58] says that every dyadic ergodic filtration \( \{\xi_n\}_n \) is lacunary standard, that is, there exists an increasing subsequence \( \{n_k\} \) of positive integers such that the filtration \( \{\xi_{n_k}\}_k \) is a \( \{2^{n_k-n_k-1} \equiv r_k\} \)-adic standard filtration. The sequences \( \{n_k\} \) for which this is true constitute the so-called scale, which is an invariant of the original dyadic filtration. The scale of a standard dyadic sequence consists of all sequences and is said to be complete. Of course, together with any sequence the scale contains all sequences cofinal with it (that is, differing from it in only finitely many positions), and the scale is monotone, since a subsequence of a standard filtration is obviously standard. It seems that these two properties exhaust the conditions on a scale. In [69], [40] non-standard dyadic sequences with various scales are constructed. For instance, \( \{\xi_{2k}\} \) can be a standard \( 4\)-adic filtration, while the filtration \( \{\xi_k\}_k \) is non-standard, and so on.

By the same method one can prove a general theorem on arbitrary inhomogeneous filtrations, not even necessarily locally finite (see [82]).

**Theorem 12.** Every ergodic filtration is lacunary standard.

Since standardness can be violated along an arbitrary sequence, classification of filtrations can hardly be efficient.

However, the role of lacunary theorems is different: they measure the degree of closeness of a filtration to a standard one.

In [64] the notion of an automorphism with a complete (dyadic) scale was introduced; in terms of the present paper, this is an automorphism for which there exists a dyadic realization with a standard tail filtration (with respect to the given measure).\(^{15}\) Later, Katok [28] (see also [29]) introduced and studied the notion of a standard automorphism as an automorphism that is monotonically equivalent to an automorphism with universal discrete spectrum. It would be interesting to find out whether these notions coincide.

6.5. Scaled and secondary entropy. Another measure of closeness to standardness is the scaled entropy of a filtration (and its analog, the scaled entropy of a group action).

Recall that the condition of the standardness criterion is that the metric \( d^n_\rho \) on the space of measured trees converges in measure to a degenerate metric.

**Proposition 7.** If a filtration is ergodic, then the sequence of metrics \( d^n_\rho \) either converges in measure to a degenerate metric, or has no limit.

\(^{15}\)In [64] there was an attempt to apply results on dyadic filtrations (standardness, scale, and so on) to the study of ergodic properties of automorphisms. However, in contrast to such applications to actions of locally finite groups, for groups that are not locally finite the requirement of dyadicity or homogeneous periodicity in the spirit of Rokhlin’s lemma complicates all the constructions. As further development of the topic showed, one should pass from homogeneous filtrations to semihomogeneous ones; this is the way that adic realizations of group actions are constructed, and the notions of standardness and complete scale become more natural.
The theory of filtrations of subalgebras

The reason for this simple but important fact is that the existence of a limit that does not reduce to a one-point space contradicts the ergodicity, since otherwise non-trivial limiting functionals would appear.

Therefore, in the case of a non-standard filtration there arises a sequence of metric spaces that has no limiting metric space, but one can follow the \( \epsilon \)-entropy of these metric spaces as a function of \( n \). It turns out that the asymptotic behaviour of this sequence under a certain normalization does not depend on the choice of a metric, and thus its class is an invariant of the filtration. It is in this way that this invariant was defined in [65], [59] in the case of a dyadic filtration with exponential (Kolmogorov) normalization. In [69], [84] it was defined in the general case and, for reasons related to the stationary case, called the secondary entropy. In the sense of this definition and under a correct normalization, the secondary entropy of the filtration corresponding to the transformation \((T, T^{-1})\) equals the entropy \( h(T) \) by a theorem in [23]. In [91] it was proved that the scaled entropy of the adic automorphism for certain measures on the graph of ordered pairs is equal to the entropy (scaling) sequence (with respect to the same measure) of the tail filtration. These theorems should be extended to arbitrary groups.

6.6. Project of classification of non-standard filtrations. The Kolmogorov zero–one law (on the triviality of the \( \sigma \)-algebra of the intersection of the \( \sigma \)-algebras of a filtration) is the simplest condition on the behaviour of a filtration at infinity. Standardness, or the Bernoulli property, is, on the other hand, the strongest such condition: it guarantees that at infinity there is no complexity, since all conditional distributions converge in the strongest sense. Here, in contrast to the traditional way of measuring the closeness of a process to a process of independent random variables, this closeness is expressed not by the rate of convergence to zero of the correlation coefficients or other characteristics, but by the smallness of the difference (in one or another metric) between the tree structures of the conditional measures. Hence, one can speak about ‘higher zero–one laws’ (meaning assertions that conditional distributions approach each other in some metric) which are intermediate between the usual zero–one law and standardness (the Bernoulli property).

As observed in the previous section, such a scale of conditions apparently exists.

Ornstein’s very weak Bernoulli (VWB) condition is also a law of this type, but as shown by examples, it does not coincide with the standardness condition: in [16], an example is given of a non-Bernoulli process with a standard ergodic dyadic filtration, and our examples of non-standardness show that a Bernoulli automorphism can have a Markov generator with respect to which the past is not standard; another example is given in [23]. However, these are only examples, and a more complete analysis of the situation is yet to be carried out. We remark also that conditions may depend on what metric we choose on the space of measured trees.

Recall that the scheme of the standardness criterion involves considering functions on \( X/\xi_n \) with values in the space of measured trees of rank \( n \), and the criterion is that for sufficiently large \( n \) the values of the functions approach one another with respect to the metric on the space of trees, that is, the sequence of functions converges to a constant in measure. In other words, the images of the measures on the

\[\text{The definition of this entropy was given in the author’s papers [73, 74], where its main properties were formulated, and they were proved later by Zatitskiy [100, 101].}\]
spaces of trees converge to a $\delta$-measure. Instead of functions on the quotient space $(X/\xi_n, \mu_\xi)$, it is convenient to consider their liftings to the space $(X, \mu)$ itself, and then on $(X, \mu)$ we obtain semimetrics $\rho_n$, in which the distance between points is the distance between their images, the values of the functions, that is, the corresponding trees.

The condition of the standardness criterion is that the sequence of these metrics collapses to a point, and if this condition is satisfied, then two finitely isomorphic filtrations are isomorphic. If there is no standardness, then the functions do not converge to a constant in measure, and this means that the sequence $(X, \mu, \rho_n, n = 1, 2, \ldots)$, where $\rho_n$ is the sequence of semimetrics, does not converge and there is no limiting metric on $(X, \mu)$.

What criteria can we use to establish an isomorphism of filtrations? Necessary conditions are the coincidence of the entropy, of the scale, and so on (see [64]). But can we state sufficient conditions? For this, one should understand what it means that two sequences of (semi)metric spaces for which there is no limiting space are asymptotically isomorphic.

The conjecture we formulate below is that sometimes finite isomorphism supplemented by a similar asymptotic isomorphism implies isomorphism. This conjecture reduces the study of all invariants of non-standard filtrations (like, for example, entropy) to a unified context.

To state this conjecture more precisely, we recall the following general idea about asymptotic isomorphism proposed in [82].

Consider a so-called metric triple, or in another terminology, an admissible triple $\theta = (X, \mu, \rho)$, that is, a space $X$ endowed with a measure $\mu$ and a metric $\rho$ that agree with each other: the metric $\rho$ is measurable as a function of two variables on the space $(X, \mu)$, the space $(X, \rho)$ is separable as a metric space, and $\mu$ is a Borel probability measure on this metric space $(X, \rho)$. By the way, note that, in contrast to the traditional practice of fixing a metric on a space and considering various Borel measures on the resulting metric space, the author has repeatedly advocated the usefulness of another approach: fixing a measure and varying metrics on the resulting measure space [71].

In [19], [71], [72] a theorem is proved on a complete system of invariants of metric triples with respect to the group of measure-preserving isometries.

**Definition 9.** The matrix distribution $D_\theta$ of a metric triple $\theta = (X, \mu, \rho)$ is the measure in the space of infinite non-negative real matrices defined as the image of the Bernoulli measure $\mu^\infty$ on $X^\infty$ under the map

$$M_\theta: (X^\infty, \mu^\infty) \to (\mathcal{M}_\mathbb{N}(\mathbb{R}_+), D_\theta),$$

$$\{x_n\}_n \mapsto \{\rho(x_i, x_j)\}_{(i,j)}.$$ 

**Theorem 13.** Two metric triples

$$\theta = (X, \mu, \rho) \quad \text{and} \quad \theta' = (X', \mu', \rho')$$

with non-degenerate measures ($\mu(U) > 0$ and $\mu'(U') > 0$ if $U$ and $U'$ are non-empty open sets in $X$ and $X'$) are isomorphic with respect to the above equivalence (that is, there exists an invertible isometry $T: X \to X'$ preserving the measure: $T\mu = \mu'$)
if and only if their matrix distributions coincide:

$$D_\theta = D_{\theta'}.$$ 

In [70], [76], [85] a similar theorem is proved on the classification of arbitrary measurable functions of several variables via matrix distributions.

