JOURNAL OF MODERN DYNAMICS
VOLUME 1, NO. 4, 2007, 545–596
WEB SITE: http://www.math.psu.edu/jmd

FIFTY YEARS OF ENTROPY IN DYNAMICS: 1958–2007

ANATOLE KATOK

PREFACE

These notes combine an analysis of what the author considers (admittedly subjectively) as the most important trends and developments related to the notion of entropy, with information of more "historical" nature including allusions to certain episodes and discussion of attitudes and contributions of various participants. I directly participated in many of those developments for the last forty three or forty four years of the fifty-year period under discussion and on numerous occasions was fairly close to the center of action. Thus, there is also an element of personal recollections with all attendant peculiarities of this genre.

These notes are meant as "easy reading" for a mathematically sophisticated reader familiar with most of the concepts which appear in the text. I recommend the book [59] as a source of background reading and the survey [44] (both written jointly with Boris Hasselblatt) as a reference source for virtually all necessary definitions and (hopefully) illuminating discussions.

The origin of these notes lies in my talk at the dynamical systems branch of the huge conference held in Moscow in June of 2003 on the occasion of Kolmogorov's one hundredth anniversary. The title of the talk was *"The first half century of entropy: the most glorious number in dynamics"*. At that time not quite half a century had elapsed after the Kolmogorov discovery, although one could arguably include some of the "pre-history" into the counting. Now this ambiguity has disappeared, and I dedicate this article to the fiftieth anniversary of the discovery of entropy in ergodic theory by Andrei Nikolaevich Kolmogorov.

There is a number of published recollections and historical notes related in various ways to our subject. I used some information from these sources with references, although this plays a secondary role to personal recollections and analysis based on the knowledge of the subject and its historical development. Naturally I took all necessary care not to contradict any source of this kind. I do recommend the account by M. Smorodinsky [140] which addresses some of the

Received April 3, 2007.
2000 *Mathematics Subject Classification:* Primary: 37-03, Secondary: 01A60.
Key words and phrases: Entropy.
Partially supported by NSF Grant 0502205.
I acknowledge great help from Boris Hasselblatt at the earlier stage of writing this paper; his contribution to the early history section has been particularly valuable.

The pictures of Ornstein, Sinai and Parry and Adler are reproduced with permission from the Mathematisches Forschungsinstitut Oberwolfach.

same themes as the first half of this paper and contains more systematic definitions and formulations as well as some complementary insights.

I also used information and insights provided by colleagues who participated in various developments described in this paper, namely, Benjamin Weiss, Don Ornstein, Roy Adler, and Jean-Paul Thouvenot. Some of the facts and interpretations I learned from them came as a surprise to me, and this certainly helped to enrich and correct the overall picture. Benjamin Weiss read several drafts of this paper and made a number of valuable comments. I am also thankful to Gregory Margulis for comments of critical nature which led me to formulate some of my opinions and comments with greater care and clarity.

Note on references. This paper is not meant as a bibliographical survey, and despite the considerable size of the bibliography, a comprehensive or systematic listing of references was not a goal. Some important papers may have been missed if they are not directly related to the main lines of our narrative. Sometimes key references are listed after the titles of subsections (and on one occasion, a section).

INTRODUCTION: KOLMOGOROV'S CONTRIBUTION TO DYNAMICS: 1953–1959 [75, 76, 77, 78, 79]

Andrei Nikolaevich Kolmogorov arguably has made the most profound impact on the shaping and development of mathematical theory of dynamical systems since Henri Poincaré.

Two cardinal branches of modern dynamics are concerned with stability of motion over long (in fact, infinite) intervals of time, and with complicated ("stochastic" or "chaotic") behavior in deterministic systems, respectively. Kolmogorov's contributions to both areas are seminal and, in fact, have determined the main trends of development for at least half a century.

In the area of stability Kolmogorov discovered (about 1953) the persistence of many (but not all!) quasiperiodic motions for a broad class of completely integrable Hamiltonian systems [76, 77]. The major branch of modern dynamics and analysis that grew out of Kolmogorov's discovery is usually referred to as the KAM (Kolmogorov–Arnol'd–Moser) theory. For brief historical accounts of the development of this area see [96] and [45], for a well-motivated introduction and overview, see [87]. The plenary talk at the International Congress of Mathematicians in Amsterdam [77] not only contained a sketch of Kolmogorov's great theorem but also an outline of a very broad program for studying long-term behavior of classical dynamical systems, both stable and ergodic, far ahead of its time. The only other paper by Kolmogorov in this area was an earlier note [75] which contained both preliminary steps for KAM and the discovery of an exotic ergodic behavior associated with abnormally fast approximation of irrational numbers by rationals.

The second great contribution of Kolmogorov to modern dynamics was the discovery of the concept of entropy for a measure-preserving transformation, together with the attendant property of completely positive entropy (K-property)

[78, 79]. It is the development of this concept and its great impact not only in dynamics but also in certain areas of geometry and number theory that is traced and discussed in this paper.

To summarize, the total length of Kolmogorov's pertinent published work is under 35 pages. There are repetitions (a part of the Amsterdam talk [77] follows [75] and [76]), and errors (in the main definition in [78], so [79] is essentially a correction of that error; also in [75] mixing is mentioned where it is in fact impossible, the correct property being weak mixing), and some of the predictions of [77] were not borne out, such as the likelihood of a great difference between ergodic properties of real-analytic and C^∞ dynamical systems.

Andrei Nikolaevich Kolmogorov

One should add for the sake of balance that some of Kolmogorov's ideas in other directions did not appear in his published work but were instead developed by his students. A well-known example is the construction of a measure-preserving transformation with simple continuous spectrum by I.V. Girsanov [34] based on the theory of Gaussian processes. But even allowing for that, the magnitude and persistence of the impact of Kolmogorov's work on the subsequent development of dynamics is amazing given the modest size of the published output and the deficiencies mentioned above.[1] A very great role in both principal cases (KAM and entropy) was played by the work of students and other mathematicians related to Kolmogorov which broadened, developed, and amplified Kolmogorov's original insights. In the case of entropy, the key role was

[1] In 1985 Kolmogorov published another corrected version [80] of his original note on entropy [78]. While this event has a purely historical interest it indicates that Kolmogorov valued his contribution very highly and wanted to be on record for having corrected the annoying error. He says: "This paper, which has been prepared specially for the jubilee edition of the Proceedings of the Steklov Institute, is a rewritten version of our note with the same title in which there was an erroneous Theorem 2; in the present version the statement and proof of that theorem have been completely changed."

played by Sinai, Rokhlin, Pinsker and Abramov, whose work during the period from 1958 till 1962 brought entropy to the center stage in ergodic theory and developed the first layer of the core theory, which remains the foundation of the subject to this day.

In the rest of the paper we will discuss the principal developments related to the notion of entropy in various areas of dynamics. We try to combine a brief historical account with explanations of certain concepts and ideas as we see them today. Due to size limitations we reduce technical details to a bare minimum and stress motivations and implications rather than (even sketches of) proofs. It goes without saying that we do not aspire to provide a comprehensive treatment of either historical or conceptual aspects of the subject. The balance between the two changes from one section to another, with a general tendency to shift toward more explanations and fewer historical comments as we progress to more recent developments.

1. EARLY HISTORY OF ENTROPY

From its origins as a fundamental—albeit, rather vague—notion in statistical mechanics, the concept of entropy became the centerpiece of mathematical theory of information created by Claude Shannon; Kolmogorov, in a masterful stroke, realized its invariant nature. In few short years, ergodic theory was revolutionized.

a. **Prehistory: 18...–1956.** [2]

Thermodynamics. The word *entropy*, amalgamated from the Greek words *en*ergy and *trop*os (meaning "turning point"), was introduced in an 1864 work of Rudolph Clausius, who defined the change in entropy of a body as heat transfer divided by temperature, and he postulated that overall entropy does not decrease (the second law of thermodynamics). Clausius was motivated by an earlier work of Sadi Carnot on an ideal engine in which the entropy (how Clausius understood it) would be constant, see [97].

While entropy in this form proved quite useful, it did not acquire intuitive meaning until James Maxwell and Ludwig Boltzmann worked out an atomistic theory of heat based on probability, *i.e.,* statistical mechanics, the "mechanical theory of heat". A central object in this endeavor was the *distribution function* on the phase space of *one* of the gas particles, which measured, for a given state of the gas as a whole (what physicists call a *microstate*) the distribution of states of individual particles; this distribution is also called a *macrostate*. Integrating this density gives 1 by definition, but Boltzmann derived a partial differential equation for it that yielded a few standard conserved quantities: Total mass is the particle mass integrated against this density; total momentum is the particle momentum integrated against it; total energy is kinetic energy of a particle integrated against this density; and entropy turns out to be the logarithm of this density integrated against the density itself, i.e., an integral of the form $-k\int f \log f$.

[2] This section was mostly written by Boris Hasselblatt.

It turns out to be natural to quantize this kind of system by partitioning the phase space (for a *single* particle) into finitely many cells (e.g., by dividing the single-particle energy into finitely many possible levels); then we can use a probability vector with entries p_i to list the proportion of particles that are in each of these partition elements. Boltzmann's description of entropy was that it is the degree of uncertainty about the state of the system that remains if we are given only the p_i, (i.e., the distribution or macrostate), and that this is properly measured by the logarithm of the number of states (microstates) that realize this distribution. For n particles, this number is $(\prod_i (np_i)!)/n!$. Stirling's formula gives the continuum limit of entropy as

$$\lim_{n\to\infty} \frac{\prod_i (np_i)!}{n \cdot n!} = \sum_i p_i \log p_i.$$

In particular, the most probable states can be found by maximizing $\sum_i p_i \log p_i$ (discrete Maxwell–Boltzmann law) [21, Lemma 2.3]. See [33] for more details.

Further development and clarification of the notion of entropy in statistical mechanics lie outside of the scope of this article.

Information theory [128, 71]. Shannon considered finite alphabets whose symbols have known probabilities p_i and, looking for a function to measure the uncertainty in choosing a symbol from among these, determined that, up to scale, $\sum_i p_i \log p_i$ is the only continuous function of the p_i that increases in n when $(p_1, \dots, p_n) = (1/n, \dots, 1/n)$ and behaves naturally with respect to making successive choices. It is the weighted average of the logarithmic size of elements and in particular it is additive for the join of independent distributions. One can argue that entropy expresses the amount of information carried by the distribution.

Accordingly, the *entropy* of a finite or countable measurable partition ξ of a probability space is given by

$$H(\xi) := H_\mu(\xi) := -\sum_{C\in\xi} \mu(C) \log \mu(C) \geq 0,$$

where $0 \log 0 := 0$. For countable ξ, the entropy may be either finite or infinite.

In the presence of a stationary random process with finitely many states (or in Shannon's language, an *information source*), probabilities of symbols are affected by preceding symbols, and the effect on the probability distribution of longer words (or the itineraries of points with respect to a partition) is captured by joint information. Thus, elementary events may be considered as measurable subsets in the space of realizations of a stationary random process, and they form a partition which one may denote as ξ.

As time progresses, the total amount of information per unit of time may only decrease. Shannon took the asymptotic amount of information per unit of time to be the entropy of an information source. We will explain this more formally shortly. A human language, for example, has lower entropy than an independent identically distributed (i.i.d.) process with the same alphabet (which should also include a symbol for the empty space between words) because the probabilities of different letters differ (which is accounted for already by the entropy of the partition) and further, because the choice of each new subsequent letter is significantly restricted by the choice of previous letters.

The seminal work of Shannon became the basis of mathematical information theory. Its relationship to the theory of stationary random processes became apparent to probabilists. In 1956, A. Ja. Khinchin [71] gave a very elegant, rigorous treatment of information theory with the entropy of a stationary random process as a centerpiece. Fully aware of the equivalence between a stationary random process with finite state space and a measure-preserving transformation, T, with a finite partition, ξ, Khinchin developed the basic calculus of entropy, which in retrospect looks like a contemporary introduction to the subject of entropy in ergodic theory.

We define the *joint partition* by

$$\xi_{-n}^{T} := \bigvee_{i=1}^{n} T^{1-i}(\xi),$$

where $\xi \vee \eta := \{C \cap D \mid C \in \xi, D \in \eta\}$. Now

$$h(T,\xi) := h_{\mu}(T,\xi) := \lim_{n\to\infty} \frac{H(\xi_{-n}^{T})}{n}$$

is called the *metric entropy* of the transformation T *relative* to the partition ξ. (It is easy to see that the limit exists.)

Via conditional entropies

$$H(\xi \mid \eta) := -\sum_{D\in\eta} \mu(D) \sum_{C\in\xi} \mu(C \mid D) \log \mu(C \mid D),$$

where $\mu(A \mid B) := \mu(A\cap B)/\mu(B)$, the entropy of a finite state random process (or, equivalently, the entropy of a measure-preserving transformation with respect to a given finite partition) can equivalently be defined as the average amount of information obtained on one step given complete knowledge of the past (meaning the sequence of partition elements to which preimages of a given point belong), *i.e.*,

$$h(T,\xi) = H(\xi \mid \xi_{-\infty}^{T}).$$

At this stage entropy provides a parameter that describes the complexity of the stationary random process with finitely many *states*, *symbols*, or *letters* in a finite *alphabet*. Recall that given a measure-preserving transformation, T, of a probability space, (X,μ), one produces such a process by fixing a finite partition ξ of X, identifying its elements with the letters of an alphabet of card(ξ) symbols and *coding* the transformation via this partition. The entropy of the random process thus obtained is usually denoted by $h(T,\xi)$.

b. **Kolmogorov's discovery** [78, 79].

Kolmogorov entropy. Naturally, the same transformation can be coded by many different partitions, and entropies of the corresponding random processes may differ. To give a trivial (and extreme) example, consider any process with a partition where one element has measure one (full measure) and the others have measure zero. Obviously, the entropy of such a process is equal to zero and all the information contained in the original process is lost; in less extreme cases, there may be a partial loss of information, *e.g.,* for an independent uniformly distributed random process with four states and with the partition accounting for the parity of the state at moment zero.

Kolmogorov realized that this can be used to define a quantity that describes the intrinsic complexity of a measure-preserving transformation. To see how, we say that a measurable partition ξ with finite entropy is a *generator* for a measure-preserving transformation T if the set of partitions subordinate to some $\bigvee_{i=-n}^{n} T^{-i}(\xi)$ is dense in the set of finite-entropy partitions endowed with the *Rokhlin metric*

$$d_R(\xi,\eta) := H(\xi \mid \eta) + H(\eta \mid \xi).$$

The key observation here is the inequality

$$|h(T,\xi) - h(T,\eta)| \le d_R(\xi,\eta).$$

Kolmogorov noted that all generators for a measure-preserving transformation T have the same entropy and defined the entropy of T to be this common value if T has a generator and ∞ otherwise. The latter choice may be defensible, but has as a consequence that the entropy of the identity is infinite. Sinai found a natural way to make this notion better-behaved by observing that generators maximize entropy relative to a partition among all partitions with finite entropy and by defining the entropy of a measure-preserving transformation T (or *metric entropy*) as $\sup\{h(T,\xi) \mid H(\xi) < \infty\}$.

K-systems. The property called *K-property*[3], also introduced by Kolmogorov in 1958, is an isomorphism invariant version of earlier regularity notions for random processes: present becomes asymptotically independent of all sufficiently long past. This notion is quite natural in the context of stationary random processes where various versions of regularity had been studied for several decades and were one of the focal points of Kolmogorov's own work in probability.[4] A fundamental observation is that K-property is equivalent to *completely positive entropy*:

$$h(T,\xi) > 0 \text{ for any partition } \xi \text{ with } H(\xi) > 0.$$

and thus is inherited by any stationary random process associated with a K-system. [121, 124]

[3] Kolmogorov's original motivation was as an abbreviation for "quasiregular," which begins with "K" in Russian, but it was quickly interpreted as the first letter of the name "Kolmogorov" and is still sometimes called the "Kolmogorov property."

[4] A version of this notion for stationary processes with continuous distribution appears in N. Wiener's 1958 monograph [143]. Among other things, it contains an incorrect proof that such a process is isomorphic (using ergodic theory language) to a sequence of independent identically distributed random variables, *i.e.,* to a Bernoulli process; even though this statement is repeated in Grenander's review in Math. Reviews, it does not seem to have been noticed by the ergodic theory community at the time, and Wiener's work had no influence on development of ergodic theory. However, Wiener's proof of existence of a Bernoulli factor in his setting is correct.

c. **Development of the basic theory: 1958–1962** [123, 124]**.** In the short period following the discovery of Kolmogorov, a group of brilliant mathematicians in Moscow realized the potential of his seminal insight and quickly developed the basic machinery which forms the core of the theory of entropy and K-systems. This group included V.A. Rokhlin, then in his early forties, for whom this was the second very productive period of work in ergodic theory after 1947–1950; the information theorist, M.S. Pinsker, who was in his early thirties; Rokhlin's student L.M. Abramov; and Kolmogorov's student, Ja.G. Sinai, in his mid-twenties, whose name will appear in this survey many times. Sinai's definition of entropy [130], mentioned above, has become standard.[5] Rokhlin's lectures [124], published in 1967 but written several years earlier, present this layer of the theory in a definitive form and serve as the model for most later textbooks and monographs.

We do not know to what extent Kolmogorov anticipated these developments.

