BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 78, Number 5, September 1972

THE ISOMORPHISM PROBLEM IN ERGODIC THEORY

BY BENJAMIN WEISS

1. **Introduction.** About ten years ago P.R. Halmos addressed this Society on *Recent progress in ergodic theory* [11]. He closed with the wish that:

"I hope that in the near future, in the course of the next twelve years, say, humanity learns sufficiently many new answers to these fascinating old questions to warrant another society address on the subject."

A few months ago, such an address was given by D. Ornstein (*Some new results in the Kolmogorov-Sinaĭ theory of entropy and ergodic theory*, Bull. Amer. Math. Soc. **77** (1971), 878–890). I propose here to survey some recent developments in the isomorphism problem and have made an attempt to present enough background material to make these results accessible to the nonspecialist.

We shall be concerned with the basic problem of classifying measure preserving transformations. For much of the background material the survey of Halmos [10] is adequate; for later developments especially for the notion of entropy (cf. below) see Rokhlin's latest survey [28] and the works quoted there. The first part of our report will contain an exposition of past work on the classification problem while the second part will treat some recent developments in greater detail. We now proceed to fix the terminology and review the basics.

A measure space $(X,\mathscr{B},m)$ consists of an underlying set X, a σ-algebra $\mathscr{B}$ of subsets of X on which a countably additive measure m is defined. We will be considering throughout finite measure spaces which have been normalized so that $m(X) = 1$, and furthermore assume that $(X, \mathscr{B}, m)$ is a separable Lebesgue space. A measure preserving mapping (m.p.m.) is a mapping φ between two measure spaces $\varphi: X_1 \to X_2$ such that

(a) $\varphi^{-1}(B_2) \in \mathscr{B}_1$ for all $B_2 \in \mathscr{B}_2$,

(b) $m_1(\varphi^{-1}(B_2)) = m_2(B_2)$, $B_2 \in \mathscr{B}_2$.

In case the two spaces coincide φ is usually called a measure preserving transformation (m.p.t.). Given the probabilistic background of this part of ergodic theory, we shall not distinguish between φ and φ' if they disagree on a set of measure zero. For example φ, a m.p.t., is said to be invertible if there exists a φ' such that $\varphi\varphi'$ = identity (a.e.) and $\varphi'\varphi$ = identity (a.e.).

An address delivered before the Summer Meeting of the Society in University Park on September 3, 1971 by invitation of the Committee to Select Hour Speakers for Annual and Summer Meetings; received by the editors February 18, 1972.

AMS 1970 *subject classifications*. Primary 28A65; Secondary 54H20.

Such m.p.t. will be called *automorphisms*, while endomorphism will be a synonym for m.p.t.

The following are some of the basic examples in the theory.

EXAMPLE 1. X is any compact group, $\mathscr{B}$ the Borel sets, m Haar measure and φ is the rotation defined by some fixed element $a \in X$, i.e. $\varphi(x) = ax$.

EXAMPLE 2. $X, \mathscr{B}$ and m as in Example 1 but φ is now a continuous algebraic endomorphism of X onto itself. One easily checks that φ preserves the Haar measure.

EXAMPLE 3. Let $(X_0, \mathscr{B}_0, m_0)$ be some fixed measure space. Form $X = \prod_{i \in I} X_i$, the cartesian product of countably many copies of X_0 where the index set I is either the positive integers or all of Z. By the usual construction endow X with the product measure of the component measure spaces. Finally define $\varphi(x)$ by requiring

$$\pi_i(\varphi(x)) = \pi_{i+1}(x)$$

where π_i is the projection of X onto its ith coordinate. This m.p.t. is called the *bilateral* or *unilateral shift*, according to what I is. It is also referred to as a *Bernoulli shift* to emphasize the fact that the measure on X is product measure and hence the coordinate functions π_i are independent in the sense of probability theory.

EXAMPLE 4. X is as in Example 3, and so is $\mathscr{B}$, i.e. $\mathscr{B}$ is the smallest σ-algebra with respect to which the mappings $\pi_i : X \to X_i$ are measurable. Now m is chosen to be any measure on X such that φ, as defined in Example 3, becomes a m.p.t. A wide variety of such measures are known from the theory of stationary stochastic processes. In fact, there is a one to one correspondence between stationary X_0-valued processes and such measures. Borrowing terminology from that theory we refer to the shift as a *Markov shift* if the sequence $\{\pi_i\}$ forms a stationary Markov chain, and so on.

It is worth emphasizing that Example 4 is the most general example of a m.p.t. In fact, given a m.p.t. we can find a real valued function $f_0 : X \to R$ that separates points in X. Defining then $f_n(x) = f_0(\varphi^n x)$ we obtain a real valued stationary stochastic process $(\ldots, f_{-1}, f_0, f_1, \ldots)$ which enables one to find a general shift isomorphic to the original m.p.t. It is a more difficult fact that if φ is ergodic and has finite entropy (see below) then it can be represented by a general shift with a *finite* state space. This result is due to Krieger [15] who improved Rokhlin's result which yielded, under the same hypothesis, a *countable* state space with finite entropy.

Two m.p.t. $(X_i, \mathscr{B}_i, m_i, \varphi_i)$ are said to be *isomorphic* if there exist m.p.m. $\psi_1 : X_1 \to X_2$, $\psi_2 : X_2 \to X_1$ such that $\psi_1\psi_2$ (resp. $\psi_2\psi_1$) is the identity on X_2 (resp. X_1) and $\varphi_2\psi_1 = \psi_1\varphi_1$, $\varphi_1\psi_2 = \psi_2\varphi_2$. Again it should be emphasized that these equations are only required to hold a.e. Unlike the case in the category of topological spaces the problem of classifying the underlying measure space is trivial, namely any separable Lebesgue space is isomorphic to the unit interval with Lebesgue measure together with at most countably many atoms. A weaker notion of isomorphism in which the conditions that $\psi_2\psi_1$ and $\psi_1\psi_2$ be the identity is dropped was introduced by Sinaĭ [31]. To date it is not known whether or not weak isomorphism implies isomorphism. There is another way of weakening the notion of isomorphism for endomorphisms that are not automorphisms. To describe this we need the idea of a natural extension due to Rokhlin [27].