**Definition 10.** Assume that for a sequence of admissible metric triples $$(X, \mu, \rho_n)$$ on the same measure space there exists a weak limit $\lim_n D_{\theta_n}$ of matrix distributions (regarded as measures on the space of matrices $M_{\mathbb{R}}(\mathbb{R}_+)$ with the usual weak topology). This limit will be called the virtual matrix distribution of the sequence of metric triples. It is an invariant with respect to a sequence of isometries $I_n$, $n = 1, 2, \ldots$.

We note that if the metric triples do not converge to any metric triple, then the limit of the matrix distributions, if it exists, is not the matrix distribution of any metric triple. On the other hand, the set of all matrix distributions of metric triples is not weakly closed, hence one can use virtual matrix distributions for classifying some diverging sequences of metric spaces.

Now we can state a conjecture on isomorphism of filtrations.

**Conjecture 1.** Consider two finitely isomorphic ergodic filtrations $\tau, \tau'$ and assume that:

1) they both are non-standard;

2) they both have virtual matrix distributions of the sequences of (semi)metric triples.

If these matrix distributions coincide, then the filtrations are isomorphic.

The validity of this conjecture, along with the standardness criterion, would allow us to believe that the classification of non-standard filtrations has been completed in the above terms and under the above assumption on the validity of the condition 2) for an arbitrary non-standard filtration. The answer would be that the virtual matrix distribution together with the finite invariants exhaust all invariants of ergodic locally finite filtrations. However, the condition 2) (more precisely, the existence of a virtual matrix distribution) can hardly always be satisfied.

Hence, the main problem is to find a complete system of invariants for a diverging sequence of metrics on a measure space. For the dyadic case, it suffices to consider invariants of sequences of metrics arising from iterations of metrics on the graph of unordered pairs, which we discuss in the next section.

But in the example in §6.2 of a Markov chain with a non-standard filtration, the virtual matrix distribution does exist. Namely, the sequence of matrix distributions of triples on the two-point space,

$$\rho_n(a, b) = 0, \quad n \equiv 0 \pmod{2}, \quad \rho_n(a, b) = 1, \quad n \equiv 1 \pmod{2},$$

converges in the sense of Cesàro to the virtual matrix distribution of this sequence, which is a measure on the set of infinite symmetric matrices with zeros on the diagonal and with independent entries $r_{i,j}$, $i > j$, taking the values 0 and 1 with equal probabilities.
7. Projective limits of simplices, invariant measures, and absolute of a graph

This section is devoted to the theory of filtrations in Borel spaces: in this case, there is no measure a priori, but only an equipment, that is, conditional measures. The main problem is to describe all Borel probability measures that agree with the equipment.

We describe a fourth language of the theory of locally finite filtrations: the language of projective limits of finite-dimensional simplices. One may say that the category of projective limits of finite-dimensional simplices is equivalent to the category of locally finite filtrations in Borel spaces, or to the category of multigraphs with equipments, or to the category of Markov compacta with cotransition probabilities. Although this is an extremely natural relation, nevertheless, as far as the author knows, the functorial equivalence of the theory of Borel and metric filtrations on the one hand and the convex geometry of projective limits of simplices on the other hand has until recently scarcely been discussed at all in the literature in this context. We use some definitions and facts from the author’s recent paper [78] (see also [36], [5], [98], [52]). The analysis of the theory of locally finite filtrations in Borel spaces, as well as the theory of central and invariant measures on path spaces of multigraphs, or the theory of Markov measures with given cotransitions, shows that the adequate and geometric language of projective limits of finite-dimensional simplices is the most appropriate language for all these theories. Below we will describe a correspondence between the problem of describing the measures that agree with an equipment and, in particular, of describing the central measures, and the geometry of projective limits of simplices (for more details, see [78]).

7.1. Projective limits of simplices, extreme points. First, we will show how, given a pair \((\Gamma, \Lambda)\), that is, an equipped graded graph, one can define a projective limit of finite-dimensional simplices in a canonical way. Later we will see that there is also a reverse transition from projective limits to equipped graphs.

Denote by \(\Sigma_n\) the finite-dimensional simplex of formal convex combinations of vertices \(v \in \Gamma_n\) at level \(n\). It is natural to view the simplex as the set of all probability measures on the set of vertices of \(\Gamma_n\). We define affine projections

\[ \pi_{n,n-1}: \Sigma_n \to \Sigma_{n-1}; \]

it suffices to define them for every vertex \(v \in \Gamma_n\). Obviously, these projections can be regarded as a system \(\Lambda\) of cotransition probabilities, and the images of vertices \(v\) are points of the previous simplex, that is, probability vectors:

\[ \pi_{n,n-1}(\delta_v) = \sum_{u: u < v} \lambda_v^u \delta_u; \quad (7.1) \]

this map can be extended by linearity to the whole simplex \(\Gamma_n\). The vertex \(\emptyset\) gives rise to the zero-dimensional simplex consisting of a single point. Degeneracies are allowed (that is, projections may glue together different vertices). Now we define
projections of simplices with arbitrary numbers:

\[ \pi_{n,m} \colon \Sigma_n \to \Sigma_m, \quad m < n; \]
\[ \pi_{n,m} = \prod_{i=m}^{n+1} \pi_{i,i-1}, \quad m > n, \quad m, n \in \mathbb{N}. \]

The collection of data \( \{ \Sigma_n, \pi_{n,m} \} \) allows one to define a projective limit on the one hand, and to recover the graph on the other hand: the vertices of the graph \( \Gamma_n \) are the vertices of \( \Sigma_n \), and the edges (and then also the paths in the graph) are recovered from the non-zero coordinates of the vectors \( \pi_{n,n-1}, n \in \mathbb{N} \).

Let \( \mathcal{M} = \prod_{n=0}^{\infty} \Sigma_n \) be the direct product of the simplices \( \Sigma_n, n \in \mathbb{N} \), endowed with the product topology.

**Definition 11.** The space of the projective limit of the family \( \{ \Sigma_n \}_n \) of simplices with respect to the system of projections \( \{ \pi_{n,m} \} \) is the following subset of the direct product \( \mathcal{M} \):

\[
\lim_{n \to \infty} (\Sigma_n, \pi_{n,m}) \equiv \{ \{ x_n \}_n; \; \pi_{n,n-1}(x_n) = x_{n-1}, \; n = 1, 2, \ldots \} \\
\equiv (\Sigma_{\infty}, \Lambda) \subset \prod_{n=0}^{\infty} \Sigma_n = \mathcal{M}.
\]

**Proposition 8.** The space of the projective limit \( \Sigma_{\infty} \) is always a non-empty, convex, closed (and hence compact) subset of the direct product \( \mathcal{M} \), and is a (possibly infinite-dimensional) Choquet simplex.

The affine structure of the direct product \( \mathcal{M} \) determines an affine structure on the limiting space; the fact that the set is non-empty and closed is obvious. It remains to check only that an arbitrary point of the limit has a unique decomposition over its Choquet boundary (see below).

We distinguish between the space of the projective limit and the ‘structure of the projective limit’, meaning that not only the limiting space itself, that is, some infinite-dimensional simplex, is important for us, but also the structure of the pre-limit simplices and their projections. In other words, we consider the category of projective limits of simplices in which an object is not a limiting simplex, but rather a sequence of finite-dimensional simplices.

Let us show that the collection of data of a projective limit of simplices allows one to recover the graph, the path space, and the system of cotransition probabilities, and the projective limit corresponding to this system and constructed according to the above rule coincides with the original one. This will establish a tautological relation between the two languages: the language of pairs (a Bratteli diagram, a system of cotransitions) on the one hand, and the language of projective limits of finite-dimensional simplices on the other hand.

Indeed, suppose we have a projective limit of finite-dimensional simplices \( \{ \Sigma_n \}, n \in \mathbb{N} \), and a coherent system of affine projections

\[ \{ \pi_{n,m} \}; \quad \pi_{n,m} \colon \Sigma_n \to \Sigma_m, \quad n \geq m, \quad n, m \in \mathbb{N}. \]

Take the vertices of \( \Sigma_n \) as the vertices at the \( n \)th level of the graph \( \Gamma \); a vertex \( u \) at level \( n \) precedes a vertex \( v \) at level \( n + 1 \) if the projection \( \pi_{n+1,n} \) sends \( v \) to
a point of $\Sigma_n$ for which the barycentric coordinate with respect to $u$ is positive. As a system of transition probabilities, we take the system of vectors $\{\lambda^u_n\}$ connected with the projections $\pi_{n+1,n}$ by the formula (7.1).

In what follows, for a given projective limit of simplices we will use the graph (of vertices of all simplices) canonically associated with it, the space of paths in this graph, and so on, and in the same way we will speak about projections of simplices canonically associated with an equipped graph.