A measure-preserving transformation, S, is called a *factor* of a measure-preserving transformation, T, if S is isomorphic to the restriction of T to an invariant σ-algebra. One can interchangeably speak about invariant partitions, invariant sub σ-algebras and factors of a measure-preserving transformation. Thus, characterization of K-systems as systems with completely positive entropy implies that any nontrivial factor of a K-system is a K-system.

The π-*partition* is the crudest partition (minimal σ-algebra) which refines every partition with zero entropy. It can also be characterized as the algebra of events which are completely determined by infinite past, however remote. This is the essence of zero entropy: a system with zero entropy is fully deterministic from the probabilistic or information-theory viewpoint. Complete knowledge of the arbitrary remote past, *i.e.,* the partition

$$\xi_{-\infty}^{T} = \bigvee_{n=-\infty}^{-N} T^n(\xi), \tag{1}$$

for any N, however large, fully determines the present and the future, *i.e.,* the partition

$$\xi^T := \bigvee_{n=-\infty}^{\infty} T^n(\xi).$$

An arbitrary system with positive entropy is then represented as the *skew product* over its largest zero entropy factor, determined by the π-partition. Thus, it has a canonical, deterministic component; to what extent one can find a complementary random, K, or Bernoulli, component was an early question which stimulated a lot of development.

For K-systems, the situation is opposite that of zero entropy systems: knowledge or arbitrary remote past gives no information about the present or future, *i.e.,* the σ-algebra of the partition (1) as $N \to \infty$ converges to the trivial σ-algebra.

The short note by Pinsker [121], where the π-partition was introduced, is one of the most important pieces of work from this early period. Pinsker also proved

[5] Metric (Kolmogorov) entropy is sometimes called Kolmogorov–Sinai entropy, especially in physics literature.

that in any system, any K-factor (if it exists) is independent from the π-partition and hence from any zero entropy factor. Using the notion introduced by H. Furstenberg several years later [32], one can reformulate this as saying that any K-system is *disjoint* from any system with zero entropy. Pinsker did not know at the time whether any ergodic positive entropy system has a K-factor (this follows from Sinai's weak isomorphism theorem proved several years later Section 2c), but based on his independence result, he formulated a bold conjecture (later known as *Pinsker Conjecture*) that every positive entropy system is a direct product of a zero entropy and a K-system.

Realization of the importance of the Kolmogorov discovery and the great impact it made on the field is attested by appearance within a couple of years of influential surveys by Rokhlin [123], and P.R. Halmos [43] centered around entropy and its impact. The list of open problems inspired by Kolmogorov's work in dynamics (not only on entropy) by S.V. Fomin [30] appeared as an appendix to the Russian translation of [42], the only textbook on ergodic theory in circulation at the time. One of Fomin's problems was whether the spectrum and entropy determine the isomorphism class of an ergodic measure-preserving transformation; it was quickly answered in the negative by Roy Adler [2].

d. **Entropy and local behavior: early insights.** It was noticed very early in the development of the entropy theory that entropy is related to the local or, in the case of smooth systems, infinitesimal, exponential rates of divergence of orbits and is somehow capturing the sum total of those rates in all directions. Thus entropy, whose genesis was in probability theory and information theory, found its way into classical dynamics. According to Sinai [138, 139], there was for a short time a misconception among the young people mentioned above that entropy of classical systems must always be equal to zero, which was dispelled by Kolmogorov. He pointed out that the linear hyperbolic map of a torus, an obvious example of a discrete time classical system (a volume-preserving diffeomorphism of a compact manifold) has positive entropy.

The formula for entropy with respect to Lebesgue measure for the automorphism F_A of the torus $\mathbb{T}^k = \mathbb{R}^k/\mathbb{Z}^k$ determined by an $k \times k$ matrix A with integer entries and determinant of absolute value one,

$$F_A x = Ax \qquad (\text{mod } 1) \quad \text{for} \quad x \in \mathbb{R}^k/\mathbb{Z}^k,$$

$$h(F_A) = \sum_{\lambda \in \operatorname{Sp} A} (\log|\lambda|)_+, \tag{2}$$

(where each eigenvalue appears according to its multiplicity and $x_+ = \frac{x+|x|}{2}$) became the prototype for entropy calculations for increasingly more general classes of smooth dynamical systems. The ultimate outcome is the formula by Ya.B. Pesin in the classical volume-preserving case and the further generalization by F. Ledrappier and L.-S. Young for arbitrary Borel probability invariant measures for smooth systems. This story belongs to a later part of this survey, Section 4, and now we mention only the earliest development in that direction.

The 1965 result by then 21-year old A.G. Kushnirenko [81] established a fundamental fact: *entropy for a classical dynamical system is finite*. Kushnirenko did it by estimating entropy of a partition by using an isoperimetric inequality and showing that the growth of the total boundary area of iterates of a partition into elements with piecewise smooth boundaries can be estimated through the Lipschtiz constant of the map in question and dimension of the manifold.

e. **Three big classes.** Thus the world of dynamics has been divided by entropy into three distinct parts:

- *Zero-entropy* or fully deterministic systems. Classical systems of that kind are characterized by subexponential growth of orbit complexity with time.
- *Finite positive entropy* systems include all finite state stationary random processes such as independent ones (Bernoulli shifts) or finite state Markov chains and classical systems with exponential growth of orbit complexity.
- *Infinite entropy* includes many important classes of stationary random processes, such as Wiener or Gaussian, with absolutely continuous spectral measure, and various infinite-dimensional dynamical systems, which appear for example in certain models of fluid mechanics or statistical physics.

This trichotomy alone already justifies calling entropy the most important numerical quantity in dynamics.

From this point on several parallel lines of development clearly emerge, and we will treat those in succession, starting with internal developments in ergodic theory.

2. Isomorphism problem: from Kolmogorov's work through the seventies

Kolmogorov entropy provides a precise instrument for complete classification of Bernoulli shifts, Markov chains, and many other natural classes of systems but fails do to do so for K-systems. As a by-product, a powerful new theory was developed which considers a weaker notion of equivalence related to time change in flows.

a. **Two isomorphism problems of Kolmogorov.** Kolmogorov's stated motivation for the introduction of entropy was to provide a new isomorphism invariant for measure-preserving transformations and flows—more specifically, to split the classes of transformations and flows with countable Lebesgue spectrum into continuum nonisomorphic classes.

In particular, *Bernoulli shifts* (independent stationary random processes) with different entropies (such as $(1/2,1/2)$ and $(1/3,1/3,1/3)$) are not isomorphic.

Two new central problems were formulated:

Are Bernoulli shifts with the same entropy *isomorphic?*

Are K-systems with the same entropy isomorphic?

The short answers turned out to be "yes" in the first case and "no" in the second, but more informative answers would roughly be "YES, and many others too", and "NO, and classification is about as hopeless as for all (measure-preserving) transformations."

b. **Early special cases of isomorphism.** [93, 13] [6]

Meshalkin examples. The first class of examples of nontrivial coding (this is what finding an isomorphism amounts to) was found by L.D. Meshalkin a few months after Kolmogorov asked his question. It can be illustrated by the simplest example of Bernoulli shifts with distributions $(1/4,1/4,1/4,1/4)$ and $(1/2,1/8,1/8,1/8,1,8)$, both with entropy $2\log 2$. The method was known internally as "re-gluing," which gives some indication of its nature.

Finitary codes and finitary isomorphism. Meshalkin and Blum–Hansen codes provide the earliest examples of what later became known as "finitary codes," which produce informationally and geometrically more satisfying constructions for isomorphism between different Bernoulli shifts than the Ornstein method. The 1967 work by Adler and Weiss [5], which achieved considerable fame for several reasons, belongs to the same line of development. It was later shown that finitary codes exist for all Bernoulli shifts as well as for transitive Markov chains with equal entropy [69, 70]. This concept and the accompanying notion of "finitary isomorphism" between stationary random processes is natural from the information theory viewpoint. It can be explained as follows.

Any isomorphism or coding between stationary processes with finite or countable sets of states associates to a realization $\mathbf{x} = (x_n)_{n\in\mathbb{Z}}$ of one process the realization $(y_n(\mathbf{x}))_{n\in\mathbb{Z}}$ of the other. The coding is called *finitary* if for almost every $\mathbf{x}$, the value $y_0(\mathbf{x})$ depends only on a finite segment of $\mathbf{x}$, whose length however depend on $\mathbf{x}$.[7] The same has to be true for the inverse.

However, this is not an isomorphism invariant. In other words, the same measure-preserving transformation T may have two generating partitions, ξ and ζ, such that stationary random processes (T,ξ) and (T,ζ) are not finitarily isomorphic, *i.e.,* one cannot be obtained from the other by a finitary code. For this reason, this line of development, while of great importance for symbolic dynamics, turned out to be conceptually less relevant for ergodic theory and its applications to classical dynamics than Ornstein's work discussed below.

c. **Sinai's weak isomorphism theorem** [132]**.** Sinai came very close to proving isomorphism of Bernoulli shifts with the same entropy. Moreover, he showed that in a natural sense, Bernoulli shifts are the simplest ergodic transformations with a given entropy.

[6] D. Kazhdan and J. Bernstein around 1965, both around age twenty at the time, found more cases of isomorphisms that cover those by Meshalkin, Blum–Hansen and much more, but this work remained unpublished.

[7] If the length is bounded, the code is called finite. This is a very restrictive notion in the context of ergodic theory.

Sinai proved that for any ergodic transformation, T, and any Bernoulli shift, B with entropy $h(T)$, there is a factor of T isomorphic to B. Hence any two Bernoulli shifts with equal entropy are *weakly isomorphic*: each is isomorphic to a factor of the other.

This, of course, gives insight into the problem of splitting of ergodic positive entropy transformations into deterministic and stochastic parts, as mentioned above: Every ergodic positive entropy transformation has not only the canonical maximal zero entropy factor but also has lots of Bernoulli (and hence K) factors of maximal entropy. Since those factors are independent of the π-factor, this can be viewed as strong evidence in favor of Pinsker's conjecture or an even a stronger statement; it looked as though the only thing left to show was that a Sinai factor and the Pinsker factor generate the whole σ-algebra. This turned out to be an illusion on two counts, as we shall soon see.

d. **Ornstein's proof of isomorphism** [100, 101, 102].

From Kolmogorov problem to Ornstein solution. Kolmogorov's isomorphism problem for Bernoulli shifts has a fairly short history, a mere dozen of years, but is it full of drama. During this period the Russian (primarily Moscow)[8] school of ergodic theory enjoyed the world domination. It should be mentioned that in Moscow, ergodic theory was not treated as a separate discipline but rather as one of the streams within dynamics. As a result, the agenda was amazingly broad and other directions in dynamics figured prominently—such as hyperbolic dynamics and its interface with celestial mechanics, smooth ergodic theory (including applications of entropy), interface with statistical mechanics, and attempts to find fruitful applications to number theory. Sinai's talk at the 1962 Stockholm ICM Congress [133] and the 1967 issue of Uspehi, based on the workings of the 1965 Humsan school [1], are representative both of this position and of the agenda. Still, the isomorphism problem was never far from the surface. I remember vividly an episode from about 1966 when an information theorist from St. Petersburg, R. Zaidman, claimed that he had a solution to the isomorphism problem for Bernoulli shifts. He came to Moscow more than once and attempted to present his solution at the (then) main ergodic theory/dynamics seminar. He did not get very far, although at the time the feeling of many listeners, shared by the author, was that he did make substantial progress in the direction of what is now called the Keane–Smorodinsky finitary isomorphism theorem. This feeling strengthened after Ornstein's solution appeared.

The announcement of Ornstein's solution in 1969 came as a shock to Moscow. The "philosophy" of Ornstein's approach was not absorbed quickly, and the efforts shifted even more than before from "pure" ergodic theory to interface with

[8]Rokhlin who moved to Leningrad (now St. Petersburg) in 1960 left ergodic theory shortly afterward for pursuits in topology and algebraic geometry; however several outstanding young mathematicians working in the area appeared in Leningrad during this period. A. M. Vershik is the most prominent of those; he was to have a long and illustrious career spanning several areas of mathematics.

other disciplines, primarily smooth dynamics and statistical mechanics. Leadership in the development of pure ergodic theory passed to Ornstein, Hillel Furstenberg, and their "schools".[9]

According to Ornstein, his interest in the isomorphism problem was sparked during his visit to Courant Institute by Jürgen Moser's enthusiasm for the recently appeared Adler–Weiss work, in which isomorphism of ergodic automorphisms of the two-dimensional torus with equal entropy was proved. We mentioned this work already and will come back to it in Section 3b. Then, Ornstein learned about Sinai's weak isomorphism theorem and started to work on the isomorphism problem.

Features of the Ornstein solution. Ornstein's solution, which we do not try to sketch here (see *e.g.,* [142] for a fairly concise and readable account), was based on a complete departure from the previous philosophy grounded in probability and coding theory. Instead it made full use of plasticity of the structure of a measure-preserving transformation free of connections with any particular class of generating partitions.

Don Ornstein, 1970

It is interesting to point out that the starting point of Ornstein's consideration was the fundamental, but quite simple, "Rokhlin Lemma," which states that every aperiodic measure-preserving transformation cyclically permutes N disjoint sets with an arbitrarily small error independent of N. In the Ornstein construction, the entropy serves as the controlling parameter, which dictates the number

[9]It is interesting to point out that at that time the author (and probably everyone at Moscow) grossly underestimated, almost ignored, Furstenberg's pioneering disjointness paper [32].

of "names" which can be put into a correspondence between two stationary random processes.

Very weak Bernoulli processes [103]. This flexible approach allowed Ornstein to do something immeasurably more important than proving isomorphism of Bernoulli shifts or mixing Markov chains with the same entropy. He found a *verifiable, necessary and sufficient condition, called* very weak Bernoulli *for a stationary random process to appear as the process generated by a measurable partition for a Bernoulli shift.* Thus if a measure-preserving transformation has a generating partition which determines a very weak Bernoulli process, it is isomorphic to any Bernoulli shift with the same entropy.

Quickly, many natural classes of systems of probabilistic, algebraic, and geometric origin were proven to possess very weak Bernoulli generating partitions and, hence, to be isomorphic to Bernoulli shifts. As an early example, let us mention the work of Yitzhak Katznelson [68], where he shows that an arbitrary ergodic (not only hyperbolic) automorphism of a torus in any dimension is Bernoulli—a big step forward from the Adler–Weiss work for the two-dimensional torus which motivated Ornstein.

Another instance which demonstrates the power of Ornstein's work and the revolutionary changes it produced in ergodic theory is the solution of the square root problem. One should keep in mind that before Ornstein, the question of whether a particular Bernoulli shift, such as (1/2, 1/2), has a square root was intractable. Of course, certain Bernoulli shifts *e.g.,* (1/4, 1/4, 1/4, 1/4) are squares of others and the isomorphism theorem for Bernoulli shifts then implies immediately that *any* Bernoulli shift with entropy $2\log 2$ has a square root. But this does not work for many other values of entropy. Furthermore, since no Bernoulli shift embeds into a flow in a natural way, the isomorphism theorem *per se* is useless in answering the question of embedding any particular Bernoulli shift into a flow. A positive solution of this problem, which is amazing from the pre-Ornstein viewpoint, is the subject of a separate paper [103] whose principal importance is that in it the very weak Bernoulli property is introduced. This paper contains a rather artificial construction of a "Bernoulli flow". But starting from the Ornstein–Weiss paper on geodesic flows [112], a pattern was established by finding very weak Bernoulli partitions in naturally appearing classes of maps, including those which are parts of flows.

e. **Development of the Ornstein theory.** Beyond checking the very weak Bernoulli property for various classes of naturally appearing and specially constructed transformations, the work of Ornstein and his collaborators and followers included two strands.

(i) Extension of the isomorphism results and accompanying criteria of finitely-determined and very weak Bernoulli form the basic case of measure-preserving transformations to flows [107], "relative" situation [141], actions of increasingly general classes of groups crowned by the Ornstein–Weiss isomorphism theory for actions of amenable groups [114]. Here, entropy, properly extended to the case of group actions, plays, as before, the role of a single controlling parameter in the otherwise very fluid situation. It should be remembered that for measure-preserving actions of finitely generated groups other than finite extensions of $\mathbb{Z}$, entropy is defined in such a way that in positive-entropy actions, many individual elements must have infinite entropy.

(ii) Development of various constructions showing existence or compatibility of various, often exotic, properties; here a “zoo” of non-Bernoulli K-systems constitutes the central exhibit. As it turned out, the gap between Bernoulli and K-systems is huge [110]. Those K-systems are extensions of their maximal entropy Bernoulli factors, with no additional entropy generated in the fibers. Not only are there many nonisomorphic K-systems with the same entropy, but also K-property can be achieved in unexpected ways, *e.g.,* by changing time in any ergodic positive entropy flow [111], or inducing (taking the first return map) of any ergodic positive entropy transformation on a certain set [113]. The trend here was from “made-to-order” artificial constructions toward showing that some natural systems exhibit exotic properties. An excellent illustration is the proof of Kalikow [47] that the extension of a Bernoulli shift with two symbols, call them 1 and -1, by moving along orbits of another Bernoulli shift forward or backward according to the value of the zero coordinate in the base is K but not Bernoulli. Soon afterward, Rudolph [125] found a classical K but not Bernoulli systems with similar behavior.

Ornstein also disproved the Pinsker Conjecture by another ingenuous counterexample [108].

