COLLOQUIUM MATHEMATICUM

VOL. 84/85 2000 PART 2

RESIDUALITY OF DYNAMICAL MORPHISMS

BY

R. M. BURTON (CORVALLIS, OR), M. S. KEANE (AMSTERDAM) AND
JACEK SERAFIN (WROCŁAW)

In memoriam Anzelm Iwanik,
teacher, colleague, and friend

Abstract. We present a unified approach to the finite generator theorem of Krieger, the homomorphism theorem of Sinai and the isomorphism theorem of Ornstein. We show that in a suitable space of measures those measures which define isomorphisms or respectively homomorphisms form residual subsets.

1. Introduction. In 1977, Burton and Rothstein [2] put forward the idea that some of the basic results in ergodic theory could be obtained using "soft" methods, and more recently Kammeyer [6] has applied this technique to relative isomorphism theory. In 1996 a simplified exposition and clarification of the original method was proposed by Serafin [13] as a part of his dissertation. This article presents a further simplification and strengthening of the method, which we hope to be useful both for understanding and further development. It presents a unified approach to the finite generator theorem (Krieger [9]), the homomorphism (Sinai [14]) and isomorphism (Ornstein [11], Keane and Smorodinsky [8]) theorems for Bernoulli schemes. It is our conviction that, although none of the results presented are new, the unification will prove to be useful for further developments.

2. Preliminaries. We begin with the basic object of our investigation, a probability space

$$(Y, \mathcal{B}, \nu)$$

together with an automorphism T of this space, which we assume to be *ergodic* and to have *finite entropy*. As we shall be concerned with classification, we also assume that the probability space $(Y, \mathcal{B}, \nu)$ is standard, i.e. isomorphic to the unit interval with Lebesgue measure; this type of space is

2000 *Mathematics Subject Classification*: 37A05, 37A35.
Research of the third author supported by KBN grant 2 P03A 039 15 1998-2001.

commonly called a nonatomic Lebesgue space (see Rokhlin [10]). For basic notions of ergodic theory, which we use here without explanation, we refer to the excellent treatises of Billingsley [1] and Walters [15]. In this situation, it has been known for forty years (see Rokhlin [10]) that T possesses a countable generator; recently Keane and Serafin [7] have given an elementary exposition using standard methods in ergodic theory, making this result easily accessible and understandable. To simplify our present exposition, we use this result, which implies that we may assume that

$$Y = \{1, 2, 3, \ldots\}^{\mathbb{Z}}$$

is a sequence space provided with the product σ-algebra $\mathcal{B}$ and the left shift transformation T, the measure ν being then given by the collection of its values on the countable collection of cylinder sets in Y. This is not necessary for what follows, but makes many of our topological statements easy to verify.

The second object we shall need is a finite shift space, considered as a measurable space but on which a variety of measures will live:

$$X = \{1, \ldots, s\}^{\mathbb{Z}}$$

with the σ-algebra $\mathcal{A}$ generated by its coordinate mappings (product σ-algebra) and the left shift S. If μ is an ergodic S-invariant probability measure on $(X, \mathcal{A})$ and if the systems

$$(Y, \mathcal{B}, \nu, T) \quad \text{and} \quad (X, \mathcal{A}, \mu, S)$$

are isomorphic (homomorphic) via a mapping

$$\phi : Y \to X$$

which carries ν to μ and T to S, then there is a probability measure ξ on

$$(Z, \mathcal{C}) := (X \times Y, \mathcal{A} \times \mathcal{B})$$

invariant under $U = S \times T$ and ergodic, such that its projections ξ^X on X and ξ^Y on Y are μ and ν respectively, and which lives on (gives mass one to) the graph of ϕ. Equivalently, under ξ we have

$$\mathcal{A} \times \mathbf{2}_Y = \mathbf{2}_X \times \mathcal{B} \quad \text{or} \quad \mathcal{A} \times \mathbf{2}_Y \subseteq \mathbf{2}_X \times \mathcal{B},$$

where $\mathbf{2}_X$ and $\mathbf{2}_Y$ are the two-set σ-algebras $\{\emptyset, X\}$ and $\{\emptyset, Y\}$, and the equality holds mod ξ in the isomorphic case, while the inclusion holds mod ξ in the homomorphic case. It should be clear that the existence of an isomorphism (homomorphism) is equivalent to the existence of a measure ξ with the above properties; we call such ξ's also *isomorphisms* (*homomorphisms*).