Consider an arbitrary projective limit of finite-dimensional simplices:

$$\Sigma_1 \leftarrow \Sigma_2 \leftarrow \cdots \leftarrow \Sigma_n \leftarrow \Sigma_{n+1} \leftarrow \cdots \leftarrow \Sigma_{\infty} \equiv \Sigma(\Gamma, \Lambda).$$

First of all, we define the limiting projections

$$\pi_{\infty,m}: \Sigma_{\infty} \to \Sigma_m$$

as the limits $\lim_n \pi_{n,m}$ for each $m$: obviously, the images $\pi_{n,m}\Sigma_n$, regarded as subsets of the simplices $\Sigma_m$, monotonically decrease as $n$ grows, and their intersections are some sets denoted by

$$\Omega_m = \bigcap_{n: n > m} \pi_{n,m}\Sigma_n;$$

these are convex closed subsets of the finite-dimensional simplices $\Sigma_m$, $m = 1, 2, \ldots,$ and the limiting projections send the limiting compactum $\Sigma_{\infty}$ epimorphically onto these sets:

$$\pi_{\infty,m}: \Sigma_{\infty} \to \Omega_m.$$

It would be more economical to consider also the projective limit

$$\Omega_1 \leftarrow \Omega_2 \leftarrow \cdots \leftarrow \Omega_n \leftarrow \cdots \leftarrow \Omega_{\infty} = \Sigma(\Gamma, \Lambda)$$

itself with the epimorphic projections $\pi_{n,m}$ restricted to the subsets $\Omega_n$ and, by definition, with the same limiting space. However, to find the sets $\Omega_n$ explicitly is an interesting and difficult problem equivalent to the main problem of describing all invariant measures.\footnote{In particular, an explicit form of the compacta $\Omega_n$ is known in very few cases, even among those where the central measures are known. Even for the Pascal graph, we obtain interesting and rather complicated convex compacta, but, for example, for the Young graph the author does not know the form of these compacta with the same degree of precision.}

Every point of the limiting compactum, that is, a sequence

$$\{x_m\}: x_m \in \Sigma_m, \quad \pi_{m,m-1}x_m = x_{m-1},$$
determines, for each $m$, a sequence of measures $\{\nu^m_n\}$ on the simplex $\Sigma_m$, namely,

$$\nu^m_n = \pi_{n,m}(\mu_n),$$

where the measure $\mu_n$ is the (unique) decomposition of the point $x_n$ into extreme points of the simplex $\Sigma_n$. Of course, the barycenter of each of the measures $\nu^m_n$ in $\Sigma_m$ is $x_m$, and this sequence of measures itself becomes coarser, in a clear sense, and converges weakly in $\Sigma_m$ as $n \to \infty$ to a measure $\nu_{x_m}$ supported by $\Omega_m \subset \Sigma_m$.\footnote{In particular, an explicit form of the compacta $\Omega_n$ is known in very few cases, even among those where the central measures are known. Even for the Pascal graph, we obtain interesting and rather complicated convex compacta, but, for example, for the Young graph the author does not know the form of these compacta with the same degree of precision.}
Obviously, in this way we obtain all points of the limiting compactum, that is, all measures with given cotransition probabilities.

A point of an arbitrary convex compactum $K$ is said to be extreme if it cannot be represented by any non-trivial convex combination of points of $K$; the collection of extreme points is called the Choquet boundary of $K$ and denoted by $\text{ex}(K)$. A point is said to be almost extreme if it lies in the closure $\text{ex}(K)$ of the Choquet boundary. We recall that an affine compactum in which every point has a unique decomposition into extreme points (over the Choquet boundary) is called a Choquet simplex.

Let us give general criteria of extremality and almost extremality for points of a projective limit of simplices.

**Proposition 9.** 1) A point $\{x_n\}$ of a projective limit of simplices is extreme if and only if for every $m$ the weak limit of the measures $\nu_{x_m}$ in the simplex $\Sigma_m$ is the $\delta$-measure at the point $x_m$:

$$\lim_n \nu^n_m \equiv \nu_{x_m} = \delta_{x_m}.$$ 

2) A point $\{x_n\}$ is almost extreme if for every $m$ and every neighbourhood $V(x_m)$ of the point $x_m \in \Sigma_m$ there exists an extreme point $\{y_n\}$ of the limiting compactum for which $y_m \in V(x_m)$.

3) For every point $\{x_n\}$ of the limiting compactum of simplices there exists a unique decomposition into extreme points (Choquet decomposition), which is determined by the measures $\nu_{x_m}$.

**Corollary 3.** The limiting compactum of a projective limit of finite-dimensional simplices is a Choquet simplex (in general, infinite-dimensional).

In the literature (see [98]) the question has been discussed of why a limit of simplices is a simplex; even in the finite-dimensional case this is not quite obvious (it was raised by Kolmogorov). The most correct way to explain this is to refer to the fact, formally much more general but actually equivalent (see, for example, [52]), that the decomposition of measures with a given cocycle into ergodic components is unique, that is, to use the probabilistic interpretation of simplices.

It is easy to prove that the converse is also true: every separable Choquet simplex can be represented as a projective limit of finite-dimensional simplices, but of course such a representation is not unique.

Slightly digressing from the main focus, we note that the simplex of invariant measures with respect to an action of a non-amenable group on a compactum is separable, though a possible approximation of it is not generated by finite approximations of the action. Therefore, there arises a non-trajectory finite-dimensional approximation of the action which apparently has not been considered in the literature.

**Remark 1.** The first two assertions of Proposition 9 can perhaps be extended to projective limits of arbitrary convex compacta.

We recall that the class of separable Choquet simplices contains Poulsen simplices, for which the set of extreme points is dense; such a simplex is unique up to an affine isomorphism and universal in the class of all affine separable simplices (see [72] and the references therein).
Proposition 10. Consider a projective limit of simplices with the following property: for every \( m \) the union of the projections
\[
\bigcup_{n,t} \{ \pi_{n,m}(t); \ t \in \text{ex}(\Sigma_n), \ n = m, m + 1, \ldots \}
\]
over all vertices of the simplices \( \Sigma_n \) and all \( n > m \) is dense in the simplex \( \Sigma_m \). Then the limiting simplex is a Poulsen simplex.

It is clear that this construction can be carried out by induction, and the extremality criterion obviously implies that in this case the extreme points are dense. With such arbitrariness in the construction, the uniqueness seems surprising. Nevertheless, verifying that a given projective limit is a Poulsen simplex is not so easy and is similar to problems related to filtrations.

A simplex whose Choquet boundary is closed is called a Bauer simplex. Between the classes of Bauer and Poulsen simplices, there are many intermediate types of simplices. This topic has been repeatedly discussed in the literature on convex geometry and the theory of invariant measures. However, it seems that these and similar properties of infinite-dimensional simplices in application to projective limits and the theory of graded graphs and corresponding algebras have never been considered. Each of these properties has an interesting interpretation in the framework of these theories. The author believes that the following class of simplices (or even convex compacta) is useful for applications: an almost Bauer simplex is a simplex whose Choquet boundary is open in its closure.

The parallelism between the study of pairs \((\Gamma, \Lambda)\) (a graded graph and an equipment of this graph) on the one hand, and of projective limits of simplices on the other hand means that the latter has a probabilistic interpretation. It is useful to describe this interpretation directly. Recall that in the context of projective limits a path is a sequence \( \{ t_n \}_n \) of vertices \( t_n \in \text{ex}(\Sigma_n) \) that agrees with the projections in the following way: for all \( n \in \mathbb{N} \) the point \( \pi_{n,n-1} t_n \) has a non-zero barycentric coordinate with respect to the vertex \( t_{n-1} \). First of all, every point \( x_\infty \in \Sigma_\infty \) of the limiting simplex is a sequence of points of simplices that agrees with the projections:
\[
\{ x_n \}; \ \pi_{n,n-1} x_n = x_{n-1}, \quad n \in \mathbb{N}.
\]
As an element of a simplex, \( x_n \) defines a measure on the vertices of this simplex, and since all these measures agree with the projections, \( x_\infty \) defines a measure \( \mu_x \) on the path space with fixed cotransition probabilities. Conversely, each such measure defines a sequence of projections. Thus, the limiting simplex is the simplex of all measures on the path space with given cotransitions.

If a point \( \mu \in \text{ex}(\Sigma_\infty) \) is extreme, this means that the measure \( \mu \) is ergodic, that is, the tail \( \sigma \)-algebra is trivial with respect to the measure \( \mu \) on the path space.

We end this section with the following conclusion.

The theory of equipped graded multigraphs and the theory of Markov compacta with systems of cotransitions are identical to the theory of Choquet simplices regarded as projective limits of finite-dimensional simplices. All three theories can be viewed as theories of locally finite filtrations in a fixed basis representation.
7.2. The fundamental problem: description of the absolute.

**Main problem.** Given a projective limit of finite-dimensional simplices, find the Choquet boundary of the limiting simplex.

It will be convenient to call this boundary (in the context of path spaces of graded graphs) the absolute of the projective limit (of a Markov chain, of a graded graph, and so on). Sometimes the absolute consists of a single point, in which case it is trivial, but this happens very rarely.

By the foregoing, this problem is equivalent to the following one:

- given a Markov compactum with a system of cotransitions, describe all ergodic measures that agree with these cotransitions,

or

- given a graded graph with an equipment, describe all ergodic measures on its path space that agree with this equipment,

or, finally,

- given a filtration on a Borel space, describe all ergodic Borel measures that agree with the cocycle of this filtration (see §3).

Actually, as we have seen, all these problems are tautologically equivalent, but still they belong to different areas of mathematics. The formulation with Markov chains is essentially due to Dynkin (see [11]).

For graphs with the canonical equipment, this is the problem of describing the central measures, or the traces of $AF$-algebras, or the characters of locally finite groups; moreover, this is also the problem of describing the invariant measures for actions of countable groups, provided that an adic approximation is chosen for them. This is an enormous circle of problems which covers a substantial part of asymptotic representation theory, combinatorics, algebra, and, of course, the theory of dynamical systems.