An automorphism $(\tilde{X},\tilde{\mathscr{B}},\tilde{m},\tilde{\varphi})$ is said to be a natural extension of $(X,\mathscr{B},m,\varphi)$ if there exists a m.p.m. $\pi:\tilde{X}\to X$ such that (1) $\varphi\pi = \pi\tilde{\varphi}$ and (2) $\tilde{\mathscr{B}}$ is the smallest $\tilde{\varphi}$-invariant σ-algebra that contains $\pi^{-1}(\mathscr{B})$. Rokhlin has shown that the natural extension exists, uniquely up to isomorphisms. The basic example to keep in mind is that of the unilateral shift whose natural extension is the bilateral shift.

To be more precise, suppose that $(X,\mathscr{B},m,\varphi)$ is a unilateral shift with m any φ-invariant measure, as in Example 4. Then for $\tilde{X}$, $\tilde{\mathscr{B}}$, $\tilde{\varphi}$ we take the corresponding bilateral shift and it remains only to define $\tilde{m}$. The standard extension theorems of measure theory imply that it suffices to define m in a consistent fashion on finitely based cylinder sets. But since a finitely based cylinder set $\tilde{C}$ in $\tilde{X}$ can also be considered as a finitely based cylinder set C in X (unique up to a shift by φ) we can use m itself to define $\tilde{m}$ on such sets, namely set $\tilde{m}(\tilde{C}) = m(C)$. It is straightforward to check that conditions (1) and (2) are satisfied.

Two endomorphisms will now be said to be *quasi-isomorphic* if their natural extensions are isomorphic as automorphisms. This concept of isomorphism is weaker than true isomorphism, since there are unilateral Markov shifts which are not isomorphic as unilateral shifts but whose bilateral extensions are isomorphic. We shall discuss this matter in greater detail later on. This notion of *quasi-isomorphism* was first used by Katznelson [14] in analyzing the ergodic endomorphisms of a finite-dimensional torus.

2. **Invariants.** A property or object associated with a m.p.t. is said to be an *invariant* if it is the same for isomorphic m.p.t.'s. The basic technique to study the classification problem is to find a sufficient number of invariants to form a complete set, i.e. a set of invariants large enough so that if all the invariants agree for two m.p.t. one can conclude that the m.p.t. are isomorphic. Historically, the first such invariants to be discussed were ergodicity and the unitary operator associated with a m.p.t. The

latter is constructed as follows: Let $H = L^2(X,\mathscr{B},m)$, and define

$$U_\varphi f\cdot(x) = f(\varphi x). \tag{2.1}$$

It is easy to verify that U_φ is an isometry; it is clearly invertible if and only if φ is. An early question in the field was to what extent does the isomorphism of φ_1 and φ_2 follow from the unitary equivalence of U_{φ_1} and U_{φ_2}. In the next section we shall recall some well-known results along these lines.

One can view ergodicity as a part of the structure of U_φ, namely, a m.p.t. φ is said to be *ergodic* if the dimension of the solution space of $U_\varphi f = f$ is one. A more familiar formulation, based on the ergodic theorem, is this: φ is said to be ergodic if, for all $A, B \in \mathscr{B}$,

$$\lim_{n\to\infty} 1/n \sum_0^{n-1} m(\varphi^{-i}A \cap B) = m(A)m(B). \tag{2.2}$$

Referring back to the examples in §1, conditions for ergodicity are:

Example 1. The cyclic group generated by "a" must be dense in X. This forces X to be a commutative monothetic group. The simplest infinite example is an irrational rotation of the circle.

Example 2. In case X is commutative the condition for ergodicity takes a simple form, namely φ is ergodic if and only if $\hat{\varphi}$, the dual transformation of the dual group $\hat{X}$, has no periodic orbits.

Example 3. All Bernoulli shifts are ergodic. This follows readily from the basic zero-one law for independent random variables.

Since all transformations can be represented in the form of Example 4, no simple criterion, apart from the definitions, should be expected here.

A condition, stronger than ergodicity, is that of *strong mixing* in which the averages of (2.2) are replaced by an ordinary limit, i.e. φ is *strongly-mixing* if

$$\lim_{n\to\infty} m(\varphi^{-n}A \cap B) = m(A)m(B), \quad \text{all } A, B \in \mathscr{B}. \tag{2.3}$$

In the examples above it is easy to see that rotations are never strongly mixing, the others, i.e. Example 2—Example 3 satisfy even stronger conditions which will be discussed in a moment. In Example 4, a criterion for strong mixing is: the shift is s.m. for a measure m if and only if (2.3) holds for all finitely based cylinder sets A, B. If a certain uniformity is introduced in this asymptotic independence we obtain an important class of transformations known as the *Kolmogorov-automorphisms*, and the *exact endomorphisms*. Still referring to Example 4 the definitions are: a bilateral (unilateral) general shift is said to be a K-automorphism (exact endomorphism) if

$$\lim_{n\to\infty} \sup_{A\in\mathscr{A}_n} |m(A\cap B) - m(A)m(B)| = 0, \tag{2.4}$$

for every fixed $B \in \mathscr{B}$, where $\mathscr{A}_n$ is the sub-σ-algebra of $\mathscr{B}$ generated by cylinder sets based on the coordinates $(n, n+1, \ldots)$. An equivalent condition, which is easier to check is

$$\bigcap_{n=1}^{\infty} \mathscr{A}_n = \{\emptyset, x\} \quad \text{a.e.}, \tag{2.5}$$

i.e. the tail σ-algebra is trivial. We have given (2.4) to emphasize the fact that this generalizes strong mixing. For transformations not given as shifts the definitions are: an automorphism $(X, \mathscr{B}, m, \varphi)$ is said to be a K-automorphism if there exists a sub-σ-algebra $\mathscr{A}_0 \subseteq \mathscr{B}$ satisfying

$$\varphi^{-1}\mathscr{A}_0 \subseteq \mathscr{A}_0; \tag{2.6}$$

$$\bigcup_{-\infty}^{\infty} \varphi^n \mathscr{A}_0 \quad \text{generates } \mathscr{B}; \tag{2.7}$$

$$\bigcap_{-\infty}^{0} \varphi^n \mathscr{A}_0 \quad \text{is trivial.} \tag{2.8}$$