The three basic theorems we want to prove can easily be interpreted using the above setting:

1. In the finite generator theorem, we are given $(Y, \mathcal{B}, \nu, T)$ with finite entropy $h_\nu(T)$, but no measure on $(X, \mathcal{A})$, and we wish to find ξ producing an isomorphism between $(X, \mathcal{A}, \xi^X, S)$ and $(Y, \mathcal{B}, \nu, T)$—any projection ξ^X will do for μ.

2. In the homomorphism theorem, we are given $(Y, \mathcal{B}, \nu, T)$ and a Bernoulli measure μ on $(X, \mathcal{A})$ with $h_\nu(T) \geq h_\mu(S)$, and we wish to find ξ producing a homomorphism such that $\xi^Y = \nu$ and $\xi^X = \mu$. Note that here we may assume that $h_\nu(T) = h_\mu(S)$ by splitting states in X, since amalgamating these states again is a homomorphism.

3. In the isomorphism theorem, we are given two Bernoulli measures μ and ν of equal entropy and need to find an isomorphism ξ with $\xi^X = \mu$ and $\xi^Y = \nu$.

Now we can explain, using the above, Burton's and Rothstein's insight which makes the finding of ξ as described relatively easy. The main difficulty is that many such ξ are possible, making it hard to define one particular ξ precisely. However, perhaps we can define a set of ξ's in such a way that "most" of the elements of this set are isomorphisms (homomorphisms), without having to point to one. For this, we need some suitable measure of largeness of sets, and they suggested that Baire category is suitable. So consider the probability measures on $(Z, \mathcal{C})$ as a metric space, using the product cylinder sets to define a metric in the usual manner giving rise to the weak topology, in which convergence of a sequence of measures corresponds to convergence of the values on every fixed cylinder set; denote this space by $\mathcal{M}$. Then $\mathcal{M}$ is clearly a compact separable metric space, and it is elementary to show (and we omit the proof) that

$$\mathcal{M}_0 := \{\xi \in \mathcal{M} : \xi \text{ is invariant and ergodic, } \xi^Y = \nu,\ h_{\xi^X}(S) \geq h_\nu(T)\}$$

and, if μ is a fixed ergodic measure on $(X, \mathcal{A}, S)$,

$$\mathcal{M}_1 := \{\xi \in \mathcal{M} : \xi \text{ is invariant and ergodic, } \xi^Y = \nu,\ \xi^X = \mu\}$$

are both Baire subsets of $\mathcal{M}$, i.e. possess the Baire property that countable intersections of (relatively) open dense subsets are dense (and, in particular, nonempty). In fact, all of the conditions in the definitions above are closed conditions (recall that entropy is an upper semicontinuous function of measures [5]) except the condition of ergodicity; the ergodic measures are the extreme points of the U-invariant measures, and a general theorem states that the extreme points of a compact convex set form a G_δ. Here one can also deduce the result from elementary considerations.

Now we can state the three theorems which have as immediate consequences the finite generator, homomorphism and isomorphism theorems, respectively.

THEOREM 1. *If* $h_\nu(T) < \log s$, *then*

$$\mathcal{M}_0^\star := \{\xi \in \mathcal{M}_0 : \xi \text{ is an isomorphism}\}$$

is a countable intersection of dense open subsets of $\mathcal{M}_0$.

THEOREM 2. *If* $h_\nu(T) = h_\mu(S)$ *and* μ *is Bernoulli, then*

$$\mathcal{M}_1^\star := \{\xi \in \mathcal{M}_1 : \xi \text{ is a homomorphism}\}$$

is a countable intersection of dense open subsets of $\mathcal{M}_1$.