In the case of filtrations, fixing some ergodic measure allows one to regard a filtration as a filtration of a measure space (rather than of merely a Borel space), with all the problems considered above, whose solution substantially depends on the choice of a specific measure. But the absolute, in a sense, combines different measures with respect to which one can consider filtrations, equivalence relations, adic actions, and so on. Therefore, the problem is not only to describe the absolute. Here is a list of natural important additional questions (we state them in terms of graded graphs).

The setting of the problem can be, for instance, as follows:

- On which graded equipped graphs does every measure in the absolute determine a standard tail filtration on the path space endowed with this measure (it is natural to call such graphs standard)?
- On which graphs does there exist a measure in the absolute with a standard tail filtration?
- On which graphs does the absolute contain no non-trivial measures with a standard tail filtration?

The same series of questions can be posed in connection with entropy and other invariants of filtrations. From the point of view of the metric theory of filtrations

---

18In [11] the terms ‘exit boundary’, ‘entrance boundary’, and so on, were used. Note that the Martin boundary can also be easily defined in our context (see [78]).
and the theory of Markov chains, these questions look especially unconventional, since we are asking how dissimilar measures with the same conditional (cotransition) probabilities can be.

7.3. Several examples. Connection with the theory of locally finite filtrations. We choose several examples of problems of describing the absolute.

1. Absolutes of groups. Consider a countable finitely generated group $G$ and choose a symmetric system of generators: $S = S^{-1}$. By the dynamical Cayley graph, or the Markov compactum of a random walk, we mean the graph in which the vertices at level $n$ are all elements of $G$ representable as words of length at most $n$, and the edges correspond to right multiplication by generators. The equipment is canonical, that is, the set of word representations of a given element is endowed with the uniform measure. The problem of describing the absolute, that is, the set of ergodic central measures, is a considerable extension of the problem of describing the Poisson–Furstenberg boundary of a random walk. While the Poisson–Furstenberg boundary reduces to a single point for many groups, the absolute is very rarely a single point; it has the structure of a generalized fibre bundle over the Poisson–Furstenberg boundary. For a free group the absolute is calculated in [90]. In this case there is a natural phase transition from ergodicity to non-ergodicity. The absolute for nilpotent groups is presently under study. A quite useful tool here is a generalization of the ergodic method (see [82]) which reduces finding the central measures to establishing their additional symmetry properties. In this sense, describing the absolute can be viewed as a far-reaching generalization of de Finetti’s theorem on measures invariant with respect to groups of permutations. In all group examples considered so far, the filtrations turned out to be standard, that is, the dynamical graph can be said to be standard.

2. On the other hand, graphs arising in the approximation theory of dynamical systems (that is, graphs providing adic realizations of automorphisms with an invariant measure) are often non-standard.

Conjecture 2. A graph providing an adic realization of an automorphism with positive entropy is not standard. The corresponding limiting simplex is a Poulsen simplex.

The graph of ordered pairs (see § 4) has many central measures with respect to which the tail filtration is not standard (see [91]).

Another, more interesting, example of this type is the ‘tower of measures’ graph.

This is the graph of unordered pairs, which appeared in 1969 as a tower of binary measures. The simplest way to introduce this graph is to define it as a projective limit of simplices in which the set of vertices of each next simplex is the set of all vertices and middle points of all edges of the previous simplex, the initial simplex being the interval $[0, 1]$ (see Fig. 11). The limiting simplex provides nice interpretations of invariants of sets with respect to non-standard dyadic filtrations. Moreover, the definition of the tower of measures was motivated by an interpretation of the standardness criterion (more precisely, an interpretation of the values of the universal projection of a set) for non-standard situations. The construction of the tower of measures is related to constructions of universal spaces in set-theoretic and algebraic topology.
3. An enormous supply of graded graphs, and hence of absolutes, filtrations, and their invariants, is provided by the representation theory of AF-algebras and locally finite groups. Each of these objects gives rise to its own Bratteli diagram, the canonical equipment determines a projective limit of simplices and the limiting simplex of central measures, and so on. These connections have been studied (but so far without taking into account the role of tail filtrations) both in Borel path spaces and in path spaces endowed with central measures. We mention only that the computation of the absolute, that is, of the traces and characters, looks isolated in each specific case; up to now, there are no sufficiently general methods for solving this problem. For example, no one has yet managed to prove the well-known Kerov conjecture [30]–[32] on central measures in the Macdonald graph, though it differs little from known solved problems. An attempt to separate solvable problems related to the absolute from ‘intractable’ ones, that is, those that have no reasonable solution, was made in [77], [79], [82], and it is also reproduced in part in this paper. This is the reason for bringing the notion of standardness of filtrations into these problems. In the papers mentioned (in more detail, in [82]) the so-called internal metric on paths of a graph is introduced, with the aim of separating the two cases. For want of space, we postpone a further analysis of these connections to future publications.

8. Historical commentary

In concluding this survey it is appropriate to add several remarks on the history of the problem and some papers, examples, and counterexamples useful in the study of filtrations.

1. In 1958 the famous Kolmogorov paper [34] appeared, containing a definition of entropy as an invariant of the metric theory of dynamical systems. In that paper Kolmogorov defined entropy in the spirit of Shannon’s concept of information transmission, basing it on a filtration (in the older terminology, the ‘sequence of past $\sigma$-algebras’) that he had often considered earlier, and acknowledging the importance of Rokhlin’s theory of partitions in his considerations. As one can learn from the recollections of various people, Kolmogorov apparently proved in an earlier lecture course in 1957 (unfortunately, any notes were lost) that Bernoulli schemes with different entropies are non-isomorphic, using a combinatorial description of entropy which, due to the subsequent definition proposed by Sinai [53], later became the main one—all the more so because this definition was at once easily extended to
arbitrary groups.\textsuperscript{19} Kolmogorov’s definition in the first paper was much closer to the traditional theory of stationary processes and to Shannon’s ideas. But a small error in this paper compelled Kolmogorov to write a second paper (see [35]). Here we present an excerpt and footnote from this second paper, in which the error in the first paper was corrected,\textsuperscript{20} since this is directly related to difficulties in the theory of filtrations:

V. A. Rokhlin pointed out to me that the proof of Theorem 2 in my paper (1) tacitly uses the following assumption. It follows from

\[ \mathcal{A}_1 \supseteq \mathcal{A}_2 \supseteq \cdots \supseteq \mathcal{A}_n \supseteq \cdots, \quad \bigcap_n \mathcal{A}_n = \mathcal{N}, \]

that

\[ \bigcap_n (\mathcal{B} \vee \mathcal{A}_n) = \mathcal{B}. \]

By the courtesy of Rokhlin, I reproduce this example. Let \( \mathcal{G}^m \) be the additive group of numbers of the form \( am^{-\beta} \) (\( m \) is a positive integer, \( \alpha \) is an integer, \( \beta \) is a non-negative integer). Denote by \( U \) the automorphism of \( \mathcal{G}^6 \) that consists in dividing by 6, by \( M \) the group of characters of \( \mathcal{G}^6 \), and by \( T \) the automorphism of \( M \) conjugate to \( U \). The subgroups \( \mathcal{G}^2 \), \( \mathcal{G}^3 \) of \( \mathcal{G}^6 \) satisfy the following obvious relations:

\[ \mathcal{G}^2 \subset U \mathcal{G}^2, \quad \bigvee_n U^n \mathcal{G}^2 = \mathcal{G}^6, \quad \bigcap_n U^n \mathcal{G}^2 = 0, \]
\[ \mathcal{G}^3 \subset U \mathcal{G}^3, \quad \bigvee_n U^n \mathcal{G}^3 = \mathcal{G}^6, \quad \bigcap_n U^n \mathcal{G}^3 = 0. \]

Therefore, the corresponding subalgebras \( \mathcal{G}^2 \), \( \mathcal{G}^3 \) of the algebra \( \mathcal{G}^6 \) of measurable sets of the space \( M \) satisfy the quasi-regularity conditions. At the same time, \( \mathcal{G}^2 \) has index 3 in \( U \mathcal{G}^2 \) and \( \mathcal{G}^3 \) has index 2 in \( U \mathcal{G}^3 \), so that

\[ MH(T \mathcal{G}^2 | \mathcal{G}^2) = \lg 3, \quad MH(T \mathcal{G}^3 | \mathcal{G}^3) = \lg 2. \]

Here

\[ \bigcap_n (\mathcal{G}^2 \vee T^n \mathcal{G}^3) \neq \mathcal{G}^2, \quad \bigcap_n (\mathcal{G}^3 \vee T^n \mathcal{G}^2) \neq \mathcal{G}^3. \]

\textsuperscript{19}Here we should mention the master’s thesis by Arov (Odessa University, 1957), mentioned by Kolmogorov in his second paper [35] and published only recently (see [1]), which contains a combinatorial definition of an invariant of an automorphism close to Sinai’s definition. A full comparison of Kolmogorov’s entropy and Arov’s entropy was carried out by Gurevich [20].