To summarize, the burst of creative activity in 1970–73 following Ornstein's proof of isomorphism of equal entropy Bernoulli shifts, revolutionized ergodic theory once more and can be compared in its impact with the 1958–62 period discussed above which followed the introduction of entropy by Kolmogorov. An important difference is that during this period Ornstein, although ably assisted by collaborators, played the role which can be compared with the combined roles of Kolmogorov (the original breakthrough), Sinai (the principal mover) and Rokhlin (another important mover and the key expositor) during the earlier period. Ornstein's role in disseminating the new theory and establishing the standards of expositions is attested by the surveys [104, 105] and the monograph [109].

f. **Kakutani (monotone) equivalence** [52, 53, 28, 113]**.** The Ornstein isomorphism theory turned out to possess a very interesting “cousin,” which, in certain aspects at least, has surpassed the original in importance and applicability. Again we will not attempt to sketch the proofs but formulate the problem, outline the results, and explain the sources of parallelism with the isomorphism problem and the Ornstein theory.

The starting point is a natural equivalence relation between ergodic systems, weaker than the isomorphism. It has two dual forms. For continuous time systems, this is a *time* (or velocity) *change*, a well-known concept in the theory of ordinary differential equations when the vector field defining a flow (a one-parameter group of transformations) is multiplied by nonvanishing scalar function ρ. Any invariant measure for the old flow multiplied by the density ρ^{-1} is invariant for the new flow. Equivalently, one may represent any measure-preserving flow as a *special flow* or *flow under a function* (often called the *roof function*) over a measure-preserving transformation. Then flows are equivalent if they are isomorphic to special flows over the same transformation.

Discrete time systems are equivalent if they appear as sections of the same flow. Equivalently, one allows for taking an induced transformation (the first return map) on a set of positive measure and the inverse operation, which is the discrete equivalent of building the flow under a function.

Both of these equivalence relations are most commonly called *Kakutani equivalence*, although the author who considered it in [52, 53] prefers a descriptive name *monotone equivalence*.

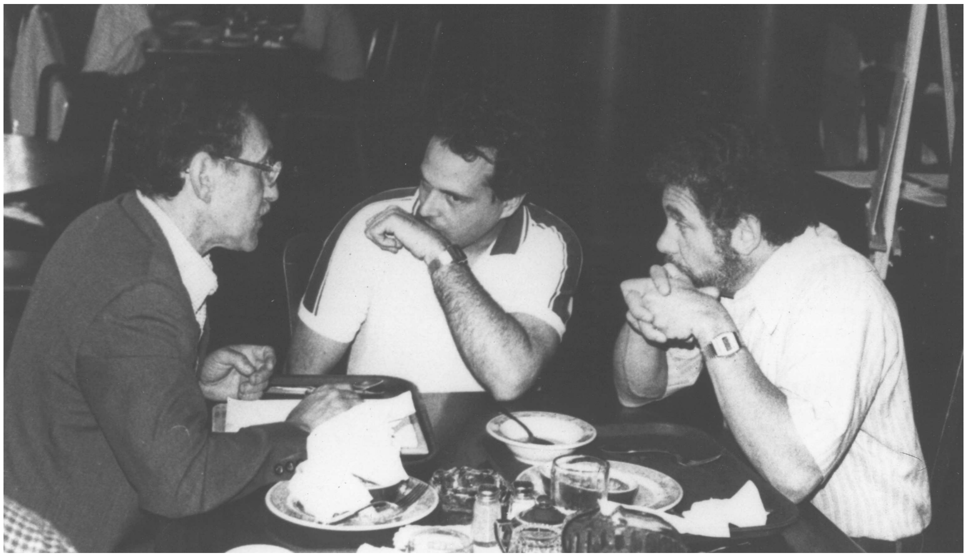
Notice that entropy is not an invariant of the monotone equivalence; however, the three big classes mentioned above—zero, positive finite, and infinite entropy—are invariant.[10]

At the basis of Ornstein isomorphism theory lies the concept of $\bar{d}$-distance between stationary random processes, see *e.g.,* [142]. It measures how well the different "n-names" which appear in those processes can be matched as $n \to \infty$. The n-name with respect to a partition ξ is simply an element of the partition $\bigvee_{i=0}^{n-1} T^n(\xi)$. For the distance between two n-names, one takes the natural *Hamming distance i.e.,* the proportion of places where the names differ. For the purposes of the monotone equivalence theory, Hamming distance is replaced by a weaker Kakutani distance, which can be defined as the normalized minimal number of *elementary operations* needed to produce one name from the other. An elementary operation consists of removing a symbol and placing another symbol in any other place. [11] After that replacement, one can follow the principal steps of the Ornstein constructions essentially verbatim and obtain necessary and sufficient conditions for a stationary random process to appear from a measure-preserving transformation monotone equivalent to a Bernoulli shift with finite or infinite entropy. Such transformations are usually called *loosely Bernoulli*.

There is, however, a remarkable difference which is responsible for the great importance of the monotone equivalence theory for zero entropy transformations. The zero entropy case of both the Ornstein theory and the monotone equivalence theory correspond to processes for which most names are simply

[10] A natural modification which amounts to fixing the average of the roof function or considering transformations which induce isomorphic transformations on sets of equal measure produces a stronger relation of *even Kakutani equivalence* which preserves the numerical value of entropy.

[11] If replacement is allowed only at the same place one naturally obtains the Hamming distance.



Jack Feldman, the author, and Don Ornstein, 1981

very close to each other in the corresponding metric. In the case of Hamming metric, this is only possible for the trivial transformation acting on the one-point space, or, equivalently, to the Bernoulli shift with zero entropy. Thus, it has no bearing on the study of zero entropy transformations on the standard Lebesgue space. However, in the case of the Kakutani metric, such a situation is possible and it defines a particular class of monotone equivalent transformations called *standard* or, sometimes, *loosely Kronecker*. It is represented, for example, by the dyadic adding machine or by any irrational rotation of the circle.

A key point in the early development of the theory was construction of a non-standard zero entropy transformation by Feldman [28].

Standard transformations are naturally the simplest transformations of Lebesgue space from the point of view of monotone equivalence: every measure-preserving transformation is equivalent to a transformation which has a given standard transformations as a factor. Notice that any zero entropy factor of a loosely Bernoulli transformation is standard.

Monotone equivalence is a useful source of counterexamples in the isomorphism theory. For example, if one constructs a K-automorphism T which is monotone equivalent to a transformation with a nonstandard zero entropy factor, then T is not loosely Bernoulli and hence not Bernoulli. This observation was used in the earliest construction of a classical system which is K but not Bernoulli [56].

3. TOPOLOGICAL ENTROPY, VARIATIONAL PRINCIPLE, SYMBOLIC DYNAMICS: 1963–1975

Topological entropy is a precise numerical measure of global exponential complexity in the orbit structure of a topological dynamical system. In a variety of situations, topological entropy is equal to the exponential growth rate for the number of periodic orbits in a dynamical system and, furthermore, it fully determines a more precise multiplicative asymptotic for that quantity. It has many important cousins which come from considering weighted orbit averages. Invariant measures capturing statistically the full complexity (plain or weighted) are often unique and possess nice statistical properties.

a. **Definition of topological entropy and comparison with Kolmogorov entropy** [3, 38, 22, 36]**.** The spectacular success of entropy in ergodic theory led to the development of a topological counterpart of entropy. One could hope that this concept would revolutionize topological dynamics in a similar way. This was borne out if not for arbitrary topological dynamical systems on compact spaces then at least for broad classes of systems especially important for applications to classical mechanics and statistical physics.

The original 1963 definition[12] imitated Shannon's development, replacing partitions by covers and, for lack of a notion of size, replacing weighted average of logarithmic size by the maximum of such expressions for a given number of elements.

Let $\mathcal{A}$ be an open cover of a compact space X. Then, $C(\mathcal{A})$ (as in "cover") denotes the minimal cardinality of a subcover, $\Phi^{-1}(\mathcal{A}) := \{\Phi^{-1}(A) \mid A \in \mathcal{A}\}$ and

$$h_{\text{top}}(\Phi) := \sup_{\mathcal{A}} \lim_{n\to\infty} \frac{1}{n} \log C(\bigvee_{i=0}^{n-1} \Phi^{-i}(\mathcal{A}))$$

where the supremum is taken over all open covers.[13]

Nowadays, an equivalent approach due to Rufus Bowen is more popular. Define

$$d_t^{\Phi}(x,y) = \max_{0\le\tau<t} d(\Phi^\tau(x), \Phi^\tau(y)), \tag{3}$$

measuring the distance between the orbit segments $\mathcal{O}^t(x) = \{\Phi^\tau(x) \mid 0 \le \tau < t\}$ and $\mathcal{O}^t(y)$. Denote by $C(\Phi,\epsilon,t)$ the minimal number of ϵ-balls with respect to d_t^Φ that cover the whole space. We define the topological entropy by

$$h(\Phi) := h_{\text{top}}(\Phi) := \lim_{\epsilon\to 0} \overline{\lim_{t\to\infty}} \frac{1}{t} \log C(\Phi,\epsilon,t) = \lim_{\epsilon\to 0} \underline{\lim_{t\to\infty}} \frac{1}{t} \log C(\Phi,\epsilon,t). \tag{4}$$

[12] The paper [3] appeared in 1965 but was submitted for publication in the fall of 1963; in fact, R. Adler lectured about topological entropy as early as spring of 1963 (B. Weiss, private communication).

[13] This is a great convenience of compactness. This definition is independent of a choice of metric. But in various partial extensions to noncompact spaces this straightforward approach does not work whereas various definitions using a metric can be extended.

It is not hard to show that these two expressions coincide and are independent of the metric.

Note that (changing ϵ by a factor of 2) one also obtains topological entropy in the same manner from the maximal number $S(\Phi,\epsilon,t)$ of points in X with pairwise d_t^Φ-distances at least ϵ. We call such a set of points (t,ϵ)-*separated*. Such points generate the maximal number of orbit segments of length t that are distinguishable with precision ϵ. Thus, entropy represents the exponential growth rate of the number of orbit segments distinguishable with arbitrarily fine but finite precision, and describes in a crude but suggestive way the total exponential complexity of the orbit structure with a single number.

To compare topological and measure-theoretic entropy nowadays, one defines analogously $C(\Phi,\epsilon,t,\delta)$ to be the minimal number of ϵ-balls with respect to d_t^Φ whose union has measure at least $1-\delta$. If Φ is ergodic and $\delta\in(0,1)$ then [55]

$$h_\mu(\Phi)=\lim_{\epsilon\to0}\overline{\lim_{t\to\infty}}\frac{1}{t}\log S(\Phi,\epsilon,t,\delta)=\lim_{\epsilon\to0}\underline{\lim_{t\to\infty}}\frac{1}{t}\log S(\Phi,\epsilon,t,\delta), \tag{5}$$

i.e., in the case of an ergodic measure-preserving transformation, its metric entropy can be characterized as the exponential growth rate for the number of *statistically significant* distinguishable orbit segments. Note that this is clearly never more than the topological entropy. However, historically it took some time to realize this relation, see Section 3e.

While in comparison with Kolmogorov entropy the notion of topological entropy lacks "structural depth" and in particular is not accompanied by proper equivalents of the notions of K-systems and π-partition,[14] this is compensated by its fundamental connections with metric (Kolmogorov) entropy via variational principle and the construction of measures capturing the global complexity for broad classes of dynamical systems. One should also mention an important line of development which provides a partial analogy with the Ornstein theory for almost topological conjugacy for topological Markov chains [4].

b. **Parry measure and Adler–Weiss partitions.** [116, 5, 6] The first step in this direction was made by William Parry.[15] For the leading class of *topological Markov chains*, Parry constructed a unique measure whose entropy turned out to be equal to the topological entropy and which is characterized by this property.

Let $A=\{a_{ij}\}$ be a transitive $N\times N$ matrix with 0-1 entries (transitivity means that some power of A has only positive entries) and consider the associated *topological Markov chain* σ_A, *i.e.,* the restriction of the shift transformation in the space of double-infinite sequences from N symbols numbered, say $\{1,\dots,N\}$, to the subset Ω_A of those sequences $\{\omega_n\}_{n\in\mathbb{Z}}$ for which $a_{\omega_n\omega_{n+1}}=1$ for all n.

[14] As a matter of record, one should point out that much later two topological counterparts of K-property were suggested by F. Blanchard [12] and a partial analogy between the measurable and topological cases was developed to a greater degree [35].

[15] Parry's note appeared before [3], but Parry must have been familiar with the notion of topological entropy at the time.

Let $q=(q_1,\dots,q_N)$ and $v=(v_1,\dots,v_N)$ be positive eigenvectors of A and A^T, respectively (those are unique up to a scalar), such that $\sum_{i=1}^N q_i v_i = 1$, and set $p=(p_1,\dots,p_N):=(q_1v_1,\dots,q_Nv_N)$ and $\pi_{ij}=a_{ij}v_i/\lambda v_j$. Then, $\sum_{i=1}^N \pi_{ij}=1$ for all j and $\pi_{ij}>0$ whenever $a_{ij}=1$, so the matrix $\Pi=\{\pi_{ij}\}_{i,j=1,\dots,N}$ is also transitive. Then, the *Parry measure* defined by

$$\mu_\Pi(\{\omega\in\Omega_A \mid \omega_i=\alpha_i,\ i=-m,\dots,m\})=\Big(\prod_{i=-m}^{m-1}\pi_{\alpha_i\alpha_{i+1}}\Big)p_{\alpha_m}$$

is invariant under σ_A, and its entropy is equal to $h_{\mathrm{top}}(\sigma_A)$.[16]

Adler and Weiss realized that hyperbolic automorphisms of two-dimensional torus can be represented as "almost" topological Markov chains by taking pieces of the stable (contracting) and unstable (expanding) eigenlines at the origin and extending them until the whole torus gets divided in several parallelograms.[17] The Markov property is self-evident from that construction which became an icon for hyperbolic dynamics. By taking an iterate of this partition and taking a finer partition into connected components of the intersection, one guarantees that this partition is a topological generator and hence can be used to code the map without essential loss of information.[18]

Adler and Weiss show that automorphisms with the same entropy with respect to Lebesgue measure are isomorphic as measure-preserving transformations. This relies on the observation that the Parry measure in the topological Markov chain (symbolic) representation corresponds to the Lebesgue measure on the torus. Notice that both are unique maximal entropy measures. Thus, for hyperbolic automorphisms of a two-dimensional torus it was shown (before Ornstein's work on isomorphism of Bernoulli shifts as we explained in Section 2dd) that entropy is the complete invariant of metric isomorphism. Notice that equality of entropies is equivalent to conjugacy of corresponding matrices $\mathbb{C}$ or $\mathbb{Q}$ but not necessarily over $\mathbb{Z}$. It is also equivalent to equality of topological entropies, since for an automorphism of a torus, the latter is equal to the metric entropy with respect to Lebesgue measure.

[16] According to Adler, Shannon had a concept equivalent to the topological entropy for topological Markov chains and more general sofic systems under the name of channel capacity of discrete noiseless channels. Furthermore, he proves the result about existence and uniqueness of measure of maximal entropy equivalent to that of Parry. However, this remained unnoticed by the dynamics community until well after Parry's work.

[17] Ken Berg in his 1967 thesis [9] arrived at the same construction. According to Adler, Berg did it before Adler and Weiss.

[18] The picture of the Adler–Weiss partition for the automorphism determined by the matrix $\begin{pmatrix}2&1\\1&1\end{pmatrix}$ has become one of the most widely recognized icons of modern dynamics. It appears in innumerable articles and texts; in particular, its plane unfolding can be found on the cover of the hardcover edition of [59] and the much more picturesque multicolored rendering on the surface of a bagel standing up on the cover of its paperback edition; see also a slanted version of the latter at the dust jacket of [20].

William Parry and Roy Adler, 1968

c. **General Markov partitions** [136, 15, 16]**.** The Adler–Weiss construction ostensibly depends on two-dimensionality of the system: it uses the fact that boundaries of partition elements consist of pieces of single stable and unstable manifolds. This is indeed necessary if elements of a partition with Markov properties are domains with piecewise smooth boundaries. This geometric structure is, however, not at all necessary in order for a Markov partition to be a useful tool for the study of dynamics of a smooth system. This was realized by Sinai who came up with a construction which is quite natural dynamically but, in general, even for automorphisms of a torus of dimension greater than two, produces Markov partitions whose elements have boundaries of fractal nature.

Finally, it was Bowen—the rising superstar of dynamics in the late 1960s and 1970s, whose meteoric rise and sudden early death gave this period a dramatic quality—who gave the theory of Markov partitions both proper generality (for locally maximal hyperbolic sets) and its elegant ultimate form based on systematic use of shadowing and free of unnecessary geometric allusions thus avoiding difficulties and pitfalls of Sinai's original more geometric approach.

d. **Topological entropy and growth of periodic orbits.** For a discrete time dynamical system $f\colon X \to X$ one defines $P_n(f)$ as the number of fixed points for f^n; for a flow Φ let $P_t(\Phi)$ be the number of periodic orbits of period $\le t$.