THEOREM 3. *If* μ *and* ν *are Bernoulli and* $h_\nu(T) = h_\mu(S)$, *then*

$$\mathcal{M}_2^\star := \{\xi \in \mathcal{M}_1 : \xi \text{ is an isomorphism}\}$$

is a countable intersection of dense open subsets of $\mathcal{M}_1$.

Note that in each of the cases above, the sets considered are nonempty, as we see that $\xi = \mu \times \nu$ (with μ Bernoulli uniform in the finite generator case) belongs to the corresponding $\mathcal{M}_0$ or $\mathcal{M}_1$. Note also that Theorem 3 is a trivial consequence of Theorem 2, simply by applying it twice! The remainder of our exposition will be devoted to simple proofs of Theorems 1 and 2.

3. Proof of Theorem 1. First we write $\mathcal{M}_0^\star$ as a countable intersection of open sets $V_{k,l}$, $k, l \geq 1$, and then we show that each $V_{k,l}$ is dense in $\mathcal{M}_0$. If P is a finite partition of Z, then we denote by $\mathcal{P}$ the algebra generated by P and by $\mathcal{P}_U$ the σ-algebra generated by all $U^t P$, $t \in \mathbb{Z}$. Similar notation is used for partitions $Q, R, \ldots$ In particular, we define $P = \{P_i : 1 \leq i \leq s\}$ by

$$P_i := \{z = (x, y) \in Z : x_0 = i\},$$

and $Q^{(l)} := \{Q_1, \ldots, Q_{l-1}, Q_l^{(l)}\}$ by

$$Q_j := \{z = (x, y) \in Z : y_0 = j\} \quad \text{and} \quad Q_l^{(l)} := Z \setminus \bigcup_{j=1}^{l-1} Q_j;$$

finally, $Q := \{Q_1, Q_2, \ldots\}$.

In general for partitions R and R' we write

$$R' \subseteq_\varepsilon \mathcal{R}_T \bmod \xi$$

if for each set R'_i of the partition R' there exists a set $\overline{R} \in \mathcal{R}_T$ such that

$$\xi(\overline{R} \,\triangle\, R'_i) < \varepsilon.$$

Now we can define the set $V_{k,l}$ to be the collection of all $\xi \in \mathcal{M}_0$ such that

$$Q^{(l)} \subseteq_{\varepsilon_k} \mathcal{P}_U \bmod \xi \quad \text{and} \quad P \subseteq_{\varepsilon_k} \mathcal{Q}_U \bmod \xi,$$

where $\varepsilon_k = 1/k$. It is easy to check that $V_{k,l}$ is open; indeed,

$$P \subseteq_{\varepsilon_k} \mathcal{Q}_U \bmod \xi$$

if and only if

$$P \subseteq_{\varepsilon_k} \bigvee_{t=-n}^{n} U^t Q^{(l)} \bmod \xi,$$

for some n and l, which is an open condition on ξ, being a finite number of strict inequalities, and similarly for the other condition. Observe that the two approximate inclusions defining $V_{k,l}$ hold if $H_\xi(P|\mathcal{Q}_U) < \varepsilon'_k$ and $H_\xi(Q|\mathcal{P}_U) < \varepsilon'_k$, where ε'_k is a function of ε_k and $\varepsilon_k \to 0$ as $k \to \infty$ if and only if $\varepsilon'_k \to 0$ as $k \to \infty$. Moreover, clearly

$$\bigcap_{k,l} V_{k,l} = \mathcal{M}_0^\star,$$

since if ξ belongs to this intersection then both

$$Q \subseteq \mathcal{P}_U \bmod \xi \quad \text{and} \quad P \subseteq \mathcal{Q}_U \bmod \xi,$$

which says that ξ is an isomorphism. Thus it only remains to show that each $V_{k,l}$ is dense in $\mathcal{M}_0$. So now, since l is fixed, we drop the superscript l, denoting by Q simply the partition $Q = \{Q_j : 1 \leq j \leq l\}$ with