\textsuperscript{20}For some reason, the publishers of Kolmogorov’s collected works decided not to publish this paper in full, but to combine a small part of it with the first paper.
This example shows the lack of continuity when passing from a filtration to its supremum with a given partition. It is interesting that this difficulty is absent in the theory of increasing sequences of partitions (σ-algebras). I mention this also because the theory of invariant partitions (in our terms, the theory of stationary filtrations) became a topic of interest for Rokhlin, Sinai, and many subsequent authors (see [47]). In our view, this deep subject is not yet exhausted, and its connections with the theory of groups of unitary or Markov operators and with scattering theory should be systematically studied. On the other hand, as one can see from the survey, the applicability of the theory of filtrations is much broader than the stationary case.

2. The question of whether there exists an ergodic filtration that is finitely isomorphic but not isomorphic to a Bernoulli filtration arose as far back as the 1950s, for example, in the work of Rosenblatt (see [49]). In slightly different terms, it was discussed in Wiener’s book *Non-linear problems in random theory* [97], in the chapters *Coding* and *Decoding*. Actually, *Decoding* contains a negative answer to this question along with tedious calculations. More precisely, it is claimed there that if a stationary Gaussian random sequence $\alpha_n$, $n < 0$, has the property that for each $n$ and for the σ-algebra $\mathcal{A}_n$ generated by the variables with indices $< n$ there exist $|n|$ Gaussian random variables independent of each other and independent of $\mathcal{A}_n$ that together with $\mathcal{A}_n$ generate the σ-algebra of all measurable sets, then the original stationary sequence is isomorphic to a one-sided Bernoulli shift. However, this is not true, since there exist non-standard stationary filtrations (which are thereby non-isomorphic to Bernoulli filtrations).

I must admit that in my first paper [58] I also claimed (for dyadic filtrations) that non-standard filtrations did not exist (without giving any supporting arguments or calculations), but I soon after corrected this error [59] and gave the first series of examples of non-standard dyadic sequences, together with the standardness criterion. Thus, classification of filtrations became a meaningful problem.

3. In speaking about the prehistory of the theory of filtrations, one should begin with the remark that the problem of analyzing the structure of stationary one-sided filtrations was raised in the 1960s by Rokhlin in the context of the study of endomorphisms (non-invertible measure-preserving transformations): many important results in this field are due to him (see [48]). He attracted to it Vinokurov and his Tashkent students (Rubshtein, Ganikhodzhaev, and others), who published a series of interesting papers on this subject (see [95], [94], [50], [51]). On the other hand, a direct need to investigate homogeneous (for example, dyadic) filtrations arose later in connection with the problem of trajectory isomorphism of automorphisms, also posed by Rokhlin irrespective of any connection with filtrations. It was successfully studied by Belinskaya [2]. In fact, by that time a positive solution in the problem of trajectory isomorphism of all ergodic transformations with an invariant measure had been already obtained by Dye [8], a fact that became known to the participants of Rokhlin’s seminar on ergodic theory only later, when I proved the theorem on lacunary isomorphism, which immediately implied Dye’s theorem. The techniques used in [8] were based on the theory of $W^*$-algebras, in contrast to an approximation proof (essentially combinatorial) in [58], [2]. Much later, the same result was proved
for all actions of amenable groups with an invariant measure [43], [3]. I was interested in trajectory theory in connection with the theory of factors even earlier, and some results for the more general case of quasi-invariant measures were obtained simultaneously with results of Krieger and Connes. For set-theoretic intersections of filtrations, I used the term ‘locally measurable partitions’, Rokhlin suggested the term ‘tame partitions’, and the generally accepted term is ‘hyperfinite partitions’.

4. Similar problems have also been posed in probability theory itself, long ago and repeatedly. For instance, the question of what stationary processes can be obtained by encoding Bernoulli schemes was studied by Rosenblatt, Itô, and many others. The clearest formulation of the problem was developed in the framework of information theory and ergodic theory: what automorphisms or endomorphisms are factors of Bernoulli automorphisms? For automorphisms, the problem was solved in the remarkable paper by Ornstein [42], who proved that a necessary and sufficient condition is the condition incongruously called VWB (very weak Bernoulli), while it would be more appropriate to call it ‘Ornstein’s condition’. For endomorphisms, it is not yet completely solved. Ornstein’s condition resembles in its nature the standardness condition, which appeared in another connection at approximately the same time [59], [66]. Their similarity is in the fact that both are requirements on the behaviour of the collection of conditional measures of finite segments of a process given a fixed past starting from some time, namely, one requires that these conditional measures approach one another as this time goes to $-\infty$. Obviously, any interpretation of a weakening of the independence condition can be expressed exactly in these terms, but still this method differs from various conditions of weak dependence which existed earlier in traditional schemes of the theory of random processes (Kolmogorov’s, Rosenblatt’s, Ibragimov’s conditions, and so on) and are usually stated as a decrease of correlation coefficients, measure densities, and so on. But in our approach, the leading role is played by the system of conditional measures and a metric that controls the convergence of systems of conditional measures. In Ornstein’s condition this is the Kantorovich metric (rediscovered by him specifically in this case) constructed from the Hamming metric on trajectories. In the case of standardness this is a modified Kantorovich metric taking into account the (tree-like) structure of the conditional measures. The difference between them is substantial, but the problems and the classes of processes are also different: Ornstein’s condition concerns stationary processes on $\mathbb{Z}$ (though is stated for their restrictions to $\mathbb{Z}_+$) and deals with the properties of a two-sided process, while standardness characterizes the properties of a one-sided filtration.\(^{21}\)

It is worth mentioning that instead of the Hamming metric (on trajectories), other metrics have also been considered, and the Kantorovich metric on measures constructed from a metric on trajectories can also be replaced by other ones: here there is wide room for experiments and analysis of asymptotic properties of processes and filtrations depending on metrics.

---

\(^{21}\)Later, Ornstein and Weiss [43], [44], Rudolph, and others extended this theory from the group $\mathbb{Z}$ to actions of arbitrary amenable groups with an invariant measure, but this generalization has no parallel in the theory of filtrations in the form used for the group $\mathbb{Z}$ (the existence of a notion of ‘past’). However, the filtrations that arise as a tool in trajectory approximation (see one of the examples of filtrations in §1) are undoubtedly related also to Ornstein’s condition for general groups.
It seems that the most important idea is that of measuring the degree of ‘being non-Bernoulli’ and ‘being non-standard’ with the help of what was called ‘secondary entropy’ and higher zero–one laws.

9. Acknowledgements

The author must remember and thank the people with whom he discussed the subject of this paper many years ago. First of all was V.A. Rokhlin, one of my teachers, who posed a series of stimulating questions on similar topics and who was very attentive and sympathetic to my ideas on this theme (see [47], [48]); A.N. Kolmogorov, who was interested in filtrations and communicated several my papers on this subject to *Proceedings of the USSR Academy of Sciences*; R. Belinskaya, who asked me one of the first specific questions about dyadic filtrations; S. Yuzvinsky, who took an active interest in the subject and wrote an important paper on applications of the scale of an automorphism [99].

In those years I had useful and fierce disputes with Ya. Sinai, B. Gurevich, V. Alexeev, A. Katok, S. Yuzvinsky, A. Stepin, V. Vinokurov, B. Rubshtein, L. Abramov, M. Gordin, I. Ibragimov, and others, as well as with A. Kirillov, who suggested a meaningful comparison of this subject with some previous works.

Later, after a long pause, interest in the subject was revived abroad; M. Yor’s question about filtrations related to Brownian motion (see also [57]) was partially answered in [7], where my standardness criterion was rediscovered. There was a remarkable flurry of activity in the theory of filtrations in 1997–98 during the semester on ergodic theory at the Israel Institute for Advanced Studies (Jerusalem): H. Furstenberg, B. Weiss, J. Feldman, D. Rudolph, D. Hoffman, D. Heicklen, and J.-P. Thouvenot published a number of papers and participated in discussions of this subject. In the 2000s, M. Émery, who gave a Bourbaki talk [12] on this and related subjects, and his students W. Schachermayer and others also worked in the theory of filtrations. In the 1990s the author had many discussions of the theory of filtrations with J. Feldman (see [15], [16]) and B. Tsirelson (see, in particular, [17]).

Most recently, especially after the subject was combined with the theory of graded graphs, the theory of filtrations and the notion of standardness attracted much interest from S. Laurent, T. de la Rue, and others; they obtained many new results mentioned in this survey (see [38]–[40], [25]). Finally, I have repeatedly discussed the subject with my students A. Gorbulyk, F. Petrov, P. Zatitskiy, and others, and this will hopefully lead to the discovery of new facts and a new understanding of this interesting and important topic. The author thanks all the people mentioned above and also those he may have forgotten to mention. Special thanks are due to A.A. Lodkin for his help with formatting the bibliography, and to A M. Minabutdinov for the figures.
Bibliography

[1] Д.З. Аров, “К истории возникновения понятия $\varepsilon$-энтропии автоморфизма пространства Лебега и понятия $(\varepsilon, T)$-энтропии динамической системы с непрерывным временем”, Теория представлений, динамические системы, комбинаторные методы. XXV, Зап. науч. сем. ПОМИ, 436, ПОМИ, СПб. 2015, с. 76–100; English transl., D.Z. Arov, “To the history of the appearance of the notion of the $\varepsilon$-entropy of an automorphism of a Lebesgue space and $(\varepsilon, T)$-entropy of a dynamical system with continuous time”, J. Math. Sci. (N. Y.) 215:6 (2016), 677–692.