For an endomorphism one can dispense with (2.6) and (2.7) and one has, φ is said to be *exact* if

$$\bigcap_{1}^{\infty} \varphi^{-n}\mathscr{B} \quad \text{is trivial.} \tag{2.9}$$

It has been shown **[26]** (**[13]** for the noncommutative case) that ergodic automorphisms (endomorphisms) of compact groups are K-automorphisms (exact). With this we conclude for the moment our brief survey of some of the invariant properties of m.p.t. and pass to some of the early results on the isomorphism problem.

3. **Algebraic isomorphism results.** A m.p.t. is said to have *discrete* or *pure point* spectrum if the set of eigenfunctions, that is, functions $f \in L^2(X, \mathscr{B}, m)$ for which there exists a complex number λ so that

$$f(\varphi x) = \lambda \cdot f(x) \quad \text{a.e.}, \tag{3.1}$$

spans $L^2(X, \mathscr{B}, m)$. The following theorem is due to von Neumann **[33]** in 1932.

THEOREM 1. *Two ergodic m.p.t. φ_1, φ_2 that have discrete spectrum are isomorphic if and only if U_{φ_1} is unitarily equivalent to U_{φ_2}.*

This was the first positive isomorphism result in ergodic theory, and remained the only one for almost twenty years. The proof may be found in any of the standard treatises, such as [10]. The algebraic nature of this result is worth emphasizing. In the course of the proof, it is first shown that the e.m.p.t. φ_i (ergodic m.p.t.) with discrete spectrum are isomorphic to a rotation of compact groups X_i by elements a_i, say, where X_i and a_i depend only on U_{φ_i}. This of course proves the theorem, but something further is also true, namely if two ergodic rotations are isomorphic then any isomorphism between them is necessarily algebraic, i.e. if

$$\varphi_i : X_i \to X_i, \qquad \varphi_i(x) = a_i x, \qquad i = 1,2, \tag{3.2}$$

and $v_1 : X_1 \to X_2$ satisfies the conditions for being a measure theoretic isomorphism then v_1 is an *algebraic* isomorphism between X_1 and X_2. This is most easily seen by considering the mapping that v_1 induces on $\hat{X}_2$ to $\hat{X}_1$, the respective dual groups.

Even in this simplest case of discrete spectrum when the hypothesis of ergodicity is dropped the situation is no longer so simple. Choksi [7] has investigated this question, and has shown that without ergodicity unitary equivalence does not even imply a weakened version of measure theoretic isomorphism. The next development in this circle of ideas was taken by Abramov [1] who gave an algebraic characterization of *totally ergodic* m.p.t. with *quasi-discrete spectrum.* Totally ergodic simply means that all powers of φ are ergodic, or in other words that no roots of unity (other than 1) occur in the spectrum of U_φ. Quasi-discrete spectrum depends for its definition on the generalized eigenvalues used by Halmos [10] to give examples of transformations that are unitarily equivalent but nonisomorphic.

Let G_0 denote the circle group $(z : |z| = 1)$, and $H_0 = \{1\}$, and let K denote the set of functions on X of constant modulus 1, $G_0 \subset K$ as the constant functions. Then define inductively

$$H_{n+1} = \{g \in G_n : \text{there exists } f \in K \text{ for which (3.4) holds}\}, \tag{3.3}$$

$$f(\varphi x) = g(x) f(x) \quad \text{a.e.}, \tag{3.4}$$

$$G_{n+1} = \{f \in K : \text{(3.4) holds for some } g \in H_{n+1}\}.$$

H_1 is simply the set of eigenvalues of φ, and G_1 the eigenfunctions of modulus 1. The H_n's and G_n's form subgroups of K, and a natural homomorphism v: $G_n \to H_n$ is defined by

$$v(f) = f(\varphi x)/f(x). \tag{3.5}$$

Note that (3.5) defines an endomorphism of K, and that $G_n = \ker(v^{n+1})$. Finally φ is said to have *quasi-discrete spectrum* if $G = \bigcup G_n$ spans

$L^2(X, \mathscr{B}, m)$. The theorem of Abramov on the classification of these transformations is:

THEOREM 2. *If φ, φ' are totally ergodic m.p.t. with quasi-discrete spectrum and there is an algebraic isomorphism between the corresponding (G, v), (G', v'), then φ is isomorphic to φ', and conversely.*

Here too the method of proof is to construct a "model" transformation for a given (G, v), which turns out to be a unipotent affine transformation of some compact group. The example to keep in mind here is the following: Let X be the n-torus and define $\varphi: X \to X$ by

$$\varphi(x_1, x_2, x_3, \ldots, x_n) = \left(x_1 + \alpha, x_1 + x_2, x_1 + x_2 + x_3, \ldots, \sum_1^n x_i\right). \tag{3.6}$$

There is also a theorem which says that any isomorphism between transformations with quasi-discrete spectrum must be algebraic. For further results along these lines see W. Parry [21], who has obtained results similar to the above, for unipotent affine transformations of nilmanifolds. The isomorphism problem is reduced to an algebraic one in the strong sense that *any* metric isomorphism is a.e. an algebraic conjugacy.