$$Q_j := \{z = (x, y) \in Z : y_0 = j\} \quad \text{if } j < l$$

and

$$Q_l := \{z = (x, y) \in Z : y_0 \geq l\}.$$

This amounts to replacing the symbols $\geq l$ in the alphabet of Y by l. Let $\xi \in \mathcal{M}_0$ be arbitrary; we have as conditions that

$$\xi^Y = \nu, \qquad h_{\xi^X}(S) \geq h_\nu(T),$$

and that ξ is U-invariant and ergodic. Also, the integers l and k are given; we want to show that there is a $\tilde{\xi} \in V_{k,l}$ as close as we want to ξ. It should be clear that the sets $V_{k,l}$ are decreasing in k and l; thus if we fix $n_0 \in \mathbb{N}$ and $\varepsilon > 0$, we desire to find $\tilde{\xi}$ which belongs to $V_{k,l}$ and

$$|\xi(A) - \tilde{\xi}(A)| < \varepsilon$$

for any atom A of $\bigvee_{t=0}^{n_0-1} U^t(P \vee Q)$. Here we may assume that k and l are fixed, but as large as we wish, the choice based upon ε and n_0. We need a number of elementary steps to accomplish our task, as follows:

STEP 1. The purpose of this step is to slightly raise the entropy $h_{\xi^X}(S)$ so that we have a strict inequality

$$h_{\xi^X}(S) > h_\nu(T).$$

At the end of the discussion, we need to do this once more. (For detailed computations see [13].) The basic idea is to perturb the x-values independently of ξ. That is, we replace ξ by ξ_1, where a realization $z^1 = (x^1, y^1)$ of ξ_1 is obtained by taking a realization $z = (x, y)$ of ξ, setting $y^1 = y$, and for each $t \in \mathbb{Z}$ flipping a coin with small success probability and, if successful, setting, say,

$$x_t^1 = x_t + 1 \pmod s;$$

otherwise retaining

$$x_t^1 = x_t.$$

It should be clear that then ξ_1^Y is still ν and

$$h_{\xi_1^X}(S) > h_{\xi_1^Y}(T),$$

and that if the success probability ε' is sufficiently small,

$$|\xi(A) - \xi_1(A)| < \varepsilon' < \varepsilon/2$$

for each atom A of $\bigvee_{t=0}^{n_0-1} U^t(P \vee Q)$; hence by replacing ε by $\varepsilon/2$ we may assume the strict inequality. A careful computation shows that in fact

$$h_{\xi_1^X}(S,P) \geq h_{\xi^X}(S,P) + \varepsilon'(\log s - h_{\xi^X}(S,P)).$$

We let $d := h_{\xi_1^X}(S) - h_\nu(T) > 0$.

Notice that it is not important to have formulae for this procedure, which can be accomplished in a multitude of ways, as long as ξ^X is not uniform Bernoulli.

STEP 2. Set

$$\delta = \min(d/8, \varepsilon/8),$$

with d and ε as above. Let $\delta_m = \xi^X([1^m])$, where 1^m is a block of m consecutive 1's, and $[B]$ is a cylinder set based on a block B. It is clear that $\delta_m \to 0$ as $m \to \infty$. Choose and fix m so large that $\delta_m < \delta/2$.

STEP 3. The purpose of this step is to choose an integer N sufficiently large so that the Shannon–McMillan–Breiman theorem for atoms of $\bigvee_{t=0}^{n-1} U^t P$, $\bigvee_{t=0}^{n-1} U^t Q$ and $\bigvee_{t=0}^{n-1} U^t(P \vee Q)$ is valid to accuracy δ, for $n \geq N$; we also need that the ergodic theorem for atoms of $\bigvee_{t=0}^{n_0-1} U^t P$, $\bigvee_{t=0}^{n_0-1} U^t Q$ and $\bigvee_{t=0}^{n_0-1} U^t(P \vee Q)$ is valid to accuracy δ, for $n \geq N$. Also, we need that the empirical frequency of the block 1^m in the X-process is correct with accuracy δ, for blocks of length $n \geq N$. For ease we omit the exact formulae at this point, and write the formulae as they occur in the sequel.