Bowen's definition of topological entropy as the exponential growth rate of number of separated orbit segments indicates that it may be a connected with the exponential growth rates of the number of *periodic* orbits, *i.e.,*

$$\limsup_{n\to\infty} \frac{\log P_n(f)}{n} \quad \text{and} \quad \limsup_{n\to\infty} \frac{\log P_t(\Phi)}{t}.$$

or the corresponding limits if they exist. In two model cases, automorphisms of an n-dimensional torus from Section 1d [19] and transitive topological Markov chains from Section 3b, a simple calculation gives correspondingly

$$P_n(F_A) = |\det(A^n - \mathrm{Id})| \tag{6}$$

and

$$P_n(\sigma_A) = \mathrm{tr}\, A^n. \tag{7}$$

Thus,

$$\lim_{n\to\infty} \frac{\log P_n(F_A)}{n} = \sum_{\lambda \in \mathrm{Sp}\, A} (\log|\lambda|)_+ \tag{8}$$

and

$$\lim_{n\to\infty} \frac{\log P_n(\sigma_A)}{n} = \log \lambda^A_{\max}, \tag{9}$$

where the first expression coincides with the right-hand side of (2) and in the second, $\lambda^A_{\max}$ is the unique eigenvalue of maximal absolute value. In the torus case, the entropy with respect to Lebesgue measure coincides with the topological entropy and for the topological Markov chain $\log \lambda^A_{\max}$ is the common value of the entropy with respect of Parry measure and topological entropy. Formulas (8) and (9) are prototypes for the expressions of the exponential growth rate of periodic orbits for larger classes of systems where those rates exist and coincide with topological entropy. Notice that both the exact expressions (7) and (6) (in the case when A has no eigenvalues of absolute value one) allow to obtain an exponential above estimate of the error terms $P_n(\sigma_A) - \exp n h_{top}(\sigma_A)$ and $P_n(F_A) - \exp n h_{top}(F_A)$ correspondingly.

Existence and statistics of closed geodesics on Riemannian manifolds and closed orbits for various special classes of Hamiltonian systems are problems of great interest not only for dynamics but also for various branches of analysis, geometry, and even number theory. The problem of existence of closed geodesics is one of the central problems in Riemannian geometry. It was noticed long before Kolmogorov's work on entropy that negative curvature leads to exponential growth of the number of closed geodesics; moreover, as early as 1961, H. Huber [46] found the multiplicative asymptotic with an error term estimate for the number of closed geodesics on a compact surface of constant negative curvature $-k^2$.[20] The exponent is k which in this case coincides with topological entropy.

[19] With the assumption that no eigenvalue of A is a root of unity which is equivalent to periodic points being isolated.

[20] The case of surfaces of constant negative curvature, both compact and especially noncompact but of finite volume, such as the modular surface, is of particular interest to number theory.

The method is based on the use of the Selberg trace formula and is hence restricted to the symmetric spaces. Margulis' principal motivation in his seminal Ph.D. thesis work [92], which at the time was published only partially [90, 91], was proving a similar multiplicative asymptotic for manifolds of variable negative curvature.

e. **The variational principle for entropy** [38, 22, 36]**.** It was noticed right after the introduction of topological entropy that in many topological or smooth systems, topological entropy coincides with entropy with respect to an invariant measure, either natural, like Lebesgue measure for an automorphism of a torus, or Haar measure for an automorphism of a general compact abelian group, or specially constructed, as the Parry measure for a transitive topological Markov chain (see Section 3b above). This naturally brings up a question about relationships between the topological entropy of a continuous transformation on a compact metrizable space and metric entropies of this transformation with respect to various Borel probability invariant measures. Of course, with the definitions of topological and metric entropy described in Section 3a the inequality

$$h_\mu(f) \leq h_{top}(f)$$

for any *ergodic* (probability Borel) f-invariant measure μ is obvious. Those definitions are however of later vintage; in particular, the one for metric entropy only appeared in my 1980 paper [55]. With original definitions using partitions and coverings, the inequality was established by L. Goodwyn and published in 1969. Relatively quickly it was shown that the inequality was sharp; in other words that there are invariant measure with entropies arbitrarily close to the topological entropy. The very substantial paper by E.I. Dinaburg which relied on an extra (unnecessary, as it turned out) assumption came a year earlier than a short and elegant note by T.N.T. Goodman, which established the *variational principle for entropy* in full generality for continuous maps of compact Hausdoff spaces:

$$h_{top}(f) = \sup h_\mu(f)$$

where supremum is taken all Borel probability f-invariant measures (or, equivalently only over ergodic ones).

The variational principle can be interpreted as a quantitative version of the classical Krylov–Bogolyubov Theorem [44], which asserts existence of at least one Borel probability invariant measure for any continuous map of a Hausdorff compact. In retrospect, proofs of both facts are fairly simple and somewhat similar: one constructs an almost invariant measure by putting a normalized δ-measure on a long orbit segment (in the case of Krylov–Bogolyubov Theorem), or on a maximal number of separated orbit segments (in the case of the Variational Principle), and takes a weak-* limit point as the orbit length tends to infinity. In the case of the Variational Principle, a simple convexity argument asserts that entropy of the limit measure is sufficiently large. The reason one in general cannot guarantee that the entropy is *equal* to the topological entropy is the need to take a limit in the degree of separation. However, if the map is *expansive, i.e.,* if any two different semiorbits (orbits in the case of a homeomorphism) diverge to a fixed distance, the argument does produce a measure of maximal entropy which, however, may still not be unique. For details see *e.g.,* [59].

f. **Maximal entropy measures for hyperbolic and symbolic systems.** [135, 91, 17]. To guarantee uniqueness of a maximal entropy measure, one needs to add some extra assumptions to expansivity. Two essential classes of positive entropy systems for which this happens are transitive topological Markov chains, discussed in Section 3b, and restrictions of smooth systems (diffeomorphisms and flows) to basic (transitive locally maximal) hyperbolic sets.

There are three different constructions of this unique measure due to Sinai, Margulis, and Bowen. Each of these constructions shows a different facet of the picture and engendered its own set of fruitful generalizations and developments.

Yakov Grigorievich Sinai, 1976

Sinai's construction goes through a Markov partition for which one can then refer to the Parry measure and needs only a small argument to show that no maximal entropy measure is lost due to the nonuniqueness of the correspondence with a Markov chain on a "small" set. However, Sinai had in mind a deeper connection with lattice models in statistical mechanics, which inspired his next great contribution [137] discussed below in Section 3g.

Margulis' motivation was to find a precise asymptotic for the growth rate of closed geodesics on a compact manifold of negative sectional curvature (see Section 3d) and his work was seminal for the whole range of applications of dynamical ideas (and entropy in particular) to Riemannian geometry. However geometrically inspired, his construction has a clear meaning in the phase space language for Anosov systems. First conditionals on stable and unstable manifolds are constructed as pullbacks of the asymptotic distribution of volume in the expanding direction of time (positive for unstable and negative for stable manifolds), and the global measure is obtained as the product of those conditionals with the time parameter along the orbits. It does not work for more general hyperbolic sets, since either stable or unstable manifolds or both are not contained in the set and hence most of the volume escapes any control.

Bowen gave an elegant axiomatic treatment of the subject using minimal essential tools. Conceptually he constructed the desired measure as the asymptotic distribution of periodic orbits in a system where orbit segments can be concatenated and closed more or less at will. The latter property is formally expressed as *specification* [59]. Let (X,d) be a compact metric space and $f\colon X\to X$ an expansive homeomorphism with the specification property. Then there is exactly one f-invariant Borel probability measure with $h_\mu(f)=h_{\text{top}}(f)$. It is called the *Bowen measure* of f and is given by

$$\mu=\lim_{n\to\infty}\frac{1}{P_n(f)}\sum_{x\in\operatorname{Fix}f^n}\delta_x, \tag{10}$$

where δ_x denotes the probability measure with support $\{x\}$.

All three constructions allow one to show that for the corresponding classes of systems the exponential growth rate for the number of periodic orbits defined in Section 3d exists and coincides with the topological entropy generalizing (8) and (9).

In fact, they give more precise asymptotic for $P_n(f)$ and $P_t(\Phi)$ than exponential.

Sinai's construction based on Markov partitions and the use of Parry measure is applicable to topologically mixing locally maximal hyperbolic sets for diffeomorphisms and provides for such a set Λ for a diffeomorphism f an exponential estimate of the error term

$$P_n(f\mid_\Lambda)-\exp nh_{top}(f\mid_\Lambda),$$

see *e.g.,* [59, Theorem 20.1.6].

Margulis' construction works for topologically mixing Anosov flows. It gives the multiplicative asymptotic

$$\lim_{t\to\infty}\frac{P_t(\Phi)\,th_{top}(\Phi)}{\exp th_{top}(\Phi)}=1. \tag{11}$$

Bowen's construction for expansive homeomorphisms with specification property gives somewhat weaker multiplicative bounds

$$0<c_1\le\liminf_{n\to\infty}\frac{P_n(f)}{\exp nh_{top}(f)}\le\limsup_{n\to\infty}\frac{P_n(f)}{\exp nh_{top}(f)}\le c_2.$$

It can be extended to the continuous time case.

g. **Pressure, Gibbs measures and equilibrium states** [137, 18, 16]**.** Bowen's construction (10) makes considering *weighted averages* of δ-measures on periodic orbits quite natural. However, originally, a construction based on weighted distribution was introduced by Sinai and was motivated by the construction of

Gibbs states in statistical mechanics. Technically, it was performed through the construction of Markov partitions and reduction to the case of topological Markov chains. For a specific way of weighting this already appeared in [135], where measures serving as asymptotic forward and backward limit of absolutely continuous distributions were constructed for Anosov systems. In full generality, this approach was developed in [137], the paper which firmly established "thermodynamical formalism" as one of the cornerstones of modern dynamics.

The basic construction of these special measures, called *equilibrium states* (as opposed to more subtle estimates), works in the setting suggested by Bowen.

First one defines a weighted version of topological entropy. Let X be a compact metric space, $f\colon X \to X$ a homeomorphism, and φ a continuous function, sometimes called the *potential*. For $x \in X$, and a natural number, n, define $S_n\varphi(x) := \sum_{i=0}^{n-1}\varphi(f^i(x))$. For $\epsilon > 0$, let

$$S(f,\varphi,\epsilon,n) := \sup\Big\{\sum_{x\in E}\exp S_n\varphi(x) \mid E\subset X \text{ is } (n,\epsilon)\text{-separated}\Big\}, \tag{12}$$

$$C(f,\varphi,\epsilon,n) := \inf\Big\{\sum_{x\in E}\exp S_n\varphi(x) \mid X = \bigcup_{x\in E} B_f(x,\epsilon,n)\Big\}. \tag{13}$$

These expressions are sometimes called *statistical sums*. Then,

$$P(\varphi) := P(f,\varphi) := \lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}}\frac{1}{n}\log C(f,\varphi,\epsilon,n)$$

is called the *topological pressure* of f with respect to φ, and, for an f-invariant Borel probability measure, μ, the *pressure* of μ is defined as

$$P_\mu(\varphi) := P_\mu(f,\varphi) := h_\mu(f) + \int\varphi\, d\mu.$$

The *variational principle for pressure* asserts that

$$P(\varphi) = \sup\{P_\mu(\varphi)\}.$$

where, as in the case of topological entropy, the supremum is taken over the set of all Borel probability invariant measures which is convenient to denote $\mathfrak{M}(f)$.

A measure $\mu\in\mathfrak{M}(f)$ is called an *equilibrium state* for φ if

$$P_\mu(f,\varphi) = P(f,\varphi).$$

Obviously, a maximal entropy measure is an equilibrium state for the function $\varphi = 0$

Existence and uniqueness of an equilibrium state μ_φ is established under the same conditions on f as for the maximal entropy measure (expansiveness and specification) for a class of potentials with reasonable behavior. The construction is a direct generalization of (10):

$$\mu_\varphi = \lim_{n\to\infty}\frac{1}{P(f,\varphi,n)}\sum_{x\in \operatorname{Fix} f^n}\exp S_n\varphi(x)\delta_x,$$

where of course $P(f,\varphi,n)$ is the statistical sum counted over all point of period n rather than over a separated set.

The following condition for coincidence of equilibrium states, which is obviously sufficient, turned out also to be necessary in this setting

$$\mu_\varphi = \mu_\psi \quad \text{iff for any } x \in \operatorname{Fix} f^n \; S_n\varphi(x) = S_n\psi(x) + cn \text{ for some } c \in \mathbb{R}.$$

Basic properties of equilibrium state such as mixing can be established in this generality. The original application of the equilibrium state concept in [137] was in the context of a transitive Anosov system (a diffeomorphism or a flow) and a Hölder potential. In this case, the unique equilibrium state exists and is in fact a Bernoulli measure. It also possesses important noninvariant properties, such as the *central limit theorem*, and under additional assumptions, the *exponential decay of correlations*. This extends to basic hyperbolic sets and Hölder potentials.

4. ENTROPY AND NONUNIFORM HYPERBOLICITY: 1975–2002

Entropy in smooth systems comes exclusively from infinitesimal exponential speed of orbit separation; it is fully determined by those in classical systems and is described by additional geometric dimension-like quantities in the general case. Nonuniform hyperbolicity also provides the setting where entropy gives the lower bound on the exponential growth rate of the number of periodic orbits. Any (nonuniformly) hyperbolic measure *can be approximated in various senses by uniformly* hyperbolic sets*; their topological entropies approximate the entropy of the measure. In the low-dimensional situations this connection is universal.*

a. **Entropy in uniformly hyperbolic setting.** It was noticed fairly early in the development of entropy theory that formula (2) expressing entropy with respect to invariant volume as the total exponential rate of expansion in all possible directions holds in the setting when this expansion rate is not constant in space. Namely, let for a volume-preserving Anosov diffeomorphism $f\colon M \to M$ J^u be the expansion coefficient of the volume in the direction of the unstable (expanding) distribution. Then,

$$h_{vol}(f) = \int_M J^u \, d(vol). \tag{14}$$

A similar statement is mentioned without elaboration already in Sinai's 1962 Stockholm ICM talk [133] and is proved in [134]; applications to nonalgebraic examples require a proof of absolute continuity of stable and unstable foliations [7, 8].

b. **Lyapunov characteristic exponents, Pesin entropy formula and Pesin theory** [117, 118]**.** There is a gap between the rather crude Kushnirenko above estimate of entropy for general smooth systems and exact formula (14). It was probably Sinai who conjectured that the latter formula could be generalized if properly defined exponential rates of infinitesimal growth were used. The question was posed already by the fall of 1965, since it was discussed during the memorable Humsan school [1]. The answer comes from the theory of *Lyapunov characteristic exponents* and the Oseledet's *Multiplicative ergodic theorem* [115] which establishes existence of those exponents in a smooth system almost everywhere with respect to any invariant measure. In fact, it seems that desire to find a general formula for entropy was the principal motivation behind Oseledet's work. Thus, the formula eventually proved by Pesin for an arbitrary $C^{1+\epsilon}$, $\epsilon > 0$ diffeomorphism f of a compact m-dimensional manifold M preserving an absolutely continuous measure μ takes the form

$$h_\mu(f) = \int_M \sum_{i=1}^m (\chi_i)_+ d\mu. \tag{15}$$

Here each Lyapunov exponent χ_i is counted as many times as it appears, *i.e.,* with multiplicity, so that one can always assume that there are exactly m (not necessarily different) Lyapunov exponents. If the measure μ is ergodic, the exponents are constant almost everywhere and hence integration can be omitted. But it is also useful to represent the entropy as the average rate of infinitesimal expansion. For that, one defines almost everywhere the unstable distribution, E^u_f, as the sum of the distributions corresponding to positive Lyapunov exponents and the unstable Jacobian, J^u, as the expansion coefficient of the volume in the direction of E^u_f. Then, formula (15) takes the form of (14).

The story of the proof of this formula is quite interesting. An early attempt was made in 1965 during the Humsan school even before a complete proof of the multiplicative ergodic theorem existed. In 1968, about the same time Oseledet's' work appeared in print, Margulis found a proof of the "≤" inequality in (15), *i.e.,* the above estimate for the entropy which thus improved on the Kushnirenko estimate. Although Margulis never published his proof, it was known in the Moscow dynamical community and his authorship was acknowledged, see [118, Section 1.6] for this and other comments about the early history of the entropy formula.

It was also an article of faith, at least within the Moscow dynamics community, that any proof of the below estimate should follow the uniformly hyperbolic model. Namely, integrability of the unstable distribution in a proper sense (allowing for existence only almost everywhere and discontinuities) should be established first; then the crucial property of absolute continuity of conditional measures on those "foliations" proved (again properly adapted). After that, the entropy inequality can be established by constructing an increasing partition whose typical elements are open subsets of the unstable manifolds, and looking into exponential rate of volume contraction in the negative direction of time along those manifolds.

This program was brilliantly realized by Pesin in the mid-seventies and, somewhat similarly to the Ornstein proof of isomorphism for Bernoulli shifts, this work achieved much more than establishing the entropy formula. Thus, it is quite justifiable to speak not only of the "Pesin entropy formula" (15) but also about the "Pesin theory". First, Pesin characterized the π-partition as the measurable hull of the partitions into either global stable or global unstable manifolds, thus showing in particular that those measurable hulls coincide. The most important part of Pesin's work deals with the case when exponents do not vanish (such measure are often called hyperbolic measures). He proves that the system has at most countably many ergodic components with respect to an absolutely continuous hyperbolic measure and that on those components the system is Bernoulli up to a finite permutation. In particular, weak mixing, mixing, K, and Bernoulli properties are all equivalent.[21] This established the paradigm of *nonuniform hyperbolicity* which is in principle capable of reconciling coexistence of large but nowhere dense sets of invariant tori established by Kolmogorov with "chaotic" or "stochastic" behavior on sets of positive measure. It also gives an accurate picture of what this latter type of behavior should look like. This picture is amply supported by a vast body of sophisticated numerical experimentation. A nonuniformly hyperbolic behavior has been established in a variety of situations, some artificial, some model-like, some quite natural and even robust. Still, establishing rigorously the coexistence of a nonuniformly hyperbolic behavior on a positive measure set with KAM behavior remains a completely intractable problem.