4. **Entropy—definition and properties.** In 1958 A.N. Kolmogorov introduced a new invariant, the entropy of a m.p.t. and succeeded in applying it to distinguish between the hitherto indistinguishable transformations —the 2-shift and the 3-shift. Since this invariant has played a central role in all subsequent work on the isomorphism problem we will give a quick account of it here. For a detailed treatment, including proofs of the results to be stated next the reader may consult [28], [22], and [4] for example. The basis for the new invariant was the notion of the information content of a countably valued random variable, or what is the same as a countable partition of a probability space, a notion first introduced and exploited by C. Shannon in his theory of transmission along noisy channels. First let us establish our notation for the calculus of partitions.

$\alpha, \beta, \gamma, \ldots$ will denote countable partitions of a finite, normalized measure space $(X, \mathscr{B}, m)$.

$\alpha = \{A_1, A_2, \ldots\}$ etc. and when it is convenient, we will assume that a definite ordering of α is given.

$\alpha \vee \beta = \{A_i \cap B_j\}$ is the *join* or least common refinement of α and β. $\bigvee_{i=1}^n \alpha_i$ is the join of α_1 and $\alpha_2, \ldots$ and α_n, while $\bigvee_{i=1}^\infty \alpha_i$ denotes the smallest σ-algebra that contains all elements of $\bigvee_{i=1}^n \alpha_i$ for $n = 1, 2, \ldots$. In general we will often identify a partition with the subalgebra of $\mathscr{B}$ that it determines, and so we will write $\alpha \subset \beta$ and mean that β is a *refine-*

ment of α, i.e. the algebra determined by β contains the algebra determined by α.

DEFINITION 1. (a) The *entropy* of α is given by

$$H(\alpha) = -\sum_{A_i \in \alpha} m(A_i) \log m(A_i). \tag{4.1}$$

(b) The class of partitions α such that $H(\alpha)$ is finite is denoted by Z. A useful extension of this notion is

DEFINITION 2. The *conditional entropy* of α *given* $\mathscr{B}'$, a subalgebra of $\mathscr{B}$ is

$$\int_X \sum_{A_i \in \alpha} m(A_i|\mathscr{B}') \log m(A_i|\mathscr{B}')\, dm, \tag{4.2}$$

where $m(A|\mathscr{B}')$ is the conditional measure of A given $\mathscr{B}'$, i.e. that $\mathscr{B}'$-measurable function that satisfies

$$\int_B m(A|\mathscr{B}')\, dm = m(A \cap B) \tag{4.3}$$

for all $B \in \mathscr{B}'$.

The basic properties of the entropy of a partition are summarized in the following proposition.

PROPOSITION. (1) $H(\alpha \vee \beta|\gamma) = H(\alpha|\gamma) + H(\beta|\alpha \vee \gamma)$.
(2) $H(\alpha|\mathscr{B}') \leqq H(\beta|\mathscr{B}')$ *if* $\alpha \subset \beta$.
(3) $H(\alpha|\mathscr{B}') \leqq H(\alpha|\mathscr{B}'')$ *if* $\mathscr{B}'' \subset \mathscr{B}'$.
(4) *If* $\mathscr{B}_n \uparrow \mathscr{B}_\infty$ *then* $H(\alpha|\mathscr{B}_n) \downarrow H(\alpha|\mathscr{B}_\infty)$.

Suppose now that $\{f_i\}$ is a countably valued stochastic process defined on $(X, \mathscr{B}, m)$ with the time shift defined by a m.p.t. $\varphi: X \to X$. Recall that this means that $f_n(x) = f_0(\varphi^n x)$. Let α denote the partition into sets of constant value for f_0, then $\varphi^{-n}\alpha$ consists of the sets of constancy for f_n. Now we will also use the notation

$$H(f_0|f_1, \ldots, f_n) = H(\alpha|\varphi^{-1}\alpha \vee \ldots \vee \varphi^{-n}\alpha). \tag{4.4}$$

By the Proposition, (4), we can pass to the limit and define

$$h(\varphi, \alpha) = \lim_{n\to\infty} H(f_0|f_1, \ldots, f_n) = H\left(\alpha \Big| \bigvee_1^\infty \varphi^{-n}\alpha\right). \tag{4.5}$$

That is to say, the entropy of φ with respect to α is the amount of information contained in f_0 given the future. If $h(\varphi, \alpha) = 0$ then we say that the process $\{f_n\}$ is *deterministic*; f_0 is measurable with respect to the future. In case the process is bilateral we could just as well have worked with $H(f_0|f_{-1}, \ldots, f_{-n})$ which explains a little better the usage of determinism.

Finally we have

DEFINITION 3. The *entropy* of a m.p.t. φ is given by

$$\sup_{\alpha \in Z} h(\varphi, \alpha) = h(\varphi). \tag{4.6}$$

Its obvious from the definition that $h(\varphi)$ is an invariant under isomorphism of m.p.t. The basic result that enables one to compute the supremum in (4.6) is the following:

PROPOSITION *If α is a φ-generator of $\mathscr{B}$, i.e. if the smallest φ-invariant σ-algebra that contains α is all of $\mathscr{B}$ then $h(\varphi) = h(\varphi, \alpha)$.*

It is now a straightforward matter to compute the entropy of a Bernoulli shift which we denote now by the probability vector that represents the distribution of the individual coordinate function. The answer is

$$h(p_1, \ldots, p_k) = -\sum p_j \log p_i = H(\alpha) \tag{4.7}$$

where α is the basic partition according to the values assumed by the zeroth coordinate. Thus we have proved

THEOREM 3 (KOLMOGOROV). *The k-shift $(1/k, 1/k, \ldots, 1/k)$ is not isomorphic to the l-shift $(1/l, \ldots, 1/l)$ for $k \neq l$.*

With this result a new era ensued in ergodic theory, and especially the isomorphism questions we are dealing with. In the same year that Kolmogorov's paper appeared, L. D. Mešalkin [16] published the first positive result by showing that for a certain class of Bernoulli shifts, entropy is a complete invariant, i.e. he showed how to construct isomorphism between special Bernoulli shifts with the same entropy. We shall give some indication of his result in the next section.