STEP 4. In this step we choose markers for both the X and Y sequences. Recall that one of our hypotheses is that $(Y, \mathcal{B}, \nu)$ is nonatomic. Hence there exist cylinder sets $[M]$ of arbitrary small measure, based on sequences M of successive symbols from $\{1, \dots, l\}$. Given such an M and a point $y \in Y$, we let $\tau_M(y)$ denote the distance between the two successive occurrences of M in y before and after the zero coordinate, and we choose M with $\nu([M])$ so small that

$$\nu(\{y : \tau_M(y) < N\}) < \delta.$$

For X-markers, we simply choose as marker the sequence 1^m, with m chosen as above.

STEP 5. For the convenience of the reader we now recall a version of the marriage lemma which we shall use in the sequel (see e.g. [4]). We start with two sets; we call elements of one set boys while elements of the other girls. Suppose that each boy knows (is acquainted with) a finite set of girls. The lemma states that it is possible for each boy to marry one of his acquaintances if and only if every finite set of n boys is collectively acquainted with a set of at least n girls. It is an immediate exercise to see that the latter condition holds if each boy knows at least K girls and each girl knows less than K boys, where K is some fixed positive number.

We now use the marriage lemma to make dictionaries necessary for the definition of $\widetilde{\xi}$. There will be one dictionary for each $t \geq N$, giving a 1-1 correspondence between x-words (girls) and y-words (boys) of length t. Fix $t \geq N$. Let B be the collection of all sequences b of length t of the Y-alphabet such that b begins with the marker M and

$$-\frac{1}{t}\log \xi^Y([b]) \leq h_{\xi^Y}(T) + \delta.$$

Let G' be the set of all sequences g' of length t of the X-alphabet such that

$$-\frac{1}{t}\log \xi^X([g']) \geq h_{\xi^X}(S) - \delta.$$

By the Shannon–McMillan–Breiman theorem, the above inequalities hold with probability at least $1-\delta$, for all $t \geq N$.

If now $g' \in G'$, we define $g := \Psi(g')$ by replacement of some (very few) symbols in g', as follows:

- If necessary, replace the first m symbols in g' by 1, so that g will start with a marker 1^m.
- Now destroy all other markers in g'; whenever 1^m occurs in g' except at the very beginning, replace the last symbol 1 by a symbol 2.

Then the mapping Ψ has collapsed some g''s into the same g; the maximal number of g''s which can give possibly the same g is bounded by

$$s^m + \sum_{i=1}^{j}\binom{t}{i}2^i,$$

where j is the number of occurrences of 1^m in g'. If now the frequency j/t of the marker occurrence differs from δ_m by no more than δ then it is an easy calculation to show that this number is exponentially small with respect to t, and if N were chosen so that

$$\max\left(\log s \cdot \frac{m}{N}, \frac{\log N}{N} + \delta_m + \delta\right) < 2\delta$$

then

$$s^m + \sum_{i=1}^{j} \binom{t}{i} 2^i < 2^{2\delta t} < 2^{dt/4}.$$

Now set $G = \{g = \Psi(g') : g' \in G'\}$. We define a relation (girl knows boy) as follows: $g \in G$ and $b \in B$ are *related* if there exists a $g' \in G'$ with $\Psi(g') = g$ such that the pair (g', b), considered as an atom of the partition

$$\bigvee_{i=0}^{t-1} U^i(P \vee Q),$$

satisfies the Shannon–McMillan–Breiman theorem for $P \vee Q$ as in step 3. That is, in particular,

$$2^{-t(h+\delta)} \le \xi([g] \times [b]) \le 2^{-t(h-\delta)},$$

where h denotes the mean entropy of $P \vee Q$ under U. From all this it is now clear that each boy knows at least

$$2^{t(h-h_{\xi^Y}(T)-4\delta)} \ge 2^{t(h-h_{\xi^Y}(T)-d/2)} =: K$$

girls, and each girl knows at most

$$2^{t(h-h_{\xi^X}(S)+2\delta)} \le 2^{t(h-h_{\xi^X}(S)+d/4)} < K$$

boys; therefore by the marriage lemma we can match each boy to a girl in a 1-1 fashion. This is our dictionary for length t.