[2] Р.М. Белинская, “Разбиения пространства Лебега на траектории, определяемые эргодическими автоморфизмами”, Функц. анализ и его прил. 2:3 (1968), 4–16; English transl., R.M. Belinskaya, “Partitions of Lebesgue space in trajectories defined by ergodic automorphisms”, Funct. Anal. Appl. 2:3 (1968), 190–199.

[3] A. Connes, J. Feldman, and B. Weiss, “An amenable equivalence relation is generated by a single transformation”, Ergodic Theory Dynam. Systems 1:4 (1981), 431–450.

[4] F. den Hollander and J.E. Steif, “Random walk in random scenery: a survey of some recent results”, Dynamics and stochastics, IMS Lecture Notes Monogr. Ser., vol. 48, Inst. Math. Statist., Beachwood, OH 2006, pp. 53–65.

[5] R.L. Dobrushin, “Perturbation methods of the theory of Gibbsian fields”, Lectures on probability theory and statistics (Saint-Flour 1994), Lecture Notes in Math., vol. 1648, Springer, Berlin 1996, pp. 1–66.

[6] R. Dougherty, S. Jackson, and A.S. Kechris, “The structure of hyperfinite Borel equivalence relations”, Trans. Amer. Math. Soc. 341:1 (1994), 193–225.

[7] L. Dubins, J. Feldman, M. Smorodinsky, and B. Tsirelson, “Decreasing sequences of $\sigma$-fields and a measure change for Brownian motion”, Ann. Probab. 24:2 (1996), 882–904.

[8] H. A. Dye, “On groups of measure preserving transformation. I”, Amer. J. Math. 81 (1959), 119–159.

[9] Е.Б. Дынкин, “Граничная теория марковских процессов (дискретный случай)”, VMH 24:2(146) (1969), 3–42; English transl., E.B. Dynkin, “Boundary theory of Markov processes (the discrete case)”, Russian Math. Surveys 24:2 (1969), 1–42.

[10] Е.Б. Дынкин, “Пространство выходов марковского процесса”, VMH 24:4(148) (1969), 89–152; English transl., E.B. Dynkin, “The space of exits of a Markov process”, Russian Math. Surveys 24:4 (1969), 89–157.

[11] E.B. Dynkin, “Entrance and exit spaces for a Markov process”, Actes du Congr`es International des Math´ematiciens, vol. 2 (Nice 1970), Gauthier-Villars, Paris 1971, pp. 507–512.

[12] M. Émery, “Espace probabilis´es filtr´es: de la th´eorie de Vershik au mouvement brownien, via des id´ees de Tsirelson”, Séminaire Bourbaki, vol. 2000/2001, Ast´erisque, vol. 282, Soc. Math. France, Paris 2002, pp. vii, 63–83, Exp. No 882.

[13] M. Émery and W. Schachermayer, “Brownian filtrations are not stable under equivalent time-changes”, Séminaire de Probabilités, vol. XXXIII, Lecture Notes in Math., vol. 1709, Springer, Berlin 1999, pp. 267–276.

[14] M. Émery and W. Schachermayer, “On Vershik’s standardness criterion and Tsirelson’s notion of cosiness”, Séminaire de Probabilités, vol. XXXV, Lecture Notes in Math., vol. 1755, Springer, Berlin 2001, pp. 265–305.
The theory of filtrations of subalgebras

[15] J. Feldman and C. C. Moore, “Ergodic equivalence relations, cohomology, and von Neumann algebras”, Bull. Amer. Math. Soc. 81:5 (1975), 921–924.

[16] J. Feldman and D. J. Rudolph, “Standardness of sequences of σ-fields given by certain endomorphisms”, Fund. Math. 157:2-3 (1998), 175–189.

[17] J. Feldman and B. Tsirelson, “Decreasing sequences of sigma-fields and a measure change for Brownian motion. II”, Ann. Probab. 24:2 (1996), 905–911.

[18] F. M. Goodman, P. de la Harpe, and V. F. R. Jones, Coxeter graphs and towers of algebras, Math. Sci. Res. Inst. Publ., vol. 14, Springer-Verlag, New York 1989, x+288 pp.

[19] M. Gromov, Metric structures for Riemannian and non-Riemannian spaces, Progr. Math., vol. 152, Birkhäuser Boston, Inc., Boston, MA 1999, xx+585 pp.

[20] Б. М. Гуревич, “К истории динамической энтропии: сравнение двух определений”, Теория представлений, динамические системы, комбинаторные методы. XXV, Зап. науч. сем. ПОМИ, 436, ПОМИ, СПб. 2015, с. 101–111; English transl., B. M. Gurevich, “Toward the history of dynamical entropy: comparing two definitions”, J. Math. Sci. (N. Y.) 215:6 (2016), 693–699.

[21] О. В. Гусева, “Классификация последовательностей измеримых разбиений”, Вестн. ЛГУ. Сер. матем., мекх., астрон. 20:1 (1965), 14–23. [O. V. Guseva, “Classification of sequences of measurable partitions”, Vestn. LGU. Ser. Matem., Mekh., Astron. 20:1, 14–23.]

[22] D. Heicklen, C. Hoffman, and D. J. Rudolph, “Entropy and dyadic equivalence of random walks on a random scenery”, Adv. Math. 156:2 (2000), 157–179.

[23] C. Hoffman and D. J. Rudolph, “A dyadic endomorphism which is Bernoulli but not standard”, Israel J. Math. 130 (2002), 365–379.

[24] É. Janvresse, T. de la Rue, and Y. Velenik, “Self-similar corrections to the ergodic theorem for the Pascal-adic transformation”, Stoch. Dyn. 5:1 (2005), 1–25.

[25] É. Janvresse, S. Laurent, and T. de la Rue, Standardness of monotonic Markov filtrations, 2015, 29 pp., arXiv:1501.02166.

[26] V. A. Kaimanovich and A. M. Vershik, “Random walks on discrete groups: boundary and entropy”, Ann. Probab. 11:3 (1983), 457–490.

[27] S. A. Kalikow, “$T, T^{-1}$ transformation is not loosely Bernoulli”, Ann. of Math. (2) 115:2 (1982), 393–409.

[28] А. Б. Каток, “Монотонная эквивалентность в эргодической теории”, Изв. АН СССР. Сер. матем. 41:1 (1977), 104–157; English transl., A. V. Katok, “Monotone equivalence in ergodic theory”, Math. USSR-Izv. 11:1 (1977), 99–146.

[29] А. Б. Каток, Б. Хасселблат, Введение в современную теорию динамических систем, Факториал, М. 1999, 768 с.; пер. с англ.: A. Katok and B. Hasselblatt, Introduction to the modern theory of dynamical systems, Encyclopedia Math. Appl., vol. 54, Cambridge Univ. Press, Cambridge 1995, xviii+802 pp.

[30] С. В. Керов, “Комбинаторные примеры в теории $AF$-алгебр”, Дифференциальная геометрия, группы Ли и механика. 10, Зап. науч. сем. ЛОМИ, 172, Изд-во “Наука”, Ленингр. отд., Л. 1989, с. 55–67; English transl., S. V. Kerov, “Combinatorial examples in the theory of $AF$-algebras”, J. Soviet Math. 59:5 (1992), 1063–1071.

[31] S. V. Kerov, “Generalized Hall–Littlewood symmetric functions and orthogonal polynomials”, Representation theory and dynamical systems, Adv. Soviet Math., vol. 9, Amer. Math. Soc., Providence, RI 1992, pp. 67–94.

[32] S. V. Kerov, Asymptotic representation theory of the symmetric group and its applications in analysis, with a foreword by A. M. Vershik, Transl. Math. Monogr., vol. 219, Amer. Math. Soc., Providence, RI 2003, xvi+201 pp.
[33] A. N. Kolmogorov, *Foundations of the theory of probability*, 2nd ed., Chelsea Publishing Co., New York 1956, viii+84 pp.

[34] A. N. Kolmogorov, “Новый метрический инвариант транзитивных динамических систем и автоморфизмов пространства Лебега”, *Dokl. AH CCCP* 119:5 (1958), 861–864. [A. N. Kolmogorov, “A new metric invariant of transitive dynamical systems and automorphisms of a Lebesgue space”, *Dokl. Akad. Nauk SSSR* 119:5 (1958), 861–864]; English transl. in *Selected works of A. N. Kolmogorov*, vol. III: Information theory and the theory of algorithms, paper 5, Kluwer, Dordrecht 1993.

[35] A. N. Kolmogorov, “Об энтропии на единицу времени как метрическом инварианте автоморфизмов”, *Dokl. AN SССР* 119:5 (1958), 861–864. [A. N. Kolmogorov, “On entropy per unit time as a metric invariant of automorphisms”, *Dokl. Akad. Nauk SSSR* 124:4 (1959), 754–755.

[36] R. Kotecký and D. Preiss, “The use of projective limits in classical statistical mechanics and Euclidean quantum field theory”, *Czechoslovak J. Phys. B* 30:1 (1980), 23–32.

[37] A. A. Lodkin and A. M. Vershik, “Approximation for actions of amenable groups and transversal automorphisms”, *Operator algebras and their connections with topology and ergodic theory* (Buşteni 1983), Lecture Notes in Math., vol. 1132, Springer, Berlin 1985, pp. 331–346.