A surprising application of entropy was the characterization by Pinsker-Rokhlin-Sinaĭ [29] of Kolomogorov automorphisms as those m.p.t.'s φ with *completely positive entropy*, which means simply that $h(\varphi, \alpha) > 0$ for every nontrivial partition α. It soon became clear that among transformations with zero entropy a wide variety of nonisomorphic types was possible (including all those algebraic examples in §3). However it was not possible to distinguish between two K-automorphisms with the same entropy, and so the suspicion grew that for K-automorphisms entropy was a complete invariant. This has turned out to be false [20], but first let us continue the story which tells to what extent entropy alone does succeed in classifying m.p.t.'s.

5. **Entropy—as a complete invariant.** We have already alluded to Mešalkin's result, which was somewhat extended by J. Blum and

D. Hanson [5]. A rather full account of this work may be found in Jacobs [12], so that I will content myself with presenting an example which will convey the flavor of the method. Observe that $h(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}) = h(\frac{1}{2},\frac{1}{8},\frac{1}{8},\frac{1}{8},\frac{1}{8})$. Our task is to construct a mapping between the 4-shift and the shift on five symbols with measure $(\frac{1}{2},\frac{1}{8},\frac{1}{8},\frac{1}{8},\frac{1}{8})$. To this end denote the four symbols by A_1, A_2, B_1, B_2 and the space of five symbols by a, b_{11}, b_{12}, b_{21}, b_{22}. Suppose then we are given a sequence of A_i's, B_j's; referring to the display (5.1), the mapping may be described in several steps.

$$
\begin{array}{lcccccccccccc}
 & ...B_1 & A_1 & A_2 & B_2 & A_2 & A_1 & B_2 & B_2 & B_2 & A_1 & A_2 & B_1 \\
(5.1) & & & & & & & & & & & & \\
 & ...b_? & a & a & b_{22} & a & a & b_{12} & b_{22} & b_{12} & a & a & b_{21} \\
 & & & & & & & & & & & & \\
 & & & & \mathrm{I} & & & \mathrm{I} & \mathrm{II} & \mathrm{III} & & & \mathrm{I}
\end{array}
$$

Step 0. Write out a sequence of lower case a's, b's underneath the given sequence, paying no attention to subscripts.

Step I. Group adjacent A_iB_j and give the b underneath the B_j the subscript ij. Cross out these A's and B's.

Step II. Group the newly adjacent A_iB_j, and give the b underneath the B_j the subscript ij. Again cross out A's and B's whose subscripts we have used.

Continue on in the same fashion. We have indicated in (5.1) at what step the subscript was attached to the lower case b's. The question-mark indicates that one has to extend the sequence to the left to decide how to subscript that b.

Naturally, it is not the case that starting from any sequence of A_i's, B_j's such a procedure will eventually assign a subscript to every b. However, since the measure on the space of A's and B's is such that A and B occur independently with equal probability, it is a consequence of the recurrence of simple random walks in one dimension that for almost every sequence such a procedure will eventually assign a mate to every B, and every A will have a B assigned to it. Indeed the recurrence implies that from any starting point, with probability one, there will eventually be an equal number of A's and B's to the right of the starting point, which implies (if the symbol with which we begin is an A) that each A is mated to a B and conversely.

From its very nature it is clear that the mapping we have defined commutes with the respective shifts, indeed since no use has been made of a starting point we have actually defined the mapping on whole orbits. The invertibility also presents no problem; one uses the same groupings to index the A's and B's. A bit of work is required to show directly that the measures are preserved, i.e. that we have an isomorphism of the m.p.t.'s. One can use instead the fact that $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ is an intrinsically ergodic system [34], which leads directly to the required conclusion.

Mešalkin's general result extends this combinatorial type of argument to any two Bernoulli shifts whose probabilities are of the form k/p^l for some prime p (in our example $p = 2$). A few years after Mešalkin, Ya. G. Sinaĭ [31] showed that any two Bernoulli shifts are *weakly isomorphic.*

THEOREM 4 (SINAĬ). *If* $(X, \mathscr{B}, m, \varphi)$ *is ergodic and*

$$h(\varphi) \geqq - \sum_{1}^{n} p_j \log p_j, \tag{5.2}$$

then there is a mapping π *from* X *onto the* $(p_1, \ldots, p_n)$*-shift* σ*, that preserves measures such that* $\sigma\pi = \pi\varphi$.

The first use of entropy to completely classify a natural class of transformations was made by R. Adler and myself [3a] in a study of the automorphisms of the 2-torus. A rather complete treatment appeared somewhat later [3b], and I shall content myself with explaining the two key ideas in our proof of the fact that two automorphisms of the torus with the same entropy are isomorphic. The first was a geometric construction of a generating partition that was Markovian, i.e. it established an isomorphism between the automorphism and a Markov shift. This idea has been carried further by Ya. G. Sinaĭ [32], for Anosov differmorphisms, R. Bowen [6] for Axiom A diffeomorphisms and M. Ratner [24] for Anosov flows. The second idea was to establish an isomorphism between two Markov shifts with the same entropy. Here we exploited the special form of the Markov transition matrix that was involved to give a coding of a combinatorial nature. As I have already indicated this has been generalized by N. Friedman and D. Ornstein [8] who showed that any Markov shifts with the same entropy are isomorphic. Last year, by using results of D. Ornstein in a more direct fashion, Y. Katznelson [14] was able to generalize our theorem to automorphisms of the torus in any number of dimensions.