Yasha Pesin, 1986

The story of the entropy formula has an interesting twist. Approximately simultaneously with Pesin, a more senior Moscow mathematician, V. M. Millionshchikov, claimed a proof of the entropy formula and published several notes to this end. Those papers contain lots of gaps and inconsistencies and their author refused to present his work at the seminar run at the time by D.V. Anosov and the author. His claim has not been recognized to this day. Since Millionshchikov's approach did not include construction of stable manifolds—not to speak of absolute continuity—no serious attempt was made in Moscow at the time to try to see whether there were valuable ideas in his papers. Several years later, Ricardo

[21] The classification for flows is similar; instead of a finite permutation, one must allow for a constant time suspension; but the last statement holds.

Mañé found another correct proof of the Pesin entropy formula [89] which did not use construction of invariant manifolds or absolute continuity. In private conversations, Mañé asserted that essential ideas of his proof came from his attempts to understand Millionshchikov's arguments and to follow his approach[22].

c. **Entropy estimates for arbitrary measures** [127, 67]**.** A version of the original Margulis argument for the above entropy estimate was applied by David Ruelle to an arbitrary invariant measure for a C^1 diffeomorphism[23]. Thus, nonuniform partial hyperbolicity (existence of some positive and some negative Lyapunov exponents) is the only source of positive entropy and hence by the variational principle also of positive topological entropy in smooth systems.

In low dimensions (two for maps and three for flows), partial hyperbolicity becomes full since, due to the symmetry of entropy with respect to time inversion, one needs at least one positive and one negative exponent. Thus, any diffeomorphism of a surface or a flow on a three-dimensional manifold with positive topological entropy must have an invariant measure (which, without loss of generality, may be assumed ergodic) with nonvanishing Lyapunov exponents.

d. **Entropy, periodic points and horseshoes** [55]**.** Unlike absolutely continuous measures, arbitrary measures considered by the Pesin theory *i.e.,* measures with nonvanishing Lyapunov exponents, may of course have complicated or even pathological intrinsic properties. However, presence of such a measure creates rich *topological* orbit structure. In particular, the exponential growth rate for the number of periodic orbits is estimated from below by the entropy with respect to any invariant measure with nonvanishing Lyapunov exponents and hence by the argument form Section 4c in the low-dimensional cases (two for maps and three for flows) by the topological entropy of the system.

Furthermore, any such measure μ is accompanied by a bevy of invariant *hyperbolic sets* with rich orbit structure. First, the entropy of the restriction of the system to such a set Λ can be made arbitrarily close to the entropy with respect to μ. Second, in the discrete time case, the restriction of a certain power of the system to Λ is topologically conjugate to a transitive topological Markov chain, and in the continuous time case, to a special flow over a transitive topological Markov chain. [58, 61].

Again, in the low-dimensional cases, entropy with respect to a measure with nonvanishing exponents can be replaced by the topological entropy with a striking conclusion: this single magic number is responsible for producing very specific complex but well-organized behavior in an arbitrary smooth system.

A remarkable corollary is lower semicontinuity of topological entropy in these low-dimensional situations. This holds already in any $C^{1+\epsilon}$ topology. As we shall see soon, in the C^∞ topology this improves to actual continuity.

[22] Mañé never claimed that Millionshchikov had a proof of the entropy formula.

[23] Notice that below entropy estimate essentially relies on the Hölder-continuity of the derivatives; C^1 situation is quite pathological in that respect.

e. **Entropy, dimension and converse to the Pesin formula** [83, 84]**.** Looking at the general Ruelle entropy inequality [127] and comparing it with the Pesin entropy formula for absolutely continuous invariant measures, one may naturally ask what is responsible for the "entropy defect" *i.e.,* the difference between the sum of positive exponents and the entropy. Since entropy in a smooth system appears only from hyperbolic behavior, it seems natural to look for the answer to the structure of conditional measures on stable and unstable manifolds. Indeed, in one extreme case, where there is no entropy at all, those conditional measures are atomic[24] and at the opposite extreme, with no defect, the conditional measures are absolutely continuous; in particular, their supports have full dimension.

The first half of the 1980s was the time when this theme was fully developed. An early observation that some sort of dimension-like characteristics of the conditional measures (in other words, their "thickness") should be responsible for entropy defect appears in the work of Lai-Sang Young [146]. Next, François Ledrappier proved the crucial inverse to the Pesin entropy formula. Namely, he characterized measures with nonvanishing exponents for which entropy is equal to the (integral of) the sum of positive exponents (absolutely continuous conditionals on unstable manifolds), or the negative to the sum of the negative exponents (absolutely continuous conditionals on stable manifolds), or both. The last class consists of precisely absolutely continuous hyperbolic measures [82].

The crowning achievement in this line of development was the definitive Ledrappier–Young formula which expressed the entropy of a diffeomorphism with respect to an arbitrary ergodic invariant measure as a weighted sum of positive Lyapunov exponents with coefficients ranging between zero and one which represent rather complicated, but purely geometric (nondynamical) dimension-like quantities. This result has become the foundation for a more detailed study of structure of invariant measures for smooth systems; in particular, it is used in a crucial way in the work on measure rigidity for actions of higher-rank abelian groups discussed in Section 6.

A note on terminology is in order. The original work of Pesin splits into two parts: the theory of stable and unstable manifolds apply to arbitrary Borel probability invariant measures, but the entropy formula and structural theory only to absolutely continuous ones. Structural conclusions in the general case first appeared in [55] in the form of a closing lemma, existence of horseshoes and such, *i.e.,* certain structures external to an original hyperbolic measure. The study of intrinsic structure of arbitrary hyperbolic measures originated in [146]. Nevertheless, nowadays the term "Pesin theory" is often—and quite justifiably—used as a synonym to the general study of nonuniformly hyperbolic behavior.

[24]This simple observation is also a starting point in the entropy rigidity theory for actions of higher-rank abelian groups discussed below in Section 6.

f. **Entropy and volume growth** [144, 98, 99]**.** While entropy with respect to an absolutely continuous invariant measure is measured by the maximal exponential rate of asymptotic infinitesimal volume growth, as in (14) and (15), topological entropy is related to the maximal exponential rate of global volume growth of embedded submanifolds. One expects that the optimal choice should correspond to unstable manifolds for the maximal entropy measure if such a measure exists. In fact, the below estimate (there is at least as much volume growth as there is entropy) follows from a proper application of Pesin theory in the broad sense. This was accomplished by Sheldon Newhouse. The crucial breakthrough was Yosif Yomdin's work which used methods and insights from real algebraic geometry and interpolation theory [145], where he showed that volume growth indeed generates entropy but only for C^∞ systems. For a map f of finite regularity C^r, as was pointed out by Margulis previously to the work of Yomdin and Newhouse, another "entropy defect" may appear which is proportional to $\frac{\|Df\|_0}{r}$. Newhouse derived two important conclusions from Yomdin's estimates:

(i) upper semicontinuity of the topological entropy in C^∞ topology (and hence continuity for low-dimensional cases), and

(ii) existence of maximal entropy measures for an arbitrary C^∞ diffeomorphisms of a compact surface and an arbitrary C^∞ flow on a compact three-dimensional manifold.

One motivation for the work of Yomdin was an "entropy conjecture" formulated by Michael Shub more than a decade earlier [129]. It makes a very plausible assertion that the topological entropy of a diffeomorphism is estimated from below by an even more global complexity invariant: the logarithm of the spectral radius of the mapping induced on the sum of real homology groups[25]. Since the homology growth can be estimated from above by the maximal volume growth of smooth simplexes, Yomdin's result implies Shub's conjecture for C^∞ diffeomorphisms. On the other hand, existence of the entropy defect shows that this method fails in finite regularity. The strongest and most natural C^1 form of the Shub entropy conjecture is still open (and remains one of the remaining entropy mysteries), the only significant advance in its direction still being the 1977 work by Michal Misiurewicz and Feliks Przytycki [94][26].

g. **Growth of closed geodesics on manifolds of nonpositive curvature** [72, 73, 74, 40]**.** There are certain classes of nonuniformly hyperbolic systems where the source of nonuniformity (or rather nonhyperbolicity) is sufficiently well localized and many features of uniformly hyperbolic behavior are present. The case

[25]This was shown to be false in general for homeomorphisms and is or course meaningless for flows since their elements are homotopic to identity.

[26]For certain manifolds, the assertion of the Shub entropy conjecture holds even for continuous maps, as was shown for example for tori by MIsiurewicz and Przytycki in another paper at about the same time. This is, however, more a statement about topology of the manifolds in question that about dynamics.

which received most attention is also of considerable interest in Riemannian geometry: geodesic flows on manifolds of nonpositive (but not strictly negative) sectional curvature.

There is an important dichotomy for compact manifolds of nonpositive curvature. If there is enough zero curvature around, the manifold splits as a Riemannian product (up to a finite cover), or is completely rigid, *i.e.,* is a symmetric space. This remarkable development inspired some of the work discussed in Section 6, but it is not directly related to entropy, so we will not discuss it here.

An alternative case (not too much zero curvature; geometric rank one in technical language)[27] was studied by Pesin [119] before the alternative was proved. Ergodic theory of geodesic flows on Riemannian manifolds of geometric rank one was an important test case for Pesin theory.

The main achievement in the quoted work of Gerhard Knieper was a construction of a maximal entropy measure using convexity properties of geodesics on manifolds of nonpositive curvature and showing for the all-important geometric rank one case that his measure is nonuniformly hyperbolic. Furthermore, Knieper showed that the restriction of the flow to the "nonhyperbolic" part of the phase space has topological entropy strictly smaller than that of the whole flow. After that, uniqueness of measure with maximal entropy follows similarly to the case of negative curvature (*i.e.,* uniformly hyperbolic).

While Knieper did not quite get the multiplicative asymptotic (11) for the growth of the number of closed geodesics[28], this was completed by R. Gunesch in his Ph.D. thesis (still not published); Gunesch also used an important result by M. Babillot who showed that the Knieper measure is mixing.

Now we switch gears and in two remaining sections discuss how entropy appears in the very popular and fruitful field called "rigidity". This is a common name for a diverse variety of phenomena appearing across dynamics, geometry, group theory, and other areas of mathematics. We will describe two different aspects: one when entropy appears as a critical parameter in certain infinite-dimensional variational problems, and the other, where it guarantees sufficient complexity of behavior which leads to rigidity.

5. ENTROPY AS A SOURCE OF RIGIDITY; 1982–1995

In classical systems, coincidence of the topological and volume entropies may lead to infinite-codimension phenomena, such as presence of an invariant rigid geometric structure; this opens a rich area of interplay between dynamics, geometry, and algebra.

---

[27] "Not too much zero curvature" should be understood in dynamical sense: along some geodesics in any two-dimensional direction, zero curvature does not persist. In fact, there are examples of geometric rank-one metrics where sectional curvature vanishes at some two-dimensional direction at any point.

[28] Some of the nonhyperbolic closed geodesics may appear in continuous families where all geodesics have the same length; naturally, in the counting each such family counts as a single geodesic.

a. **Comparing global and statistical complexity.** The variational principle for entropy provides a general framework in which one can ask how far a particular invariant measure is from capturing the maximal dynamical complexity of the system. This is evidently an interesting question if the measure in question has certain special significance; an absolutely continuous invariant measure, if there is one, is obviously a candidate for such consideration; for dissipative systems, SRB measures may play a similar role. Furthermore, there are classes of systems where an invariant volume (an invariant measure given by smooth positive density in local coordinate systems) exists by structural reasons. Two principal examples are restrictions of Hamiltonian systems of classical mechanics to non-critical constant energy surfaces, and geodesic flows on compact Riemannian manifolds[29].

The general question for such systems is when the Kolmogorov entropy with respect to the invariant volume coincides with the topological entropy, or, equivalently, when the volume is a (and usually "the") maximal entropy measure? Let us call this the entropy rigidity problem. The situation is quite different for the discrete and continuous time systems. Since these questions make sense primarily for systems with uniform hyperbolic behavior, *i.e.,* Anosov diffeomorphisms and Anosov flows, which are structurally stable, topological entropy in the discrete time case is locally constant in appropriate spaces of dynamical systems. For the continuous time case, structural stability does not guarantee conjugacy of flows but only existence of a homeomorphism between the orbit foliations. To explain the specifics, we consider two simplest low-dimensional situations where the entropy rigidity problem has been solved early on.

b. **Entropy rigidity on the two-dimensional torus.** Every Anosov diffeomorphism f of a torus in any dimension has hyperbolic "linear part"[30] and is topologically conjugate to the hyperbolic automorphism F_A to which it is homotopic. Notice that for an automorphism, the metric and topological entropies coincide (see Section 3d). Since both numbers are invariant under smooth conjugacy, the entropy rigidity problem for Anosov automorphisms of a torus can be specified as follows:

Under what conditions

$$h_{vol}(f) = h_{top}(f) \tag{16}$$

implies that f is smoothly conjugate to an automorphism?

For volume-preserving Anosov diffeomorphisms of the k-dimensional torus, the answer is: "always" for $k = 2$, and "not always" for $k \geq 3$.

Counterexamples in higher dimension are not particularly interesting; the easiest way to construct such an example is to consider a perturbation of a hyperbolic automorphism which preserves either stable or unstable foliation (one

[29] The latter class can be included into the former by passing to the cotangent bundle via the Legendre transform.

[30] I.e., the induced action on the first homology group has no eigenvalues of absolute value one.

needs to pick one which has dimension ≥ 2), preserves volume and changes eigenvalues at some periodic point.

Proof for $k = 2$ starts from the observation that (16) implies that the conjugacy is volume-preserving (this is true in any dimension). In particular, it takes conditional measures on stable and unstable manifolds for the automorphism into corresponding conditional measures for f. Both families of conditional measures are defined up to a scalar and are given by smooth densities which depend continuously on the leaves. Hence the conjugacy is smooth along stable and unstable foliations and is C^1 globally. Sobolev inequalities then imply that the conjugacy has higher regularity, *e.g.,* C^∞ if f is C^∞.

In higher dimension, interesting problems still appear if one asks about entropy rigidity for more narrow classes of diffeomorphisms than volume-preserving ones, for example symplectic.

c. **Conformal estimates and entropy rigidity for surfaces** [57]**.** A totally different and much more interesting picture appears for a special class of Anosov flows on three-dimensional manifolds, namely, geodesic flows on surfaces of negative curvature.

Let us consider a Riemannian metric σ on a compact surface M of genus $g \geq 2$; $E = 2-2g$ its Euler characteristic, and v the total area. By the Koebe regularization theorem, there is a unique metric σ_0 of constant negative curvature with the same total area conformally equivalent to σ, *i.e.,* $\sigma = \varphi\sigma_0$ where φ is a scalar function. Since the areas are equal one has $\int_M \varphi dvol = 1$ and hence $\rho := \int_M (\varphi)^{1/2} dvol \geq 1$ with the strict inequality for any nonconstant φ. Then,

$$h_{top} \geq \rho\left(\frac{-2\pi E}{v}\right)^{1/2}. \tag{17}$$

Let λ be the Liouville measure in the unit tangent bundle SM; λ is invariant under the geodesic flow. If σ is a metric without focal points (*e.g.,* of negative or nonpositive curvature) then also,

$$h_\lambda \leq \rho^{-1}\left(\frac{-2\pi E}{v}\right)^{1/2}. \tag{18}$$

Inequality (17) immediately implies that metrics of constant negative curvature $\frac{2g-2}{v}$ strictly minimize the topological entropy of the geodesic flow among all (not only negatively curved) metric on a surface of genus $g \geq 2$ with the given total area v.

Similarly, (18) implies that metrics of constant negative curvature $\frac{2g-2}{v}$ strictly maximize the entropy of the geodesic flow with respect to the Liouville measure among all metric on a surface of genus $g \geq 2$ with the given total area v and with no focal points (in particular, among all metrics of nonpositive curvature).

In particular, this solves the entropy rigidity problem for metrics with no focal points on surfaces of genus greater than one:

$$h_{top} = h_\lambda \Leftrightarrow \text{ constant curvature.} \tag{19}$$

d. **Entropy rigidity for locally symmetric spaces** [10, 11]**.** The entropy rigidity for surfaces (19) naturally suggests a higher-dimensional generalization. The model Riemannian manifolds of negative curvature where two entropies coincide are locally symmetric spaces of noncompact type and real rank one. The converse has been known for about twenty five years as "Katok entropy conjecture":[31]

For a negatively curved metric on a compact manifold

$$h_{top} = h_{\lambda} \Leftrightarrow \text{locally symmetric space}.$$

Arguments for conformally equivalent metrics extend to arbitrary dimension, but the main problem is that conformal equivalence to symmetric metric is a specific two-dimensional phenomenon. Even if a symmetric metric exists on M, it is not true any more that any other metric is conformally equivalent to a symmetric one.