STEP 6. We now construct a measure $\overline{\xi}$ with all properties required except for $h_{\overline{\xi}^X}(S) \ge h_\nu(T)$. This is very simple: to obtain a typical point $(\overline{x}, \overline{y})$ for $\overline{\xi}$, first choose a point (x, y) according to the measure ξ, and then set $\overline{y} = y$. To obtain $\overline{x}$, use the markers M occurring in y and the dictionaries to replace pieces of x. If the words do not occur in the dictionaries or if $\tau_M(y) < N$, then leave x unchanged in those coordinates. It is easy to see that the zero coordinate of a point x is coded using dictionaries with probability at least $1 - 2\delta$. It is then elementary to check that the desired properties hold.

STEP 7. In this step we raise the entropy of $\overline{\xi}^X$, without disturbing the other properties of $\overline{\xi}$. Clearly the two properties of the mean entropy

$$h(U, R) = h(U, R \vee T^{-1}R), \qquad h(U, R) - h(U, R') \le H(R|R')$$

together imply

$$|h(U, R) - h(U, R')| \le H(R|\mathcal{R}'_U) + H(R'|\mathcal{R}_U).$$

As a consequence we have

$$\begin{aligned} |h_{\overline{\xi}^X}(S, P) - h_{\overline{\xi}^Y}(T, Q)| &= |h_{\overline{\xi}}(U, P \times \mathbf{2}_Y) - h_{\overline{\xi}}(U, \mathbf{2}_X \times Q)| \\ &\le H_{\overline{\xi}}(P|\mathcal{Q}_U) + H_{\overline{\xi}}(Q|\mathcal{P}_U) < 2\varepsilon'_k. \end{aligned}$$

We desire here that ε'_k be so small that the above inequality implies

$$Q \subseteq_{\varepsilon_k} \mathcal{P}_U \bmod \overline{\xi}, \qquad P \subseteq_{\varepsilon_k} \mathcal{Q}_U \bmod \overline{\xi}.$$

Use now the same method as in step 1 to raise the entropy of the $\overline{\xi}^X$. This corrresponds to moving $\overline{\xi}$ to $\widetilde{\xi}$ (which is then our desired measure) by an amount that depends upon ε'_k, and which can be made smaller than $\varepsilon/2$ in the distribution distance. The upper semicontinuity of entropy then guarantees that $H_{\widetilde{\xi}}(P|\mathcal{Q}_U)$ and $H_{\widetilde{\xi}}(Q|\mathcal{P}_U)$ remain small and consequently $\widetilde{\xi} \in V_{k,l}$, the inclusions defining $V_{k,l}$ being open conditions. ∎

4. Proof of Theorem 2. Assume now that μ is a Bernoulli measure on $(X, \mathcal{A}, S)$ and P is an independent generating partition. Define

$$\mathcal{M}_{1,n} := \{\xi \in \mathcal{M}_1 : P \subseteq_{1/n} \mathcal{B} \bmod \xi\}.$$

As P generates we have

$$\mathcal{M}_1^\star = \bigcap_n \mathcal{M}_{1,n}.$$

Obviously the sets $\mathcal{M}_{1,n}$ are open so it suffices to show that they are dense. Fix $\xi \in \mathcal{M}_1$, $\varepsilon > 0$, and $n \geq 1$. We shall find $\widetilde{\xi} \in \mathcal{M}_{1,n}$ such that $d(\xi, \widetilde{\xi}) < \varepsilon$, where d is the usual metric inducing the weak topology on $\mathcal{M}$.