6. A necessary and sufficient condition for a transformation to be isomorphic to a Bernoulli shift. Work, in the last few years, on the isomorphism problem has been dominated by the astonishing results of D. Ornstein [17] - [20]. He developed new techniques that enabled him to show —first that B-shifts with the same entropy are isomorphic, and then to extend these isomorphism results to wider and wider classes of transformations. Although we will not be able to give proofs we shall formulate some of the results that have already been widely applied. A key notion has been that of approximate independence, which we proceed to formulate.

DEFINITION. Two partitions α and β are said to be ε-independent if

$$\sum_{i,j} |m(A_i \cap B_j) - m(A_i)m(B_j)| \leqq \varepsilon. \tag{6.1}$$

The relation between entropy and this notion of almost independence is provided by the easy

PROPOSITION. *If $H(\alpha) - H(\alpha|\beta) \leqq \delta$ then α and β are $\varepsilon(\delta)$ independent, where $\varepsilon(\delta) \to 0$ with δ.*

Using this it is possible to define what a weakly Bernoulli partition is.

DEFINITION. α is *weakly Bernoulli* for a m.p.t. φ if given ε there exists a $K = K(\varepsilon)$ such that $\bigvee_1^n \varphi^{-j}\alpha$ is ε-independent of $\bigvee_{n+K}^{2n+K} \varphi^{-j}\alpha$ for all n.

Notice that this makes sense even for noninvertible transformations. For easy examples of weakly Bernoulli partitions one need look no further than aperiodic ergodic finite state Markov chains. To place this notion in the framework of the mixing conditions that we discussed in §2 observe that for φ to be a K-automorphism, if α is a generator, it is necessary and sufficient that given ε, and N there exists a $K = K(\varepsilon, N)$ such that $\bigvee_1^N \varphi^{-j}\alpha$ and $\bigvee_{N+K}^{N+K+n} \varphi^{-j}\alpha$ be ε-independent for all n.

THEOREM 5. (a) *If α is a weakly Bernoulli generator for invertible φ then φ is isomorphic to any Bernoulli shift having the same entropy as φ.*

(b) *In case φ is not invertible, if α is a weakly Bernoulli generator then $\tilde{\varphi}$, the natural extension of φ, is isomorphic to any Bernoulli shift with the same entropy.*

This is the main result of [**8**]. I should like to emphasize that these techniques do not yield isomorphisms of noninvertible transformations qua noninvertible transformations. It is easy to give examples of transformations whose natural extensions are isomorphic but which are not isomorphic. For example, it was observed in [**9**] that the one sided Markov chain determined by $\binom{p\ q}{q\ p}$, $p \neq q$, is not isomorphic to the one sided B-shift with probabilities (p, q), while it follows from the above theorem that their natural extensions are isomorphic. We shall return to this point again later.

In a fascinating continuation of this work, D. Ornstein has actually given necessary and sufficient conditions for a transformation to be isomorphic to a Bernoulli shift. To formulate this result we need to define two metrics on sets of partitions of the form $\{\alpha, \varphi^{-1}\alpha, \ldots, \varphi^{-n+1}\alpha\}$.

The first is denoted by

$$d\left(\bigvee_0^{n-1} \varphi^{-j}\alpha, \quad \bigvee_0^{n-1} \psi^{-j}\beta\right) \tag{6.2}$$

and is simply the ordinary L^1 distance between the distribution of the ordered partition $\bigvee_0^{n-1} \varphi^{-j}\alpha$. The second is

$$\bar{d}(\{\varphi^{-j}\alpha\}_0^{n-1}, \{\psi^{-j}\beta\}_0^{n-1}) \tag{6.3}$$

and is defined as

$$\inf \frac{1}{n} \sum_0^{n-1} D(\alpha_i, \beta_i) \tag{6.4}$$

where the infimum is over all sets $\{\alpha_i\}$ and $\{\beta_i\}$ such that $d(\bigvee_0^{n-1}\alpha_i, \bigvee_0^{n-1}\varphi^{-i}\alpha) = 0$, $d(\bigvee_0^{n-1}\beta_i, \bigvee_0^{n-1}\psi^{-j}\beta) = 0$, and $D(\alpha, \beta) = \sum_i m(A_i \,\Delta\, B_i)$ where Δ denotes symmetric difference. Now a partition α is said to be *finitely determined* for φ if given ε there exists a δ and n such that if ψ is mixing, $h(\psi) \geqq h(\varphi)$, and β satisfies

(1) $d(\bigvee_0^{n-1}\varphi^{-i}\alpha, \bigvee_0^{n-1}\psi^{-i}\beta) < \delta$,

(2) $|h(\varphi, \alpha) - h(\psi, \beta)| < \delta$,

then, for *all* m,

$$\bar{d}(\{\varphi_\alpha^{-i}\}_0^m, \{\psi_\beta^{-i}\}_0^m) < \varepsilon. \tag{6.5}$$

What this means is that the joint distribution of a finite number of $\varphi^j\alpha$'s already fixes the entire distribution in some sense. In **[18]** Ornstein proved

THEOREM 6. *φ is a Bernoulli shift if and only if every finite partition is finitely determined.*

In particular any factor of a B-shift is a B-shift, if T^k is a B-shift so is T, and so on. It was by using this condition that he was later able to construct examples of K-automorphisms that are not B-shifts.

7. **Recent results and problems.** Some idea of the recent activity in the subject will now be given by listing some classes of transformations that have recently been shown to be isomorphic to Bernoulli shifts.

(a) *f-expansions.* Let φ be a piecewise continuous mapping of $(0, 1)$ into itself satisfying the properties:

(1) φ is twice differentiable on its intervals of continuity.

(2) φ maps its maximal intervals of continuity onto $(0, 1)$. Such transformations are associated with various number theoretical expansions of real numbers, such as the continued fraction expansion. R. Adler and I have shown that if φ satisfies the following two conditions then

there is an invariant measure for φ absolutely continuous with respect to Lebesgue measure and the natural extension of φ with respect to this measure is isomorphic to any Bernoulli shift with the same entropy. The two conditions are

(a) for some k, $|d\varphi^k(x)/dx| \geqq c > 1$ where φ^k is the kth iterate of φ;
(b) $\sup_I \sup_{x,y\in I} |\varphi''(x)|/|\varphi'(y)^2| \leqq M$, where I ranges over the intervals of continuity of φ.