More than ten years after the conjecture had been formulated, a major advance was made by Gérard Besson, Gilles Courtois, and Sylvestre Gallot. They proved that on a compact manifold which carries a locally symmetric metric any such metric strictly minimizes h_{top} for the geodesic flow among metrics of fixed volume. Their proof heavily relies on extremal properties of harmonic maps.

At about the same time, Livio Flaminio [29] showed that already in the next lowest-dimensional case after surfaces, there is no analog of (18). Namely, there are arbitrarily small perturbations of a constant negative curvature metric on certain compact three-dimensional manifolds which have variable curvature, preserve the total volume and *increase* the metric entropy with respect to the Liouville measure. In those examples, however, the topological entropy increases even more than the metric one. Flaminio's proof involves calculating the leading quadratic term in the expression for the metric entropy in a small neighborhood of a metric of constant curvature and using representation theory for $SO(3,1)$ to demonstrate that this quadratic form is not negative definite. Flaminio shows that nevertheless locally in a neighborhood of a constant negative curvature the metric Liouville entropy remains smaller than the topological entropy.

6. RIGIDITY OF POSITIVE ENTROPY MEASURES FOR HIGHER-RANK ABELIAN ACTIONS; 1990–2006

For smooth commuting maps and flows with hyperbolic behavior, a panoply of rigidity properties appears in contrast with the classical dynamics; entropy is a dominant actor again. Individual maps are Bernoulli and hence have huge centralizers by the Ornstein theory, but the measurable structure of the actions is completely rigid: the hard whole is built from soft elements.

[31] Not called that way by its author of course, but the name has been added in particular to distinguish it from the older "Shub entropy conjecture", see Section 4f.

a. **Contrast between rank one and higher rank.** While the notion of entropy and principal blocks of the abstract ergodic theory including the Ornstein isomorphism theory and, somewhat surprisingly, even the Kakutani equivalence theory (which from the first glance depends on the notion of order in $\mathbb{Z}$ or $\mathbb{R}$) extend from maps and flows to actions of general abelian groups such as $\mathbb{Z}^k$ and $\mathbb{R}^k$, the relationships between ergodic theory and finite-dimensional dynamics for actions of such groups changes dramatically. In the classical cases, the most natural abstract models such as Bernoulli maps and corresponding flows provide also models for ergodic behavior of classical systems. For essentially all other finitely generated discrete and connected locally compact groups, this is not the case anymore.

The explanation for this contrast is quite simple. We restrict ourselves to the case of discrete groups to avoid some purely technical complications. Entropy measures the growth of number of essential distinct names with respect to a partition against the size of a growing collection of finite sets which exhausts the group. For the group $\mathbb{Z}$, the collection in question is comprised of segments $[-n, n]$. For other groups, say for $\mathbb{Z}^k$, one needs to use collections of sets, such as parallelepipeds, with the number of elements growing faster than its diameter. However for a smooth system, the total number of essential names will still grow only exponentially with respect to any natural norm on the group. For example, if one consider the collection of squares in $\mathbb{Z}^2$ with the side n in order to have positive entropy the number of essential names should grow no slower that $\exp cn^2$ for some constant $c > 0$, while for a smooth action it can grow no faster than $\exp cn$. This disparity forces the entropy of any smooth action to vanish. See [66] for a more detailed explanation.

However, there are smooth actions of higher-rank abelian groups or semigroups where individual elements have positive entropy and hence exhibit a certain degree of hyperbolicity. We first list representative examples and mention results about rigidity of measures for which some elements have positive entropy, and then describe some essential elements of the structure and certain features used in the proofs.

b. **Basic examples.**

Furstenberg's $\times 2, \times 3$ *action* [32]. Consider the action of the semigroup $\mathbb{Z}_+^2$ on the circle generated by two expanding endomorphisms:

$$E_2 \colon S^1 \to S^1 \;\; x \mapsto 2x, \quad (\text{mod } 1),$$

$$E_3 \colon S^1 \to S^1 \;\; x \mapsto 3x, \quad (\text{mod } 1).$$

Aside from Lebesgue measure, there are infinitely many atomic ergodic invariant measures supported by orbits of rational numbers whose denominators are relatively prime to 2 and 3. Furstenberg asked whether those are the only ergodic invariant measures for the action generated by E_2 and E_3. Notice that for an action with a singles generator, there is an enormous variety of invariant measures, whereas if one consider action by multiplications by all natural numbers or even by all squares or all cubes, the only invariant measures are Lebesgue and the δ-measure at 0.

Furstenberg's question stimulated a lot of developments and it is not answered to this day. Moreover, we consider it as one of the most difficult or "resistant" open problems in dynamics, since essentially new methods seem to be needed for its solution. It is even unclear which answer to expect.

A significant advance was made by Dan Rudolph more than twenty years after the question had been posed [126]. Rudolph proved that if at least one (and hence any non-identity) element of the action has positive entropy with respect to an ergodic invariant measure, then the measure is Lebesgue[32]. In this way, entropy made a reappearance in the context of higher-rank smooth actions.

Commuting toral automorphisms [64, 48]. An invertible example which is in essence very similar to the previous one is the action generated by two commuting hyperbolic automorphisms of the three-dimensional torus.

Let $A, B \in SL(3,\mathbb{Z})$, $AB = BA$, $A^k = B^l \Rightarrow k = l = 0$, A, B hyperbolic. The $\mathbb{Z}^2$ action is generated by automorphisms of the torus $\mathbb{T}^3 = \mathbb{R}^3/\mathbb{Z}^3$:

$$F_A\colon x \mapsto Ax, \quad (\text{mod } 1),$$

$$F_B\colon x \mapsto Bx, \quad (\text{mod } 1).$$

Shortly after Rudolph's paper appeared, I suggested to consider invariant measures for this action in an attempt to find a more geometric proof of Rudolph's theorem. This was successful and resulted in our work with Ralf Spatzier [64][33] in which not only rigidity of positive entropy measure for a broad class of actions of $\mathbb{Z}^k$, $k \geq 2$ by automorphisms of a torus is proved, but also many noninvertible cases (including naturally actions by expanding endomorphisms of the circle) are covered.

In order to see a similarity between the action on the three-dimensional torus and the multiplications by 2 and by 3, we pass to the natural extension for the latter action. The natural extension is the invertible action on the space of all "pasts" for the original noninvertible action. The phase space for this natural extension of the $\times 2, \times 3$ action is a solenoid, the dual group to the discrete group $\mathbb{Z}(1/2, 1/3)$. It is locally modeled on the product of $\mathbb{R}$ with the groups of 2-adic and 3-adic numbers. Thus, while topologically the solenoid is one-dimensional, there are three "directions" there: one real and two non-Archimedean. They play the role very similar to that of three common eigendirections for A and B in the toral case.

Weyl chamber flow (WCF)[25, 27]. Let $M = SL(n,\mathbb{R})/\Gamma$, $n \geq 3, \Gamma$ a lattice in $SL(n,\mathbb{R})$, D the group of positive diagonal matrices which is isomorphic to $\mathbb{R}^{n-1}$. The action of D on M by left translations is called the *Weyl chamber flow*. This

[32] A couple of years earlier, Russel Lyons [88] proved a somewhat weaker result using a method based on harmonic analysis.

[33] Which appeared only in 1996 but was submitted already in 1992.

example (assume for simplicity $n = 3$) stands in the same relation to the previous one as the geodesic flow on a surface of constant negative curvature to a hyperbolic automorphism of a torus. This is an Anosov or normally hyperbolic action, *i.e.,* there are elements which are partially hyperbolic diffeomorphisms whose neutral foliation is the orbit foliation. For Weyl chamber flows, the notion of rigidity of positive entropy measure has to be somewhat modified. For certain lattices Γ there are closed subgroups of $SL(n,\mathbb{R})$ which project to the factor space as closed sets. On some of those invariant sets, the Weyl chamber flow may become essentially a rank one action. Those examples were found by Mary Rees in the early 1980s. So in the description of positive entropy invariant measures, one needs to allow first for Haar measures on some closed subgroups whose projections to the homogeneous space are closed, and second, for invariant measures supported on the projections of the subgroups in Rees-type examples.

c. **Lyapunov exponents and Weyl chambers for actions of $\mathbb{Z}_+^k, \mathbb{Z}^k$, and $\mathbb{R}^k$ [48].** Let α be such an action, μ an α-invariant ergodic measure; by passing to the natural extension and/or suspension can always reduce the situation to the case of $\mathbb{R}^k$, which has certain advantages in visualization of "time"—although it adds extra directions in space.

Lyapunov exponents for an $\mathbb{R}^k$ action with respect to an ergodic invariant measure are linear functionals $\chi_i \in (\mathbb{R}^k)^*$.

Lyapunov hyperplanes are the kernels of nonzero Lyapunov exponents .

Weyl chambers are connected components of $\mathbb{R}^k \setminus \bigcup_i \operatorname{Ker}\chi_i$. Elements of the action which do not belong to any of the Lyapunov hyperplanes are called *regular*. Thus, Weyl chambers are connected components of the set of all regular elements.

The measure is called hyperbolic is all Lyapunov exponents other than k coming from the orbit directions are nonzero.

For the algebraic examples discussed in Section 6b, Lyapunov exponents are obviously independent of the measure. For $k = 2$, Lyapunov hyperplanes become lines and Weyl chambers sectors in the plane. The iconic picture which appears for both examples in Section 6b.2 and in Section 6b.3 is of three different Lyapunov lines dividing the plane into six Weyl chambers. However, in the former case, there are only three Lyapunov exponents χ_1, χ_2, χ_3, and the only relation among them is $\chi_1 + \chi_2 + \chi_3 = 0$, corresponding to the preservation of volume. In the latter case, there are three pairs of exponents, each pair consisting of two exponents of opposite sign, and thus corresponding to the same Lyapunov line.

In order to incorporate the case of $\times 2$, $\times 3$ action to this scheme, one considers the non-Archimedean Lyapunov exponents for the natural extension in addition to the single real exponent. Since the multiplication by 2 is an isometry in the 3-adic norm and vice versa, the Lyapunov lines in this case are the two axis and the line $x \log 2 + y \log 3 = 0$, which does not intersect the first quadrant, the only one "visible" if one considers the original noninvertible action[34].

Weyl chambers can be characterized as the loci of element of the action for which the stable and unstable distributions and their integral foliations are the same[35]. Maximal, nontrivial intersections of stable distributions and foliations correspond to half-spaces on each side of a Lyapunov hyperplane; they are called *coarse Lyapunov* distributions (foliations); in all our examples, those are one-dimensional, and coarse Lyapunov distributions coincide with the *Lyapunov distributions* which appear in the Multiplicative Ergodic theorem for higher-rank actions. While Lyapunov distributions are not always uniquely integrable, the coarse Lyapunov distributions are. On the torus, of course, all Lyapunov distributions are integrable and they commute, but for the Weyl chamber flows, there are nontrivial commutation relations between the pairs of distributions corresponding to the exponents of the opposite sign. This noncommutativity turns out to be a saving grace, both in the measure rigidity and in the differentiable rigidity theory for Weyl chamber flows.

The main source of rigidity is the combination of global recurrence and local isometry for action along the singular directions, *i.e.,* for elements of the action in the Lyapunov hyperplanes. This is true for measure rigidity, which we discuss now, as well as differentiable rigidity, which we barely mention, since it is not explicitly related to entropy. Global recurrence may mean, depending on the context, topological transitivity or some sort of ergodicity. The other property means that within the leaves of the corresponding Lyapunov foliation, the action is isometric, as in our algebraic examples, or nearly so, as for small perturbations in the local differentiable rigidity theory or for actions with Cartan homotopy data in [50, 62]. Then a geometric structure along the leaves of a Lyapunov foliation is carried over to the whole space or its substantial part due to recurrence. In the differentiable rigidity theory, the geometric structures in question are flat affine structures or, more generally, the structures with finite-dimensional automorphisms group coming from the "nonstationary normal forms theory" [41]. In the measure rigidity, those are conditional measure induced on the leaves of the foliations and it is here where entropy makes its entrance.

d. **Entropy function.** Entropy function for an $\mathbb{R}^k$ action α is defined on $\mathbb{R}^k$ by

$$h(\mathbf{t}) =: h_\mu(\mathbf{t}) =: h_\mu(\alpha(\mathbf{t})).$$

It has the following properties:

1. $h(\lambda \mathbf{t}) = |\lambda| \cdot h(\mathbf{t})$ for any scalar λ;
2. $h(\mathbf{t}+\mathbf{s}) \le h(\mathbf{t}) + h(\mathbf{s})$;
3. h is linear in each Weyl chamber.

[34]See [64] for a general treatment of noninvertible actions on tori and actions on solenoids along these lines.

[35]In the general case, one should understand the terms "distributions" and "foliations" in the sense of Pesin theory, *i.e.,* as families of subspaces and smooth submanifolds defined almost everywhere; in our algebraic examples, those objects are, of course, smooth and homogeneous

The Ledrappier–Young entropy formula [84], discussed in Section 4e, is used in an essential way in the proofs of (2) and (3).

e. **Conditional measures.** Positivity of metric entropy for an ergodic transformation is equivalent to continuity (nonatomicity) of conditional measures on stable and unstable manifolds. The basic argument, which is sufficient to prove entropy rigidity for both the action by automorphisms of $\mathbb{T}^3$ and the natural extension of the $\times 2$, $\times 3$ action, is showing that those conditional measures must be invariant under almost every translation and hence Lebesgue. The argument uses as a basic tool the well-known Luzin theorem from real analysis and also a "π-partition trick", which is based on the fact that π-partitions can be calculated either from the infinite past or from the infinite future.

For Weyl chamber flows, this argument is not sufficient primarily because not every pair of Lyapunov distributions belong to the stable distribution of an element. There are two more methods which are known as the "high entropy" [25] and the "low entropy" [27][36] methods; they both use special character of commutation relations between various Lyapunov distributions.

f. **Isomorphism rigidity** [60]**.** One of the striking consequences of measure rigidity justifying the expression "the hard whole built from soft elements" is a very special character of measurable isomorphisms between actions. For example, any isomorphism between two actions by hyperbolic automorphisms of $\mathbb{T}^3$ (and for much more general "genuinely higher-rank" actions by hyperbolic automorphisms of a torus) is algebraic *i.e.,* coincides almost everywhere with an affine map. A remarkable corollary is that the entropy function is not a full isomorphism invariant here, since equivalence of entropy functions only implies conjugacy over $\mathbb{Q}$ but not over $\mathbb{Z}$. Any two actions on $\mathbb{T}^3$ isomorphic over $\mathbb{Q}$ are finite factors of each other, but if the class number of the corresponding cubic algebraic number field (which is obtained by adding eigenvalues of A and B to the rationals) is greater than one, there are examples not isomorphic over $\mathbb{Z}$ and hence not isomorphic metrically. Furthermore, the centralizer of the action in the group of measure-preserving transformation coincides with the algebraic centralizer and is in fact a finite extension of the action itself. Factors and joinings are similarly rigid [49].

All this stands in a dramatic contract with the case of a single automorphism. Being a Bernoulli map, it has a huge centralizer which actually defies description and an enormous variety of factors and self-joinings. Also, automorphisms of tori with the same entropy are isomorphic, but the entropy in dimension greater than two does not even determine the conjugacy class over $\mathbb{Q}$, or, equivalently, over $\mathbb{C}$. And of course, isomorphisms between two automorphisms with the same entropy are incredibly diverse.

g. **Applications** [86, 27]**.** Measure rigidity for higher-rank actions found applications in several areas of mathematics beyond ergodic theory or dynamics. The

[36]This method was developed by E. Lindenstrauss earlier in a different context

first spectacular application was Elon Lindenstrauss's work on arithmetic quantum chaos [85, 86], a breakthrough in the area which for along time resisted efforts of several outstanding mathematicians working by methods from classical analysis and number theory.

Another application is a partial solution of the Littlewood conjecture on multiplicative Diophantine approximations [27]. It follows from the measure rigidity for the Weyl chamber flow on the noncompact but finite volume symmetric space $SL(3,\mathbb{R})/SL(3,\mathbb{Z})$. In this case, one can show that the only invariant ergodic positive entropy measure is Lebesgue. This implies, via arguments related to the variational principle and to the connection between entropy and Hausdorff dimension, that the set of bounded orbits has zero Hausdorff dimension.

Littlewood, around 1930, conjectured the following property of multiplicative Diophantine approximation. For every $u, v \in \mathbb{R}$,

$$\liminf_{n\to\infty} n\langle nu\rangle\langle nv\rangle = 0, \tag{20}$$

where $\langle w\rangle = \min_{n\in\mathbb{Z}} |w-n|$ is the distance of $w \in \mathbb{R}$ to the nearest integer.

The corollary of measure rigidity described above implies the following partial result toward Littlewood's conjecture:

Let

$$\Xi = \left\{(u,v)\in\mathbb{R}^2 : \liminf_{n\to\infty} n\langle nu\rangle\langle nv\rangle > 0\right\}.$$

Then the Hausdorff dimension $\dim_H \Xi = 0$. In fact, Ξ is a countable union of compact sets with box dimension zero.

h. **Measure rigidity beyond uniform hyperbolicity [50, 62, 51].** So far, we discussed measure rigidity for algebraic actions. It might look as a natural next step to consider uniformly hyperbolic, nonalgebraic actions. However, another aspect of higher-rank rigidity, the differentiable rigidity, indicates that those objects should be essentially algebraic, so not much would be gained. On the other hand, some of the key features of the algebraic situation in a properly modified way appear in a very general context where indeed nontrivial behavior is possible. So far, the case of actions of the highest possible rank has been considered: $\mathbb{Z}^k$ actions of $k+1$-dimensional manifolds for $k \geq 2$. The actions on $\mathbb{T}^3$ from Section 6b are the simplest, nontrivial models; they and their higher-dimensional counterparts are called *linear Cartan actions*.