Let $\xi \in \mathcal{M}_1$. Observe that $\mathcal{M}_1$ is a subset of $\mathcal{M}_0$, so Theorem 1 implies the existence of a measure ξ_1 such that

$$d(\xi, \xi_1) < \varepsilon' < \varepsilon/2, \quad \xi_1^Y = \nu, \quad h_{\xi_1^X}(S) \geq h_\nu(T), \quad P \times \mathbf{2}_Y \subset \mathbf{2}_X \times \mathcal{B} \bmod \xi_1,$$

where ε' is to be determined later. It is easy to see that the last condition implies that in fact $h_{\xi_1^X}(S) = h_\nu(T) = h_\mu(S)$. Standard calculation shows that for every positive integer m and positive δ we can choose ε' small enough so the condition $d(\xi, \xi_1) < \varepsilon'$ implies $|\mathrm{dist}(\xi_1^X, \bigvee_{i=0}^{m-1} S^{-i}P) - \mathrm{dist}(\mu, \bigvee_{i=0}^{m-1} S^{-i}P)| < \delta$. At this point it is more convenient to have measures μ and ξ_1^X "living" on two separate spaces, so let us consider a copy $(\widetilde{X}, \widetilde{\mathcal{A}}, \widetilde{P}, \mu)$ of $(X, \mathcal{A}, P, \mu)$. As the process (S, P, μ) is independent and consequently finitely determined, it follows that $(S, \widetilde{P}, \mu)$ and (S, P, ξ_1^X) are close in the $\bar{d}$-metric. Equivalently (see [12]), there exists an ergodic $S \times S$-invariant ϱ which projects to ξ_1^X and μ such that

$$\varrho\Big(\bigcup_{i=1}^k \widetilde{P}_i \times P_i\Big) \geq 1 - \varepsilon_1 > 1 - \varepsilon/2,$$

with ε_1 small to be determined later. Since the measures ϱ and ξ_1 have a common factor ξ_1^X, we can use Furstenberg's construction ([3, pp. 110–115]) in order to find a common extension η on $\widetilde{X} \times X \times X \times Y$, an independent joining over a common factor ξ_1^X. First decompose ξ_1 and ϱ over a factor measure ξ_1^X. Let $A \times B \in \widetilde{\mathcal{A}} \times \mathcal{A}$ and $C \times D \in \mathcal{A} \times \mathcal{B}$. We have

$$\begin{aligned}\varrho(A \times B) &= \int_X \varrho_x(A \times B)\, d\xi_1^X(x) = \int_X \mathbf{E}_\varrho(\chi_{A\times B}|X)(x)\, d\xi_1^X(x) \\ &= \int_X \chi_B(x)\mathbf{E}_\varrho(\chi_{A\times X}|X)(x)\, d\xi_1^X(x)\end{aligned}$$

and similarly

$$\xi_1(C \times D) = \int_X \xi_{1,x}(C \times D)\, d\xi_1^X(x) = \int_X \chi_C(x)\mathbf{E}_{\xi_1}(\chi_{X\times D}|X)(x)\, d\xi_1^X(x).$$

The independent joining η is now given by

$$\eta(A \times B \times C \times D) = \int_X \chi_{B\cap C}(x)\mathbf{E}_\varrho(\chi_{A\times X}|X)(x)\mathbf{E}_{\xi_1}(\chi_{X\times D}|X)(x)\, d\xi_1^X(x).$$

We can define the projection ξ_2 of η to $\widetilde{X} \times Y$ by

$$\begin{aligned}\xi_2(A \times D) &= \eta(A \times X \times X \times D) \\ &= \int_X \mathbf{E}_\varrho(\chi_{A\times X}|X)(x)\mathbf{E}_{\xi_1}(\chi_{X\times D}|X)(x)\, d\xi_1^X(x).\end{aligned}$$

Finally, put ξ_2 on $X \times Y$ and call it $\widetilde{\xi}$. It is a straightforward computation that $\widetilde{\xi}^Y = \nu$ and $\widetilde{\xi}^X = \mu$. Moreover, the independent joining η can be constructed ergodic, so $\widetilde{\xi}$ as a factor measure is also ergodic. Let us now estimate $d(\xi_1, \widetilde{\xi})$. We have, for $A \times B \in \mathcal{A} \times \mathcal{B}$,

$$|\xi_1(A \times B) - \widetilde{\xi}(A \times B)| = |\xi_1(A \times B) - \xi_2(\widetilde{A} \times B)|,$$

where $\widetilde{A}$ is the $\widetilde{\mathcal{A}}$-copy of a set $A \in \mathcal{B}_X$. Comparing the decompositions of ξ_1 and ξ_2 we see that