The method of proof involves exploiting (b) to control the deviation of the iterates of φ from a piecewise linear function. The details will appear elsewhere. Related results have been obtained by S. Rudolfer **[30]**. As examples of transformations that satisfy our conditions one can take $\varphi(x) = 1/x^k - n$ on the interval $[(n+1)^{-1/k}, n^{-1/k}]$, $k = 1, 2, \ldots$. The case $k = 1$ corresponds to the continued fraction expansion.

(b) *β-transformation.* Here the mapping is simply $\varphi: [0, 1] \to [0, 1]$, $\varphi(x) = \beta x \pmod 1$ where $\beta > 1$. For β not an integer the difficulty in the analysis of this transformation arises from the fact that φ does not map every maximal interval of continuity onto $(0, 1)$. Here again R. Adler and I, and independently M. Smorodinsky have shown that φ, as a m.p.t. with respect to the invariant measure discovered by Renyi **[25]**, that is absolutely continuous with respect to Lebesgue measure, is quasi-isomorphic to any Bernoulli shift with entropy $\log \beta$.

Since there are values of β for which it is known that φ is a one sided Markov chain, for some values of β I know that φ is not isomorphic to a one sided Bernoulli shift. The corresponding question for f-expansions seems to be still wide open.

(c) *Automorphism of nilmanifolds.* Here the results are still rather fragmentary. By using the methods of Y. Katznelson, I have been able to show that a skew product with a Bernoulli shift as a base and ergodic automorphisms of the torus on the fibers is isomorphic to a Bernoulli shift. Now it is well known that an automorphism of a nilmanifold is built up as a finite sequence of skew products with automorphisms of tori. The difficulty lies in the fact that the conditions for ergodicity of the automorphism do not necessitate the ergodicity of all of the toral automorphisms that occur in building up the manifold automorphism. From these remarks the interested reader can easily reconstruct the results that I have obtained and since I am not sure how promising this line is I will not spell things out anymore.

(d) *Skew products.* If φ is a K-automorphism then there is a maximal subalgebra such that φ is a Bernoulli shift when we restrict to that subalgebra. Hence any K-automorphism can be viewed as a skew product with a Bernoulli shift as a base, i.e. a transformation of the form $(x, y) \to$

$(\sigma x, \varphi(x)y)$ where $\sigma: X \to X$ is a Bernoulli shift and $\varphi(x)$ is a m.p.t. of Y for every x. Thus a natural question is under what circumstances is a skew product with B-shift a base again a B-shift. We mentioned one such result under (c). Another has been obtained by R. Adler and P. Shields [2]. It says the following:

THEOREM 7. *Let (X, σ) be a Bernoulli shift on two symbols and Y the circle. If $\varphi(x)y$ is defined by*

$$\begin{aligned} \varphi(x)y &= y + \alpha \quad \text{if} \quad x \in A_1, \\ &= y + \beta \quad \text{if} \quad x \in A_2, \end{aligned}$$

where (A_1, A_2) is the independent generator of (X, σ), and $\alpha - \beta$ is irrational, then the skew product defined by $\psi(x, y) = (\sigma x, \varphi(x)y)$ is a Bernoulli shift.

This result can be extended somewhat by replacing rotations by other isometries, but at the present the generalization to functions φ which depend on X in a more complicated way seems to be difficult.

(e) *Automorphisms of a solenoid.* By using the method of Y. Katznelson I have been able to show that the ergodic automorphisms of a solenoidal group, that is the dual of a subgroup of the rationals, is a Bernoulli shift. Details will appear elsewhere.

I would like to conclude this report with a few open problems of a more general character.

(1) Replace the powers of a single transformation φ by a group generated by several commuting transformations. The model for a Bernoulli shift should now be an array x_{ij} of independent identically distributed random variables with a horizontal shift and a vertical shift. It is likely that a generalization of entropy to such a situation will prove necessary.

(2) Study what kind of skew products can be Bernoulli shifts if the base is given as a B-shift. Here the following simple example is as yet undecided: (X, σ) a B-shift and (Y, φ) another.

$$\psi(x, y) = (\sigma x, \varphi(x)y)$$

where $\varphi(x) = \varphi$ or φ^{-1} depending upon whether $x \in A_1$ or A_2 where (A_1, A_2) is an independent generator. It is not difficult to show that ψ is a K-automorphism. Is it a B-shift?

(3) The explicit codes of [3] and [16] have certain properties that are not shared by the general isomorphisms of Ornstein. Is there some stronger notion of isomorphism that underlies this "almost continuous" coding?

(4) The classification of noninvertible transformations *qua* transformations and not via their natural extensions is still wide open. Here even the problem of endomorphisms of the 2-torus is yet to be settled, namely are they isomorphic to one sided B-shifts? I should point out that this is not the case for all of the β-transformations. For particular values of β, I can show that they are *not* isomorphic to one sided B-shifts. Recently some striking examples of the type of phenomenon that can occur have been discovered by Parry and Walters [23].

ADDED IN PROOF. Problem (1) has been settled by Y. Katznelson and myself in a forthcoming paper entitled *Commuting of measure preserving transformations*, Israel J. Math. Similar results have been obtained by J.-P. Thouvenot based on work of J. P. Conze.

BIBLIOGRAPHY

1. L. M. Abramov, *Metric automorphisms with a quasi-discrete spectrum,* Izv. Akad. Nauk SSSR Ser. Mat. **26** (1962), 513–530; English transl., Amer. Math. Soc. Transl. (2) **39** (1964), 37–56. MR **26** #606.