Now we consider a $\mathbb{Z}^k$ action on $\mathbb{T}^{k+1}$ such that the induced action on the first homology coincides with a linear Cartan action; we will say that such an action has *Cartan homotopy data*. It turns out that such an action α has a uniquely defined invariant measure μ, called *large*, which is absolutely continuous such that the action α is metrically isomorphic to the corresponding linear Cartan action α_0 with Lebesgue measure. Furthermore, the restriction of this isomorphism to a compact (noninvariant in general) set of arbitrarily large measure μ (and hence of positive Lebesgue measure) is smooth in the sense of Whitney. This is a rather remarkable case of existence of an invariant geometric structure determined by a purely homotopy information.

Furthermore, if one considers $\mathbb{Z}^k$ actions on an arbitrary $(m+1)$-dimensional compact manifold with an invariant measure for which the Lyapunov hyperplanes are in general position[37], then assuming that at least one element of the action has positive entropy, we deduce that the measure μ is absolutely continuous.

7. EPILOGUE: SOME (MOSTLY OLD) OPEN PROBLEMS

Questions related to entropy inspired and continue to inspire new advances in ergodic theory, topological and differentiable dynamics, and related areas of mathematics. Open problems listed here, one for each section, mostly old and well-known, some mentioned in passing in the text, still remain major challenges. A serious advance toward the solution of any of these problems would require fresh new insights and will almost surely stimulate more progress.

a. **Entropy for the standard map.** The following family of area-preserving maps f_λ of the cylinder $C = S^1 \times \mathbb{R}$:

$$f_\lambda(x, y) = (x + y, y + \lambda \sin 2\pi(x + y)).$$

is often called standard.

PROBLEM 7.1. *Is the metric entropy $h_{area}(f_\lambda)$ positive*
(i) for small λ, or
(ii) for any λ *(assuming y is periodic too)?*

A positive answer would imply existence of ergodic components of positive measure by Pesin theory.

This seems to be the only major problem related to entropy which originated in the sixties (in fact, this one has been known since an early part of that decade) which is still open. It also looks extremely difficult, especially part (i), due to the existence of invariant curves by the Moser Invariant Curve Theorem [95]. Invariant circles divide the cylinder into annular domains of instability with positive topological entropy and complex orbits of various kinds. But the boundaries of those domains are not smooth KAM-type circles but some degenerate ones, most likely not even C^1, although they are always graphs of Lipschitz functions $S^1 \to \mathbb{R}$. If, as expected, ergodic components come close to these circles, estimates needed to establish nonuniform hyperbolicity become hardly imaginable.

For large values of λ, Moser circles disappear, but those around elliptic points remain. For part (ii) there is a hope, based on parameter exclusion techniques pioneered by Michael Jakobson, that for some parameter values, all elliptic points can be eliminated and estimates carried out almost everywhere.

[37] For $k = 2$, this simply means that Lyapunov lines are different

b. **Weak Pinsker property.** After Pinsker's conjecture was disproved by Ornstein, Thouvenot suggested a similar property which he calls the *weak Pinsker property*:

$(\mathcal{WP})$ A measure-preserving transformation T is the direct product of a Bernoulli shift and a transformation with arbitrarily small entropy.

PROBLEM 7.2. *Does any ergodic measure-preserving transformation T satisfy the weak Pinsker property?*

To the best of our knowledge, this is the only major structural open problem related to the basic properties of entropy and the Ornstein isomorphism theory. Thouvenot thinks that the weak Pinsker property does not always hold and that there is a certain invariant which he calls the "entropy production rate" whose vanishing characterizes systems with $(\mathcal{WP})$.

c. **Shub entropy conjecture for C^1 maps.** Let f be a C^1 diffeomorphism (or, more generally, a C^1 map) of a compact differentiable manifold M to itself;

$$f_*: \sum_{i=1}^{n} H(m,\mathbb{R}) \to \sum_{i=1}^{n} H(m,\mathbb{R})$$

the induced map of the homology with real coefficients.

PROBLEM 7.3. *Is it true that $h_{top}(f) \geq \log s(f_*)$, where s is the spectral radius (the maximal absolute value of an eigenvalue)?*

A positive answer is known as the (C^1) Shub entropy conjecture. Although for C^∞ maps a positive answer follows from the work of Yomdin discussed in Section 4f, Yomdin's volume estimate fails in finite smoothness. My survey [54] from ten years before the work of Yomdin, which appeared in English translation just before Yomdin's work, still accurately reflects the state of the subject beyond C^∞.

d. **Growth of periodic orbits for surface diffeomorphisms.** As was mentioned in Section 3d, we proved in [55] that for a $C^{1+\epsilon}$ diffeomorphism f of a compact surface

$$\limsup_{n\to\infty} \frac{\log P_n(f)}{n} \geq h_{top}(f),$$

i.e., the *exponential* growth rate of the number of periodic orbits is estimated from below by topological entropy. Notice that this remains true if one considers only hyperbolic periodic orbits.

PROBLEM 7.4. *Is it true that*

$$\limsup_{n\to\infty} \frac{P_n(f)}{\exp n h_{top}(f)} \geq c > 0$$

for any (i) $C^{1+\epsilon}$ or (ii) C^∞ diffeomorphism f of a compact surface?

The answer, at least in the C^∞ case, is likely to be positive so we dare to call this a conjecture. On the other hand, Jerome Buzzi thinks that in finite regularity the answer may be negative.

A possible strategy for (i) would be to construct an infinite Markov partition which captures full topological entropy and use results of B.M. Gurevich and others for asymptotics of periodic orbit growth for certain infinite topological Markov chains.

For (ii) there is another approach based on existence of a maximal entropy measure discussed in Section 4f. A hopeful development in the last few years is the theory of symbolic extensions and entropy structure by M. Boyle, T. Downarowicz, Newhouse and others [19, 24], as well as detailed analysis of maximal entropy measures in low dimension by Buzzi. It is based on a much weaker but much more general construction than that of Markov partitions, called principal symbolic extensions.

e. **Entropy rigidity conjecture.** This was already discussed in Section 5d. We just repeat the question.

PROBLEM 7.5. *Suppose that for a Riemannian metric of negative sectional curvature on a compact manifold, topological entropy of the geodesic flow coincides with the entropy with respect to the smooth (Liouville) measure. Does this imply that the metric is locally symmetric?*

As we already mentioned, a presumed positive answer is known as the "Katok entropy conjecture". It stimulated a remarkable cycle of papers by Besson, Courtois and Gallot and lots of other interesting work.

f. **Values of entropy for $\mathbb{Z}^k$ actions.** This is inspired by results discussed in Section 6h. Let $k \geq 2$, α be a $C^{1+\epsilon}$, $\epsilon > 0$ action of $\mathbb{Z}^k$ on a $k+1$-dimensional manifold, μ an ergodic $alpha$-invariant measure with no proportional Lyapunov exponents, and at least one element of α has positive entropy. According to [51], such a measure is absolutely continuous.

PROBLEM 7.6. *What are possible values of entropy for elements of an action α as above?*

We conjecture that the entropy values are algebraic integers of degree at most $k+1$ as in actions with Cartan homotopy data [50] on torus and other known examples on a variety of different manifolds [62]

This is the only recent problem on our list; it was formulated at the 2006 Oberwolfach conference in geometric group theory, hyperbolic dynamics and symplectic geometry. All other problems have been known for more than twenty years.

REFERENCES

[1] *Cycle of papers in ergodic theory*, (Russian) Ushehi Mat. Nauk, **22** (1967), 3–172; English translation: Russian Math. Surveys, **22** (1967), 1–167.

[2] Roy L. Adler, *On a conjecture of Fomin*, Proc. Amer. Math. Society, **13** (1962), 433–436.

[3] Roy L. Adler, Alan G. Konheim and M. Harry McAndrew, *Topological entropy*, Transactions of the American Mathematical Society, **114** (1965), 309–319.
[4] Roy L. Adler and Brian Marcus, *Topological entropy and equivalence of dynamical systems*, Memoires. of AMS, **20** (1979).
[5] Roy L. Adler and Benjamin Weiss, *Entropy, a complete metric invariant for automorphisms of the torus*, Proceedings of the National Academy of Sciences, **57** (1967), 1573–1576.
[6] Roy L. Adler and Benjamin Weiss, *Similarity of automorphisms of the torus*, Memoirs of the American Mathematical Society, American Mathematical Society, Providence, R.I., 1970.
[7] Dmitry V. Anosov, *Geodesic flows on closed Riemannian manifolds of negative curvature* (Russian) Trudy Mat. Inst. Steklov., 1967, 209; English translation: American Mathematical Society, Providence, R.I. 1969, iv+235.
[8] Dmitry V. Anosov and Yakov G. Sinaĭ, *Certain smooth ergodic systems* Russian Math. Surveys 22, 1967 (137), 103–167.
[9] Kenneth Berg, "On the Conjugacy Problem for K-Systems," Ph.D. thesis, University of Minnesota, 1967.
[10] Gérard Besson, Gilles Courtois and Sylvestre Gallot, *Entropies et rigidités des espaces localement symétriques de courbure strictement négative* (French), Geometric and Functional Analysis, **5** (1995), 731–799.
[11] Gérard Besson, Gilles Courtois and Sylvestre Gallot, *Minimal entropy and Mostow's rigidity theorems*, Ergodic Theory and Dynamical Systems, **16** (1996), 623–649.
[12] F. Blanchard, *Fully positive topological entropy and topological mixing* Contemp. Math |bf 135 Amer. Math. Soc. Providence, (1992), 95–105.
[13] J. R. Blum and D. L. Hanson, *On the isomorphism problem for Bernoulli schemes*, Bulletin of the American Mathematical Society, **69** (1963), 221–223.
[14] Rufus Bowen, *Topological entropy and axiom A*, in: Global Analysis, Proceedings of Symposia in Pure Mathematics, **14**, American Mathematical Society, Providence, RI 1970, 289–297.
[15] Rufus Bowen, *Markov partitions for Axiom A diffeomorphisms*, American Journal of Mathematics, **92** (1970), 725–747.
[16] Rufus Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Mathematics **470**, Springer-Verlag, Berlin, 1975.
[17] Rufus Bowen, *Periodic points and measures for Axiom A diffeomorphisms*, Transactions of the American Mathematical Society, **154** (1971), 377–397.
[18] Rufus Bowen and David Ruelle, *The ergodic theory of Axiom A flows*, Inventiones Mathematicae, **29** (1975), 181–202.
[19] Micheal Boyle and Tomasz Downarowicz, *The entropy theory of symbolic extensions* Inventiones Math., **156** (2004), 119–161.
[20] Michael Brin, Boris Hasselblatt and Yakov Pesin, editors, "Modern Theory of Dynamical Systems and Applications," Cambridge University Press, 2004.
[21] I. Csiszár and J. Körner, Information theory, *Coding theorems for discrete memoryless sources*, Akadémiai Kiadó, Budapest, 1981.
[22] Efim I. Dinaburg, *A correlation between topological entropy and metric entropy*, Dokl. Akad. Nauk SSSR, **190** (1970), 19–22.
[23] Efim I. Dinaburg, *A connection between various entropy characterizations of dynamical systems*, Izv. Akad. Nauk SSSR Ser. Mat., **35** (1971), 324–366.
[24] Tomasz Downarowicz and Sheldon Newhouse, *Symbolic extensions and smooth dynamical systems*, Invent. Math., **160** (2005), 453–499.
[25] Manfred Einsiedler and Anatole Katok, *Invariant measures on G/Γ for split simple Lie Groups G*, Comm. Pure. Appl. Math., **56**, (2003), 1184–1221.
[26] Manfred Einsiedler and Anatole Katok, *Rigidity of measures -The high entropy case and non-commuting foliations*, Israel Math. J., **148** (2005), 169–238.

[27] Manfred Einsiedler, Anatole Katok and Elon Lindenstrauss, *Invariant measures and the set of exceptions in Littlewood's conjecture*, Ann. Math., (2), **164** (2006), 513–560.

[28] Jacob Feldman, *New K-automorphisms and a problem of Kakutani*, Israel J. Math., **24** (1976), 16–38.

[29] Livio Flaminio, *Local entropy rigidity for hyperbolic manifolds* Comm. Anal. Geom., **3** (1995), 555–596.

[30] Sergey Vasil'evish Fomin, *Some new problems and results in ergodic theory*, (Russian) Appendix to Russian translation of Lectures on Ergodic theory by P.R. Halmos, Moscow, Mir, 1960, 134–145.

[31] Nathaniel A. Friedman and Donald Ornstein, *On isomorphism of weak Bernoulli transformations*, Advances in Mathematics, **5** (1970), 365–394.

[32] Harry Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Mathematical Systems Theory, **1** (1967), 1–49.

[33] Hans-Otto Georgii, *Probabilistic aspects of entropy*, in Entropy, Andreas Greven, Gerhard Keller, Gerald Warnecke, eds., Princeton University Press, (2003), 37–54.

[34] Igor V. Girsanov, *Spectra of dynamical systems generated by stationary Gaussian processes.* (Russian) Dokl. Akad. Nauk SSSR (N.S.), **119** (1958), 851–853.

[35] Eli Glasner and Benjamin Weiss, *On the interplay between measurable and topological dynamics*, in Handbook in dynamical systems, Elsevier, Amsterdam, **1B** 2006, 597–648.

[36] Tim N. T. Goodman, *Relating topological entropy and measure entropy*, Bulletin of the London Mathematical Society, **3** (1971), 176–180.

[37] Tim N. T. Goodman, *Maximal measures for expansive homeomorphisms*, Bulletin of the London Mathematical Society (2) 5 (1972), 439–444.

[38] L. Wayne Goodwyn, *Topological entropy bounds measure-theoretic entropy*, Proceedings of the American Mathematical Society, **23** (1969), 679–688.

[39] L. Wayne Goodwyn, *Comparing topological entropy with measure-theoretic entropy*, American Journal of Mathematics, **94** (1972), 366–388.

[40] Roland Gunesch, "Precise Asymptotic for Periodic Orbits of the Geodesic Flow in Nonpositive Curvature," Ph. D. Thesis, Pennsylvania State University, 2002. http://etda.libraries.psu.edu/theses/approved/WorldWideFiles/ETD-198/etd.pdf

[41] Misha Guysinsky and Anatole Katok, *Normal forms and invariant geometric structures for dynamical systems with invariant contracting foliations*, Math. Res. Letters, **5** (1998), 149–163.

[42] Paul R. Halmos, "Lectures on Ergodic Theory," Publications of the Mathematical Society of Japan, The Mathematical Society of Japan, 1956.

[43] Paul R. Halmos, *Recent progress in ergodic theory*, Bulletin of the American Mathematical Society, **67** (1961), 70–80.

[44] Boris Hasselblatt and Anatole Katok, *Principal structures*, in "Handbook of Dynamical Systems," North-Holland, Amsterdam, **1A** (2002), 1–203.

[45] Boris Hasselblatt and Anatole Katok *The development of dynamics in the 20th century and contribution of Jürgen Moser*, Erg. Theory and Dynam. Systems, **22** (2002), 1343–1364.

[46] Heinz Huber, *Zur analytischen Theorie hyperbolischer Raumformen und Bewegungsgruppen. II.* (German) Math. Ann., **142** (1961), 385–398.

[47] Steven Kalikov, *The T, T^{-1}-transformation is not loosely Bernoulli.*, Annals of Mathematics, **115** (1982), 393–409.

[48] Boris Kalinin and Anatole Katok, *Invariant measures for actions of higher rank abelian groups*, in "Smooth Ergodic Theory and its Applications" (Seattle, WA, 1999), 593–637, Proceedings of Symposia in Pure Mathematics, **69**, American Mathematical Society, Providence, RI, 2001.

[49] Boris Kalinin and Anatole Katok, *Measurable rigidity and disjointness for $\mathbb{Z}^k$ actions by toral automorphisms*, Ergodic Theory and Dynamical Systems, **22** (2002), 507–523.