[38] S. Laurent, “On standardness and I-cosiness”, *Séminaire de Probabilités*, vol. XLIII, Lecture Notes in Math., vol. 2006, Springer, Berlin 2011, pp. 127–186.

[39] S. Laurent, “On Vershikian and I-cosy random variables and filtrations”, *Theory Probab. Appl.* 55:1 (2011), 54–76.

[40] S. Laurent, “Standardness and nonstandardness of next-jump time filtrations”, *Electron. Commun. Probab.* 18 (2013), 56, 11 pp.

[41] X. Mélá and K. Petersen, “Dynamical properties of the Pascal adic transformation”, *Ergodic Theory Dynam. Systems* 25:1 (2005), 227–256.

[42] D. S. Ornstein, *Ergodic theory, randomness, and dynamical systems*, Yale Math. Monogr., vol. 5, Yale Univ. Press, New Haven, CT–London 1974, vii+141 pp.

[43] D. S. Ornstein and B. Weiss, “Ergodic theory of amenable actions. I. The Rohlin lemma”, *Bull. Amer. Math. Soc. (N. S.)* 2:1 (1980), 161–164.

[44] D. S. Ornstein and B. Weiss, “Entropy and isomorphism theorems for actions of amenable groups”, *J. Anal. Math.* 48 (1987), 1–141.

[45] V. A. Rokhlin, “Об основных понятиях теории меры”, *Matem. cб.* 25(67):1 (1949), 107–150; English transl., V. A. Rokhlin, *On the fundamental ideas of measure theory*, Amer. Math. Soc. Transl., vol. 1952, № 71, Amer. Math. Soc., Providence, RI 1952, 55 pp.

[46] V. A. Rokhlin, “Метрическая классификация измеримых функций”, *UMN* 12:2(74) (1957), 169–174. [V. A. Rokhlin, “Metric classification of measurable functions”, *Uspekhi Mat. Nauk* 12:2(74) (1957), 169–174.]

[47] V. A. Rokhlin, “Лекции по энтропийной теории преобразований с инвариантной мерой”, *UMN* 22:5(137) (1967), 3–56; English transl., V. A. Rokhlin, “Lectures on the entropy theory of measure-preserving transformations”, *Russian Math. Surveys* 22:5 (1967), 1–52.

[48] V. A. Rokhlin, *Избранные работы. Воспоминания о В. А. Рокхлине, материалы к биографии*, 2-e изд., МЦНМО 2010, 576 c. [V. A. Rokhlin, *Selected works. Reminiscences of V. A. Rokhlin, biographical materials*, 2nd ed., MTsNMO, Moscow 2010, 576 pp.]
[49] M. Rosenblatt, *Markov processes. Structure and asymptotic behavior*, Grundlehren Math. Wiss., vol. 184, Springer-Verlag 1971, xiii+268 pp.

[50] B. A. Рубштейн, “Об убывающих последовательностях измеримых разбиений”, *Докл. АН СССР* 205:3 (1972), 526–530; English transl., B. A. Rubštein, “On decreasing sequences of measurable partitions”, *Soviet Math. Dokl.* 13 (1972), 962–965.

[51] B.-Z. Rubshtein, “Lacunary isomorphism of decreasing sequences of measurable partitions”, *Israel J. Math.* 97 (1997), 317–345.

[52] К. Schmidt, “A probabilistic proof of ergodic decomposition”, *Sankhyā Ser. A* 40:1 (1978), 10–18.

[53] Я.Г. Синай, “О понятии энтропии динамической системы”, *Докл. АН СССР* 124:4 (1959), 768–771. [Ya.G. Sinai, “On the concept of entropy of a dynamical system”, *Dokl. Akad. Nauk SSSR* 124:4 (1959), 768–771.]

[54] Я.Г. Синай, “О понятии энтропии динамической системы”, *Докл. АН СССР* 124:4 (1959), 768–771. [Ya.G. Sinai, “On the concept of entropy of a dynamical system”, *Dokl. Akad. Nauk SSSR* 124:4 (1959), 768–771.]

[55] R. P. Stanley, *Ordered structures and partitions*, Mem. Amer. Math. Soc., vol. 119, Amer. Math. Soc., Providence, RI 1972, iii+104 pp.

[56] А.М. Степин, “Об энтропийном инварианте убывающих последовательностей измеримых разбиений”, *Функц. анали́з и его прил.* 5:3 (1971), 80–84; English transl., A. M. Stepin, “On entropy invariants of decreasing sequences of measurable partitions”, *Funct. Anal. Appl.* 5:3 (1971), 237–240.

[57] N. Tsilevich, A. Vershik, and M. Yor, “An infinite-dimensional analogue of the Lebesgue measure and distinguished properties of the gamma process”, *J. Funct. Anal.* 185:1 (2001), 274–296.

[58] А.М. Вершик, “Теорема о лакунарном изоморфизме монотонных последовательностей разбиений”, *Функц. анали́з и его прил.* 2:3 (1968), 17–21; English transl., A. M. Vershik, “Theorem on lacunary isomorphisms of monotonic sequences of partitions”, *Funct. Anal. Appl.* 2:3 (1968), 200–203.

[59] А.М. Вершик, “Убывающие последовательности измеримых разбиений и их применения”, *Докл. АН СССР* 199:4 (1970), 748–751; English transl., A. M. Vershik, “Decreasing sequences of measurable partitions and their applications”, *Soviet Math. Dokl.* 11 (1970), 1007–1011.

[60] А.М. Вершик, “Неизмеримые разбиения, траекторная теория, алгебры операторов”, *Докл. АН СССР* 199 (1970), 1004–1007; English transl., A. M. Vershik, “Nonmeasurable decompositions, orbit theory, algebras of operators”, *Soviet Math. Dokl.* 12 (1971), 1218–1222.

[61] П. Б. Затицкий, Ф. В. Петров, “Об исправлении метрик”, *Теория представлений, динамические системы, комбинаторные методы*. ХХ, Зап. науч. сем. ПОМИ, 390, ПОМИ, СПб. 2011, с. 201–209; English transl., P. B. Zatitskiy and F. V. Petrov, “Correction of metrics”, *J. Math. Sci. (N. Y.)* 181:6 (2012), 867–870.

[62] А.М. Вершик, “Континуум попарно неизоморфных диадических последовательностей”, *Функц. анали́з и его прил.* 5:3 (1971), 16–18; English transl., A. M. Vershik, “Continuum of pairwise nonisomorphic dyadic sequences”, *Funct. Anal. Appl.* 5:3 (1971), 182–184.

[63] А.М. Вершик, “Устройство ручных разбиений”, *УМН* 27:3(165) (1972), 195–196. [A. M. Vershik, “Construction of tame partitions”, *Uspekhi Mat. Nauk* 27:3(165) (1972), 195–196.]

[64] А.М. Вершик, “Четыре определения шкалы автоморфизма”, *Функц. анали́з и его прил.* 7:3 (1973), 1–17; English transl., A. M. Vershik, “Four definitions of the scale of an automorphism”, *Funct. Anal. Appl.* 7:3 (1973), 169–181.
[65] A. M. Vershik, *Approximation in measure theory*, Diss. ... dокт. физ.-матем. наук, ЛГУ, Л. 1973, 302 с. [A. M. Vershik, *Approximation in measure theory*, Doctoral dissertation in the physical and mathematical sciences, Leningrad State Univ., Leningrad 1973, 302 pp.]

[66] A. M. Vershik, “О работах Д. Орнштейна, условиях слабой зависимости и о классах стационарных мер”, *Теория вероятн. и ее примен.* 21:3 (1976), 673–675; English transl., A. M. Vershik, “On D. Ornstein’s papers, weak dependence conditions and classes of stationary measures”, *Theory Probab. Appl.* 21 (1977), 655–657.

[67] A. M. Vershik, “Равномерная алгебраическая аппроксимация операторов сдвига и умножения”, *Докл. АН СССР* 259:3 (1981), 526–529; English transl., A. M. Vershik, “Uniform algebraic approximation of shift and multiplication operators”, *Soviet Math. Dokl.* 24:1 (1981), 97–100.

[68] A. M. Vershik, “Теорема о марковской периодической аппроксимации в эргодической теории”, *Краевые задачи математической физики и смежные вопросы теории функций*. 14, Зап. науч. сем. ЛОМИ, 115. Изд-во “Наука”, Ленингр. отд., Л. 1982, с. 72–82; English transl., A. M. Vershik, “A theorem on the Markov periodic approximation in ergodic theory”, *J. Soviet Math.* 28:5 (1985), 667–674.

[69] A. M. Vershik, “Теория убывающих последовательностей измеримых разбиений”, *Алгебра и анализ* 6:4 (1994), 1–68; English transl., A. M. Vershik, “Theory of decreasing sequences of measurable partitions”, *St. Petersburg Math. J.* 6:4 (1995), 705–761.

[70] A. M. Vershik, “Классификация измеримых функций нескольких аргументов и инвариантно распределенные случайные матрицы”, *Функц. анали. и его прил.* 36:2 (2002), 12–27; English transl., A. M. Vershik, “Classification of measurable functions of several variables and invariantly distributed random matrices”, *Funct. Anal. Appl.* 36:2 (2002), 93–105.