2. R. L. Adler and P. Shields, *Skew products of Bernoulli shifts with rotations,* Israel J. Math. (to appear).

3a. R. L. Adler and B. Weiss, *Entropy, a complete metric invariant for automorphisms of the torus,* Proc. Nat. Acad. Sci. U.S.A. **57** (1967), 1573–1576. MR **35** #3031.

3b. ———, *Similarity of automorphisms of the torus,* Mem. Amer. Math. Soc. No. 98 (1970). MR **41** #1966.

4. P. Billingsley, *Ergodic theory and information,* Wiley, New York, 1965. MR **33** #254.

5. J. R. Blum and D. L. Hanson, *On the isomorphism problem for Bernoulli schemes,* Bull. Amer. Math. Soc. **69** (1963), 221–223. MR **26** #1412.

6. R. Bowen, *Markov partitions for axiom A diffeomorphisms,* Amer. J. Math. **92** (1970), 725–747. MR **43** #2740.

7. J. R. Choksi, *Non-ergodic transformations with discrete spectrum,* Illinois J. Math. **9** (1965), 307–320. MR **30** #4903.

8. N. Friedman and D. Ornstein, *On isomorphism of weak Bernoulli transformations,* Advances in Math. **5** (1970), 365–394. MR **43** #478c.

9. H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation,* Math. Systems Theory **1** (1967), 1–49. MR **35** #4369.

10. P. R. Halmos, *Lectures on ergodic theory,* Publ. Math. Soc. Japan, no. 3, Math. Soc. Japan, Tokyo, 1956. MR **20** #3958.

11. ———, *Recent progress in ergodic theory,* Bull. Amer. Math. Soc. **67** (1961), 70–80. MR **23** #A290.

12. K. Jacobs, *Lecture notes on ergodic theory.* Parts I, II, Matematisk Institut, Aarhus Universitet, Aarhus, 1963. MR **28** #3238; errata, p. 1247.

13. S. A. Juzvinskiĭ, *Metric properties of endomorphisms of compact groups,* Izv. Akad. Nauk SSSR Ser. Mat. **29** (1965), 1295–1328; English transl., Amer. Math. Soc. Transl. (2) **66** (1968), 63–98. MR **33** #2798.

14. Y. Katznelson, *Ergodic automorphisms of T^n are Bernoulli shifts,* Israel J. Math. **10** (1971), 186–195.

15. W. Krieger, *On entropy and generators of measure-preserving transformations,* Trans. Amer. Math. Soc. **149** (1970), 453–464. MR **41** #3710.

16. L. D. Mešalkin, *A case of isomorphism of Bernouilli schemes,* Dokl. Akad. Nauk SSSR **128** (1959), 41–44. (Russian) MR **22** #1650.

17. D. Ornstein, *Bernoulli shifts with the same entropy are isomorphic,* Advances in Math. **4** (1970), 337–352. MR **41** #1973.

18. ———, *Two Bernoulli shifts with infinite entropy are isomorphic,* Advances in Math. **5** (1970), 339–348. MR **42** #478a.

19. ———, *Factors of Bernoulli shifts are Bernoulli shifts,* Advances in Math. **5** (1970), 349–364. MR **42** #478b.

20. ———, *An example of a Kolmogorov automorphism that is not a Bernoulli shift* (to appear).

21. W. Parry, *Metric classification of ergodic nilflows and unipotent affines,* Amer. J. Math. **93** (1971), 819–828.

22. ———, *Entropy and generators in ergodic theory,* Benjamin, New York, 1969. MR **41** #7071.

23. W. Parry and P. Walters, *Endomorphisms of a Lebesgue space*, Bull. Amer. Math. Soc. **78** (1972), 272–276.

24. M. Ratner, *Markov partitions for Anosov flows on n-dimensional manifolds*, Math. Systems (to appear).

25. A. Rényi, *Representations for real numbers and their ergodic properties,* Acta Math. Acad. Sci. Hungar. **8** (1957), 477–493. MR **20** #3843.

26. V. A. Rohlin, *Metric properties of endomorphisms of compact commutative groups,* Izv. Akad. Nauk SSSR Ser. Mat. **28** (1964), 867–874; English transl., Amer. Math. Soc. Transl. (2) **64** (1962), 244–252. MR **29** #5955.

27. ———, *Exact endomorphisms of a Lebesgue space,* Izv. Akad. Nauk SSSR Ser. Mat. **25** (1961), 499–530; English transl., Amer. Math. Soc. Transl. (2) **39** (1964), 1–36. MR **26** #1423.

28. ———, *Lectures on the entropy theory of transformations with invariant measure,* Uspehi Mat. Nauk **22** (1967), no. 5 (137), 3–56; Russian Math. Surveys **22** (1967), no. 5, 1–52. MR **36** #349.

29. V. A. Rohlin and Ja. G. Sinaĭ, *Construction and properties of invariant measurable partitions,* Dokl. Akad. Nauk SSSR **141** (1961), 1038–1041=Soviet Math. Dokl. **2** (1961), 1611–1614. MR **27** #2604.

30. S. Rudolfer, *The independence properties of certain number-theoretic endomorphisms,* Proc. Sympos. on Topology, Dynamics and Ergodic Theory, University of Kentucky, Lexington, Ky., 1971, 68–69.

31. Ja. G. Sinaĭ, *Weak isomorphism of transformations with invariant measure,* Mat. Sb. **63 (105)** (1964), 23–42; English transl., Amer. Math. Soc. Transl. (2) **57** (1966), 123–143. MR **28** #5164b.

32. ———, *Construction of Markov partitionings,* Funkcional. Anal. i Priložen. **2** (1968), no. 3, 70–80. (Russian) MR **40** #3591.

33. J. von Neumann, *Zur operatoren Methoden in der klassischen Mechanik,* Ann. of Math. **33** (1932), 587–642.

34. B. Weiss, *Intrinsically ergodic systems,* Bull. Amer. Math. Soc. **76** (1970), 1266–1269. MR **42** #1978.