[50] Boris Kalinin and Anatole Katok, *Measure rigidity beyond uniform hyperbolicity: Invariant Measures for Cartan actions on Tori*, Journal of Modern Dynamics, **1** (2007), 123–146.
[51] Boris Kalinin, Anatole Katok and Federico Rodriguez Hertz, *Nonuniform measure rigidity*, preprint, 2007.
[52] Anatole. B. Katok, *Time change, monotone equivalent and standard dynamical systems* Soviet Math. Dokl., **16** (1975), 789–792.
[53] Anatole. B. Katok, *Monotone equivalence in ergodic theory*, Izv. Akad. Nauk SSSR, **41** (1977), 104–157; English translation: Math USSR Izvestija, **11** (1977), 99–146.
[54] Anatole B. Katok, *A conjecture about entropy*, in Smooth Dynamical Systems, Moscow, Mir, 1977, 181–203; English translation: AMS Transl (2), **133** (1986), 91–107.
[55] Anatole Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Publications Mathématiques de l'Institut des Hautes Études Scientifiques, **51** (1980), 137–173.
[56] Anatole Katok, *Smooth non-Bernoulli K-automorphisms*, Invent. Math., **61** (1980), 291–299.
[57] Anatole Katok, *Entropy and closed geodesics*, Ergodic Theory and Dynamical Systems, **2** (1982), 339–367.
[58] Anatole Katok, *Non-uniform hyperbolicity and Periodic orbits for Smooth Dynamical Systems*, in Proc. Intern. Congress of Math., **2**, PWN-North Holland, (1984), 1245–1253.
[59] Anatole Katok and Boris Hasselblatt, "Introduction to the modern theory of dynamical systems," Cambridge University Press, Cambridge, 1995.
[60] Anatole Katok, Svetlana Katok and Klaus Schmidt, *Rigidity of measurable structure for $\mathbb{Z}^d$-actions by automorphisms of a torus*, Commentarii Mathematici Helvetici, **77** (2002), 718–745.
[61] Anatole Katok and Leonardo Mendoza, *Dynamical systems with non-uniformly hyperbolic behavior*, Supplement to *Introduction to the modern theory of smooth dynamical systems*, Cambridge university press, (1995), 659–700.
[62] Anatole Katok and Federico Rodriguez Hertz, *Uniqueness of large invariant measures for Z^k actions with Cartan homotopy data*, Journal of Modern Dynamics, **1** (2007), 287–300.
[63] Anatole Katok and Ralf J. Spatzier, *First cohomology of Anosov actions of higher rank abelian groups and applications to rigidity*, Publications Mathématiques de l'Institut des Hautes Études Scientifiques, **79** (1994), 131–156.
[64] Anatole Katok and Ralf J. Spatzier, *Invariant measures for higher-rank hyperbolic abelian actions*, Ergodic Theory and Dynamical Systems, **16** (1996), 751–778.
[65] Anatole Katok and Ralf J. Spatzier, *Differential rigidity of Anosov actions of higher rank abelian groups and algebraic lattice actions*, Tr. Mat. Institute. Steklova **216** (1997), Din. Sist. i Smezhnye Vopr., 292–319; Proceedings of the Steklov Institute of Mathematics 1997, (216), 287–314.
[66] Anatole Katok and Jean-Paul Thouvenot, *Slow entropy type invariants and smooth realization of commuting measure-preserving transformations*, Ann. Inst. H.Poincaré, **33** (1997), 323–338.
[67] Svetlana Katok, *The estimation from above for the topological entropy of a diffeomorphism*, in Global theory of dynamical systems Lecture Notes in Math., **819**, Springer, Berlin, 1980, 258–264.
[68] Yitzhak Katznelson, Ergodic automorphisms of $\mathbb{T}^n$ are Bernoulli shifts, Israel. J. Math., **12** (1971), 161–173.
[69] Michael Keane and Meir Smorodinsky, *Bernoulli schemes of the same entropy are finitarily isomorphic*, Annals of Mathematics, **109** (1979), 397–406.
[70] Michael Keane and Meir Smorodinsky, *Finitarily isomorphism of irreducible Markov shifts*, Israel Journal of Mathematics, **34** (1979), 281–286.
[71] Alexander Ya. Khinchin, *On the basic theorems of information theory*, Uspehi Mat. Nauk (N.S.), **11** (1956), 17–75.
[72] Gerhard Knieper, *On the asymptotic geometry of nonpositively curved manifolds*, Geometric and Functional Analysis, **7** (1997), 755–782.

[73] Gerhard Knieper, *The uniqueness of the measure of maximal entropy for geodesic flows on rank* 1 *manifolds*, Annals of Mathematics (2), **148** (1998), 291–314.
[74] Gerhard Knieper, *Closed geodesics and the uniqueness of the maximal measure for rank 1 geodesic flows*, in: Smooth ergodic theory and its applications (Seattle, WA, 1999), 573–590, Proceedings of Symposia in Pure Mathematics, **69**, American Mathematical Society, Providence, RI, 2001.
[75] Andrei N. Kolmogorov, *On dynamical systems with an integral invariant on the torus*, Doklady Akademii Nauk SSSR (N.S.), **93** (1953), 763–766.
[76] Andrei N. Kolmogorov, *On conservation of conditionally periodic motions for a small change in Hamilton's function*, Doklady Akademii Nauk SSSR (N.S.) **98** (1954), 527–530; English translation in *Stochastic behavior in classical and quantum Hamiltonian systems*, Volta Memorial conference, Como, 1977, Lecture Notes in Physics, **93**, Springer, 1979.
[77] Andrei N. Kolmogorov, *Théorie générale des systèmes dynamiques et mécanique classique* (French), in *Proceedings of the International Congress of Mathematicians, Amsterdam, 1954, Vol. 1*, 315–333, Noordhoff, Groningen; North-Holland, Amsterdam, 1957; English translation: Ralph Abraham, *Foundations of Mechanics*, Benjamin, 1967, Appendix D.
[78] Andrei N. Kolmogorov, *A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces*, Doklady Akademii Nauk SSSR (N.S.), **119** (1958), 861–864.
[79] Andrei N. Kolmogorov, *Entropy per unit time as a metric invariant of automorphisms*, Doklady Akademii Nauk SSSR, 124 (1959), 754–755.
[80] Andrei N. Kolmogorov, *A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces* (Russian) in Topology, ordinary differential equations, dynamical systems. Trudy Mat. Inst. Steklov., **169** (1985), 94–98.
[81] Anatoli G. Kushnirneko, *An upper bound of the entropy of classical dynamical systems*(Russian), Sov. Math. Dokl., **6** (1965), 360–362.
[82] François Ledrappier, *Proprétés ergodiques des mesures de Sinaï.* (French) Inst. Hautes Čtudes Sci. Publ. Math. IHES, **59** (1984), 163–188.
[83] François Ledrappier and Lai-Sang Young, *The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula*, Annals of Mathematics (2), **122** (1985), 509–539.
[84] François Ledrappier and Lai-Sang Young, *The metric entropy of diffeomorphisms. II. Relations between entropy, exponents and dimension*, Annals of Mathematics (2), **122** (1985), 540–574.
[85] Elon Lindenstrauss, *On quantum unique ergodicity for* $\Gamma\backslash\mathbb{H}\times\mathbb{H}$, International Mathematics Research Notices, **2001**, 913–933.
[86] Elon Lindenstrauss, *Invariant measures and arithmetic quantum unique ergodicity*, Annals of Mathematics, **163** (2006), 165–219.
[87] Rafael de la Llave, *A tutorial on KAM theory*, in "Smooth Ergodic Theory and its Applications" (Seattle, WA, 1999), 175–292, Proceedings of Symposia in Pure Mathematics, **69**, American Mathematical Society, Providence, RI, 2001.
[88] Russel Lyons, *On measures simultaneously 2- and 3-invariant* Israel J. Math., **61** (1988), 219–224.
[89] Ricardo Mañé, *A proof of Pesin entropy formula*, Ergodic Theory and Dynamical Systems, **1** (1981), 95–102
[90] Grigoriĭ A. Margulis, *Certain applications of ergodic theory to the investigation of manifolds of negative curvature*, Akademija Nauk SSSR. Funktsionalnyi Analiz i ego Prilozhenija **3** (1969), 89–90; English translation: Functional Analysis and its Applications, **3** (1969), 335–336.
[91] Grigoriĭ A. Margulis, *Certain measures that are connected with Y-flows on compact manifolds*, Akademija Nauk SSSR. Funktsionalnyi Analiz i ego Prilozhenija, **4** (1970), 62–76.
[92] G.A. Margulis, *On some aspects of the theory of Anosov systems*, With a survey by Richard Sharp, *Periodic orbits of hyperbolic flows*, Springer Verlag, 2003.

[93] Lev D. Mešalkin, *A case of isomorphism of Bernoulli schemes*, Dokl. Akad. Nauk SSSR, *128* (1959), 41–44.

[94] Michal Misiurewicz and Feliks Przytycki, *Topological entropy and degree of smooth mappings* Bull. Acad. Polon. Sci, ser Math. Astron Phys., **25** 1977), 573–574.

[95] Jürgen K. Moser, *On invariant curves of area-preserving mappings of an annulus*, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II **1962** (1962), 1–20; *Remark on the paper: "On invariant curves of area-preserving mappings of an annulus"*, Regul. Chaotic Dyn., **6** (2001), 337–338.

[96] Jürgen K. Moser, *Recollections. Concerning the early development of KAM theory*, in *The Arnoldfest (Toronto, ON, 1997)*, 19–21, American Mathematical Society, Providence, 1999.

[97] Ingo Müller, *Entropy: a subtle concept in thermodynamics*, in "Entropy, Andreas Greven", Gerhard Keller, Gerald Warnecke, eds., Princeton University Press, 2003.

[98] Sheldon Newhouse, *Entropy and volume*, Ergodic Theory and Dynamical Systems, **8*** (Conley Memorial Issue, (1988), 283–300.

[99] Sheldon Newhouse, *Continuity properties of entropy*, Annals of Mathematics (2), **129** (1989), 215–235; *Corrections to "Continuity properties of entropy"*, Annals of Mathematics (2), **131** (1990), 409–410.

[100] Donald S. Ornstein, *Bernoulli shifts with the same entropy are isomorphic*, Advances in Mathematics, **4** (1970), 337–352.

[101] Donald S. Ornstein, *Two Bernoulli shifts with infinite entropy are isomorphic*, Advances in Mathematics, **5** (1970), 339–348.

[102] Donald S. Ornstein, *Factors of Bernoulli shifts are Bernoulli shifts*, Advances in Mathematics, **5** (1970), 349–364.

[103] Donald S. Ornstein, *Imbedding Bernoulli shifts in flows*, 1970 Contributions to Ergodic Theory and Probability (Proc. Conf., Ohio State Univ., Columbus, Ohio, 1970), 178–218 Springer, Berlin.

[104] Donald S. Ornstein, *Some new results in the Kolmogorov–Sinai theory of entropy and ergodic theory*, Bulletin of the American Mathematical Society, **77** (1971), 878–890.

[105] Donald S. Ornstein, *Measure-preserving transformations and random processes*, Amer. Math. Monthly, **78** (1971), 833–840.

[106] Donald S. Ornstein, *An example of a Kolmogorov automorphism that is not a Bernoulli shift*, Advances in Mathematics **10** (1973) 49–62.

[107] Donald S. Ornstein, *A mixing transformation for which Pinsker's conjecture fails*, Advances in Mathematics, **10** (1973), 103–123.

[108] Donald S. Ornstein, *The isomorphism theorem for Bernoulli flows*, Advances in Mathematics, **10** (1973), 124–142.

[109] Donald S Ornstein, *Ergodic theory, randomness, and dynamical systems*, Yale Mathematical Monographs, No. 5. Yale University Press, New Haven, Conn.-London, 1974.

[110] Donald.S. Ornstein and Paul Shields, *An uncountable family of K-automorphisms*, Advances in Mathematics, **10** (1973), 63–88.

[111] Donald. S. Ornstein and Meir Smorodinsky, *Ergodic flows of positive entropy can be time changed to become K-flows*, Israel Journal of Mathematics, **29** (1977), 75–83.

[112] Donald S. Ornstein and Benjamin Weiss, *Geodesic flows are Bernoullian*, Israel J. Math., **14** (1973), 184–198.

[113] Donald S Ornstein, Daniel J Rudolph and Benjamin Weiss, *Equivalence of measure preserving transformations*, Mem. Amer. Math. Soc., **37** (1982).

[114] Donald. S. Ornstein and Benjamin Weiss, *Entropy and isomorphism theorems for the action of an amenable group,* Journal d'Analyse Mathematique, *48* (1987), 1–141.

[115] Valery I. Oseledets, *A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems*, (Russian) Trudy Moskov. Mat. Obšč., **19** (1968), 179–210.

[116] William Parry, *Intrinsic Markov chains*, Transactions of the American Mathematical Society, **112** (1964), 55–66.

[117] Jakov B. Pesin, *Families of invariant manifolds corresponding to nonzero characteristic exponents*, Mathematics of the USSR, Isvestija, **10** (1976), 1261–1305.
[118] Jakov B. Pesin, *Characteristic Ljapunov exponents, and smooth ergodic theory*, Russian Math.Surveys, **32** (1977), 55–114.
[119] Jakov B. Pesin, *Geodesic flows in closed Riemannian manifolds without focal points*, Izv. Akad. Nauk SSSR Ser. Mat., **41** (1977), 1252–1288.
[120] Jakov B. Pesin, *Formulas for the entropy of the geodesic flow on a compact Riemannian manifold without conjugate points*, Mat. Zametki, **24** (1978), 553–570.
[121] M. S. Pinsker, *Dynamical systems with completely positive or zero entropy*, Dokl. Akad. Nauk SSSR, **133** (1960), 937–938.
[122] M. S. Pinsker, *Informatsiya i informatsionnaya ustoĭchivost sluchaĭnykh velichin i protsessov*, [Information and informational stability of random variables and processes] (Russian) Problemy Peredači Informacii, Vyp. 7 Izdat. Akad. Nauk SSSR, Moscow, 1960.
[123] Vladimir Abramovich Rokhlin, *New progress in the theory of transformations with invariant measure*, Russian Math. Surveys, **15** (1960), 1–22.
[124] Vladimir Abramovich Rokhlin, *Lectures on entropy theory of measure preserving transformations*, Russian Math. Surveys, **22** (1967), 1–52.
[125] D. Rudolph, *Asymptoticall Brownian skew-products give non-loosely Bernoulli K-automorphisms*, Inv. Math., **91** (1988), 105–128.
[126] Daniel J. Rudolph, $\times 2$ *and* $\times 3$ *invariant measures and entropy*, Ergodic Theory and Dynamical Systems, **10** (1990), 395–406.
[127] David Ruelle, *An inequality for the entropy of differentiable maps*, Bol. Soc. Brasil. Mat., **9** (1978), 83–87.
[128] Claude Shannon, *A mathematical theory of communication*, Bell Systems Technical Journal, **27** (1948), 379–423, 623–656; Republished, University of Illinois Press Urbana, IL, 1963.
[129] Michael Shub, Dynamical systems, filtrations and entropy, Bull Amer. Math. Soc, **80** (1974), 27–41.
[130] Yakov Sinaĭ, *On the concept of entropy for a dynamic system*, Dokl. Akad. Nauk SSSR, **124** (1959), 768–771.
[131] Yakov Sinaĭ, *Dynamical systems with countable Lebesgue spectrum. I*, Izv. Akad. Nauk SSSR Ser. Mat., **25** (1961), 899–924.
[132] Yakov G. Sinaĭ, *A weak isomorphism of transformations with invariant measure*, Dokl. Akad. Nauk SSSR **147** (1962) 797–800; *On a weak isomorphism of transformations with invariant measure*, Mat. Sb. (N.S.), **63** (1964), 23–42.
[133] Yakov Sinaĭ, *Probabilistic concepts in ergodic theory*, Proceedings. Internat. Congr. Mathematicians (Stockholm, 1962), 540–559 Institute Mittag-Leffler, Djursholm, 1963.
[134] Yakov Sinaĭ, *Classical dynamical systems with countable Lebesgue spectrum. II*, Izv. Akad. Nauk SSSR Ser. Mat., **30** (1966), 15–68.
[135] Yakov G. Sinaĭ, *Markov partitions and Y-diffeomorphisms* Functional Analysis and its Applications, **2** (1968), 64–89.
[136] Yakov G. Sinaĭ, *Construction of Markov partitionings*, Funktsionalnyi Analiz i ego Prilozhenija, **2** (1968), 70–80.
[137] Yakov G. Sinaĭ, *Gibbs measures in ergodic theory*, Uspehi Mat. Nauk, **27** (1972), 21–64.
[138] Yakov G. Sinaĭ, *About A. N. Kolmogorov's work on the entropy of dynamical systems,* Ergodic Theory Dynam. Systems, **8** (1988), 501–502.
[139] Yakov G. Sinaĭ, *Kolmogorov's work on ergodic theory*, Ann. Probab., **17** (1989), 833–839.
[140] Meir Smorodinsky, *Information, entropy and Bernoulli systems*, Development of mathematics 1950–2000, 993–1012, Birkhäuser, Basel, (2000).
[141] Jean-Paul Thouvenot, *Quelques propriétés des systèmes dynamiques qui se décomposent en un produit de deux systèmes dont l'un est un schéma de Bernoulli* (French), Israel J. Math., **2** (1975), 177–207.

[142] Jean-Paul Thouvenot, *Entropy, isomorphism and equivalence in ergodic theory*, in Handbook of Dynamical Systems, vol. **1A**, North-Holland, Amsterdam, 2002, 205–237.
[143] Norbert Wiener, *Nonlinear problems in random theory*, MIT press, 1958, ix+131pp..
[144] Yosif Yomdin, *Volume growth and entropy*, Israel Journal of Mathematics, **57** (1987), 285–300.
[145] Yosif Yomdin, *Metric properties of semialgebraic sets and mappings and their applications in smooth analysis*, Géométrie algébrique et applications, III (La Rábida, 1984), 165–183, Travaux en Cours, **24**, Hermann, Paris, 1987.
[146] Lai-Sang Young, *Dimension, entropy and Lyapunov exponents*, Ergodic Theory Dynamical Systems, **2** (1982), 109–124.

ANATOLE KATOK <katok_a@math.psu.edu>: Department of Mathematics, The Pennsylvania State University, University Park, PA 16802-6401