$$\begin{aligned}|\xi_1(A \times B) - \xi_2(\widetilde{A} \times B)| &\le \int_X |\chi_A(x) - \mathbf{E}_\varrho(\chi_{\widetilde{A}\times X}|X)(x)|\, d\xi_1^X(x) \\ &= \int_X |\mathbf{E}_\varrho(\chi_{\widetilde{X}\times A}|X)(x) - \mathbf{E}_\varrho(\chi_{\widetilde{A}\times X}|X)(x)|\, d\xi_1^X(x) \\ &\le \int_X \mathbf{E}_\varrho(\chi_{\widetilde{X}\times A\triangle\widetilde{A}\times X}|X)(x)\, d\xi_1^X(x) \\ &= \int_{\widetilde{X}\times X} \chi_{\widetilde{X}\times A\triangle\widetilde{A}\times X}\, d\varrho \\ &= \varrho(\widetilde{X} \times A \triangle \widetilde{A} \times X) < \varepsilon_1,\end{aligned}$$

so the triangle inequality implies $d(\xi, \widetilde{\xi}) < \varepsilon$, as required. A judicious choice of ε_1 and upper semicontinuity of entropy guarantee that

$$H_{\widetilde{\xi}}(P \times \mathbf{2}_Y \,|\, \mathbf{2}_X \times \mathcal{B}) \le H_{\xi_1}(P \times \mathbf{2}_Y \,|\, \mathbf{2}_X \times \mathcal{B}) + 1/n = 1/n.$$

The proof is complete. ■

References

[1] P. Billingsley, *Ergodic Theory and Information*, Wiley, New York, 1965.

[2] R. Burton and A. Rothstein, *Isomorphism theorems in ergodic theory*, Technical Report, Oregon State Univ., 1977.

[3] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton Univ. Press, Princeton, NJ, 1981.

[4] P. Halmos and H. Vaughan, *The Marriage Problem*, Amer. J. Math. 72 (1950), 214–215.

[5] K. Jacobs, *Lecture Notes on Ergodic Theory*, Aarhus Univ., 1962.

[6] J. Kammeyer, *The isomorphism theorem for relatively finitely determined $\mathbb{Z}^n$-actions*, Israel J. Math. 69 (1990), 117–127.

[7] M. Keane and J. Serafin, *On the countable generator theorem*, Fund. Math. 157 (1998), 255–259.

[8] M. Keane and M. Smorodinsky, *Bernoulli schemes of the same entropy are finitarily isomorphic*, Ann. of Math. 109 (1979), 397–406.

[9] W. Krieger, *On entropy and generators of measure-preserving transformations*, Trans. Amer. Math. Soc. 149 (1970), 453–464.

[10] V. A. Rokhlin, *Lectures on the entropy theory of measure-preserving transformations*, Russian Math. Surveys 22 (1967), no. 5, 1–52.

[11] D. Ornstein, *Bernoulli shifts with the same entropy are isomorphic*, Adv. Math. 4 (1970), 337–352.

[12] —, *Ergodic Theory, Randomness and Dynamical Systems*, Yale Univ. Press, New Haven, 1974.

[13] J. Serafin, *Finitary codes and isomorphisms*, Ph.D. Thesis, Technische Universiteit Delft, 1996.

[14] Ya. Sinai, *A weak isomorphism of transformations with an invariant measure*, Dokl. Akad. Nauk SSSR 147 (1962), 797–800 (in Russian).

[15] P. Walters, *An Introduction to Ergodic Theory*, Springer, 1982.

Department of Mathematics
Oregon State University
Corvallis, OR 97331-4665
U.S.A.
E-mail: burton@math.orst.edu

Centre for Mathematics and Computer Science (CWI)
Post Office Box 94079
1090 GB Amsterdam, The Netherlands
E-mail: keane@cwi.nl

Institute of Mathematics
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
E-mail: serafin@banach.im.pwr.wroc.pl

Received 3 August 1999;
revised version 1 December 1999

(3806)