[71] A. M. Vershik, “Случайные метрические пространства и универсальное пространство Урысона”, *Фундаментальная математика сегодня*, МЦНМО, М. 2003, с. 54–88. [A. M. Vershik, “Random metric spaces and the universal Uryson space”, *Fundamental mathematics of today*, MTsNMO, Moscow 2003, pp. 54–88.]

[72] A. M. Vershik, “Случайные метрические пространства и универсальность”, *УМН* 59:2(356) (2004), 65–104; English transl., A. M. Vershik, “Random metric spaces and universality”, *Russian Math. Surveys* 59:2 (2004), 259–295.

[73] A. M. Vershik, “Информация, энтропия, динамика”, *Математика XX века. Взгляд из Петербурга*, МЦНМО, М. 2010, с. 47–76. [A. M. Vershik, “Information, entropy, dynamics”, *Mathematics of the 20th century. The view from St. Petersburg*, MTsNMO, Moscow 2010, pp. 47–76.]

[74] A. M. Vershik, “Dynamics of metrics in measure spaces and their asymptotics invariants”, *Markov Process. Related Fields* 16:1 (2010), 169–184.

[75] A. M. Vershik, “Автоморфизм Паскаля имеет непрерывный спектр”, *Функц. анали. и его прил.* 45:3 (2011), 16–33; English transl., A. M. Vershik, “The Pascal automorphism has a continuous spectrum”, *Funct. Anal. Appl.* 45:3 (2011), 173–186.

[76] A. M. Vershik, “О классификации измеримых функций нескольких переменных”, *Теория представлений, динамические системы, комбинаторные методы. XXI*, Зап. науч. сем. ПОМИ, 403, ПОМИ, СПб. 2012, c. 35–57; English transl., A. M. Vershik, “On classification of measurable functions of several variables”, *J. Math. Sci. (N. Y.)* 190:3 (2013), 427–437.
А.М. Вершик, “Задача о центральных мерах на пространствах путей градуированных графов”, Функциональный анализ и его приложения 48:4 (2014), 26–46; English transl., A. M. Vershik, “The problem of describing central measures on the path spaces of graded graphs”, Funct. Anal. Appl. 48:4 (2014), 256–271.

А.М. Вершик, “Оснащенные градуированные графы, проективные пределы симплексов и их границы”, Теория представлений, динамические системы, комбинаторные методы. XXIV, Зап. науч. сем. ПОМИ, 432, ПОМИ, СПб. 2015, с. 83–104; English transl., A. M. Vershik, “Equipped graded graphs, projective limits of simplices, and their boundaries”, J. Math. Sci. (N. Y.) 209:6 (2015), 860–873.

А.М. Вершик, “Стандартность как инвариантная формулировка независимости”, Функциональный анализ и его приложения 49:4 (2015), 18–32; English transl., A. M. Vershik, “Standardness as an invariant formulation of independence”, Funct. Anal. Appl. 49:4 (2015), 253–263.

А. М. Вершик, “Several remarks on Pascal automorphism and infinite ergodic theory”, Armen. J. Math. 7:2 (2015), 85–96.

А.М. Вершик, “Smoothness and standardness in the theory of AF-algebras and in the problem on invariant measures”, Probability and statistical physics in St. Petersburg. Proc. Sympos. Pure Math., vol. 91, Amer. Math. Soc., Providence, RI 2016, pp. 423–436.

А. М. Вершик, “Asymptotic theory of path spaces of graded graphs and its applications”, Jpn. J. Math. 11:2 (2016), 151–218.

А. М. Вершик, А. Л. Фёдоров, “Траекторная теория”, Итоги науки и техн. Сер. Соврем. пробл. матем. Нов. достиж., 26, ВИНИТИ, М. 1985, с. 171–211; English transl., A. M. Vershik and A. L. Fedorov, “Trajectory theory”, J. Soviet Math. 38:2 (1987), 1799–1822.

А. М. Вершик, А. Л. Горбульский, “Масштабированная энтропия фильтраций σ-алгебр”, Теория вероятности и ее применение 52:3 (2007), 446–467; English transl., A. M. Vershik and A. D. Gorbul’skii, “Scaled entropy of filtrations of σ-fields”, Theory Probab. Appl. 52:3 (2008), 493–508.

А. М. Вершик и У. Хабоэкк, “On the classification problem of measurable functions in several variables and on matrix distributions”, Вероятность и статистика. 22, Зап. науч. сем. ПОМИ, 441, ПОМИ, СПб. 2015, с. 119–143; J. Math. Sci. (N. Y.) 219:5 (2016), 683–699.

А. М. Вершик, С. В. Керов, “Локально полупростые алгебры. Комбинаторная теория и K₀-функтор”, Итоги науки и техн. Сер. Соврем. пробл. матем. Нов. достиже., 26, ВИНИТИ, М. 1985, с. 3–56; English transl., “Locally semisimple algebras. Combinatorial theory and the K₀-functor”, J. Soviet Math. 38:2 (1987), 1701–1733.

А. М. Вершик, С. В. Керов, “Асимптотика максимальной и типичной размерностей неприводимых представлений симметрической группы”, Функциональный анализ и его приложения 19:1 (1985), 25–36; English transl., A. M. Vershik and S. V. Kerov, “Asymptotic of the largest and the typical dimensions of irreducible representations of a symmetric group”, Funct. Anal. Appl. 19:1 (1985), 21–31.

А. М. Вершик, И. П. Корнфельд, “Общая эргодическая теория групп преобразований с инвариантной мерой. Гл. 4. Периодические аппроксимации и их приложения. Эргодические теоремы, спектральная и энтропийная теория для действий общих групп”, Динамические системы – 2, Итоги науки и техн. Сер. Соврем. пробл. матем. Фундам. направлении, 2, ВИНИТИ, М. 1985, с. 70–89; English transl., I. P. Kornfel’d and A. M. Vershik, “General ergodic theory of transformation groups with invariant measure. Chapter 4. Periodic
approximations and their applications. Ergodic theorems, spectral and entropy theory for general group actions”, Dynamical systems II. Ergodic theory with applications to dynamical systems and statistical mechanics, Encyclopaedia Math. Sci., vol. 2, Springer, Berlin 1989, pp. 59–77.

[89] A. M. Vershik and A. N. Livshits, “Adic models of ergodic transformations, spectral theory, substitutions, and related topics”, Representation theory and dynamical systems, Adv. Soviet Math., vol. 9, Amer. Math. Soc., Providence, RI 1992, pp. 185–204.

[90] A. M. Vershik, A. V. Malyshev, “Phase transition in the exit boundary problem for random walks on groups”, Funct. Anal. Appl. 49:2 (2015), 86–96.

[91] A. M. Vershik, P. B. Zatitskii, “Universal adic approximation, invariant measures and scaled entropy”, Izv. RAN. Ser. matem., 2017 (in press); Preprint POMI 16/2016, POMI, SPb. 2016.

[92] A. M. Vershik, P. B. Zatitskiy, and F. V. Petrov, “Geometry and dynamics of admissible metrics in measure spaces”, Cent. Eur. J. Math. 11:3 (2013), 379–400.

[93] A. M. Vershik, P. B. Zatitskiy, and F. V. Petrov, “Virtual continuity of measurable functions and its applications”, Russian Math. Surveys 69:6 (2014), 1031–1063.

[94] V. G. Vinokurov, N. N. Ganikhodzhaev, “Conditional functions in the trajectory theory of dynamical systems”, Izv. AN CCCP. Ser. matem. 42:5 (1978), 928–964; English transl., V. G. Vinokurov and N. N. Ganikhodzhaev, “Conditional functions in the trajectory theory of dynamical systems”, Math. USSR-Izv. 13:2 (1979), 221–252.

[95] V. G. Vinokurov, B. A. Rubinstein, “Extensions of decreasing sequences of measurable partitions of a Lebesgue space”, Random processes and related questions, vol. vyp. 2, Fan, Tashkent 1972, pp. 49–61.

[96] J. von Neumann, “Zur Operatorenmethode in der klassischen Mechanik”, Ann. of Math. (2) 33:3 (1932), 587–642.

[97] N. Wiener, Nonlinear problems in random theory, Technology Press Research Monographs, Technology Press of M.I.T., John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London 1958, ix+131 pp.

[98] G. Winkler, Choquet order and simplices with applications in probabilistic models, Lecture Notes in Math., vol. 1145, Springer-Verlag, Berlin 1985, vi+143 pp.

[99] S. A. Yuzvinsky, “Distinction of K automorphisms by the scale”, Funct. Anal. Appl. 7:4 (1973), 312–316.

[100] P. B. Zatitskiy, “On a scaling entropy sequence of a dynamical system”, Funct. Anal. Appl. 48:4 (2014), 291–294.

[101] P. B. Zatitskiy, “Masstabilizierende entropische Verzweigung: Invariantzung und praktische Anwendungen”, Teorija predstavlenij, dinamicheskie sistemy, kombinatornye metody. XXIV, Zap. nauč. sem. POMI, 432, POMI, SPb.
2015, c. 128–161; English transl., P. B. Zatitskiy, “Scaling entropy sequence: invariance and examples”, J. Math. Sci. (N. Y.) 209:6 (2015), 890–909.

Anatoly M. Vershik
St. Petersburg Department
of the Steklov Mathematical Institute;
St. Petersburg State University;
Institute for Information Transmission Problems
E-mail: avershik@gmail.com

Received 24/JAN/17
Translated by N. TSILEVICH