COLLOQUIUM MATHEMATICUM

VOL. 84/85 2000 PART 2

ISOMORPHIC RANDOM BERNOULLI SHIFTS

BY

V. M. GUNDLACH AND G. OCHS (BREMEN)

Dedicated to the memory of Anzelm Iwanik

Abstract. We develop a relative isomorphism theory for random Bernoulli shifts by showing that any random Bernoulli shifts are relatively isomorphic if and only if they have the same fibre entropy. This allows the identification of random Bernoulli shifts with standard Bernoulli shifts.

1. Introduction. Shift systems arise naturally in ergodic theory in the following two ways: as representations of stationary stochastic processes and via symbolic dynamics for smooth dynamical systems with hyperbolic properties. These two interpretations are combined in chaos theory to associate stochastic features to deterministic systems which can be modeled by shifts. In particular such an identification allows a classification of systems according to stochastic properties. The extreme case is the identification with Bernoulli shifts, which are models of independent and identically distributed stochastic processes. This makes the class of dynamical systems which are isomorphic to Bernoulli shifts especially interesting.

In order to show that Bernoulli shifts need not be isomorphic, Kolmogorov introduced the notion of entropy as an extremely successful isomorphy invariant. It gained even more importance when Ornstein [Orn70] showed that entropy is a complete invariant for Bernoulli shifts and even shifts with the so-called weak-Bernoulli property, i.e. that any such shifts with the same entropy are isomorphic. This result was extended by Thouvenot [Th75a] in a relative isomorphism theory for the so-called relative, conditional or fibre entropy to the case of factors of skew products.

In the theory of random dynamical systems random shifts as introduced in [BG92] arise naturally via symbolic dynamics for smooth systems with hyperbolic properties evolving under the influence of noise (see also [GK99] and [Gun99]). They can be seen as representations of stationary stochastic

2000 *Mathematics Subject Classification*: Primary 37A35, 37H05; Secondary 28D20, 60J99.

Key words and phrases: random dynamical system, random shift, fibre entropy, relative isomorphism.

processes in random environments. For the same reasons as in the deterministic situation random Bernoulli shifts, their classification and a suitable notion of entropy are of interest. Another interesting question arises from the fact that "deterministic" shifts are trivial examples of random shifts: when can one identify random and "deterministic" shift systems? The answer to this question is somehow amazing, but not unexpected: also for random Bernoulli shifts the fibre entropy is a complete invariant and hence allows the decoupling of noise and shift systems. For experts in the field of random dynamical systems it means a surprise, as they are used to new features in the classification of random systems, as the topological classification of hyperbolic linear random dynamical systems in [Con97] shows.

In our main result we will show explicitly that random Bernoulli shifts are (relatively) isomorphic if and only if they have the same entropy. This result could also be obtained by an application of an extension of [Lin77, Appendix] of the relative isomorphism theory of [Th75a], which guarantees that a property of the generator known as relatively very weak Bernoulli is sufficient for the fibre entropy to be a complete invariant. We prefer a more direct approach to pay tribute to the random features of our shift systems.

This paper is also a result of many discussions we had with J.-P. Thouvenot during several German–Polish Conferences on Dynamical Systems and Ergodic Theory. At these occasions we also experienced the organizing skills, determination and the enthusiasm, but most of all the kindness of Anzelm Iwanik.

2. The main result. Throughout this paper we fix a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ together with an invertible $\mathbb{P}$-preserving ergodic transformation ϑ. Consider a compact metric space X with Borel σ-algebra $\mathcal{B}$, and a set $E \subset \Omega \times X$ measurable with respect to the product σ-algebra $\mathcal{F} \otimes \mathcal{B}$ such that the fibres $E_\omega := \{x \in X : (\omega, x) \in E\}$, $\omega \in \Omega$, are compact. Then a continuous bundle random dynamical system (RDS) in time $\mathbb{Z}$ is generated by invertible mappings $f(\omega) : E_\omega \to E_{\vartheta\omega}$ with iterates given by

$$f(n,\omega) := \begin{cases} f(\vartheta^{n-1}\omega)\ldots f(\vartheta\omega)f(\omega) & \text{for } n \geq 1, \\ \mathrm{id} & \text{for } n = 0, \\ f(\vartheta^{n}\omega)^{-1}\ldots f(\vartheta^{-1}\omega)^{-1} & \text{for } n \leq -1, \end{cases}$$

for $n \in \mathbb{Z}$, $\omega \in \Omega$ such that $(\omega, x) \mapsto f(\omega)x$ is measurable and $x \mapsto f(\omega)x$ is continuous for $\mathbb{P}$-almost all ω. With the help of the skew product transformation $\Theta : E \to E$, $\Theta(\omega, x) = (\vartheta\omega, f(\omega)x)$, we call a probability measure μ on $(E, \mathcal{F} \otimes \mathcal{B} \cap E)$ *f-invariant* if it is invariant under Θ and has marginal $\mathbb{P}$ on Ω. Any such measure μ disintegrates via $d\mu(\omega, x) = d\mu_\omega(x)d\mathbb{P}(\omega)$ with disintegrations satisfying $f(\omega)\mu_\omega = \mu_{\vartheta\omega}$ $\mathbb{P}$-a.s. For these invariant measures μ we introduce the *fibre* (or *relative*) *entropy* of f with respect to μ according to [Kif86] as $h_\mu(f) = h_\mu(\Theta|(\mathrm{Pr}_\Omega|_E)^{-1}(\mathcal{F}))$ where the right-hand side is the conditional entropy of Θ with respect to $(\mathrm{Pr}_\Omega|_E)^{-1}(\mathcal{F})$ and $\mathrm{Pr}_\Omega|_E$ is the natural projection from E to Ω. Equivalently one can define $h_\mu(f)$ using countable partitions $P = \{A_i\}$ of X into measurable sets A_i as

$$h_\mu(f) = \sup_P h_\mu(f, P) = \sup_P \lim_{n\to\infty} H_{\mu_\omega}\Big(\bigvee_{i=0}^{n-1} f(i,\omega)^{-1} P_{\vartheta^i\omega}\Big) \quad \mathbb{P}\text{-a.s.}$$

where $H_{\mu_\omega}(P_\omega)$ denotes the entropy of the partition $\{A_i \cap E_\omega\}$ of E_ω and the supremum is taken over all partitions P such that $\int H_{\mu_\omega}(P_\omega)\, d\mathbb{P}(\omega) < \infty$. This representation of fibre entropy can be found in [Bog93] together with a version of the Kolmogorov–Sinai Theorem which states that for partitions $\mathcal{P}$ satisfying $\bigvee_{i=-\infty}^{\infty} f(i,\omega)^{-1} P_{\vartheta^i\omega} = \mathcal{B} \cap E_\omega$, so-called generators, one has $h_\mu(f) = h_\mu(f, P)$.

We will be mainly interested in the case where $X := \prod_{i=-\infty}^{\infty} \overline{\mathbb{Z}}^+$ where $\overline{\mathbb{Z}}^+ = \mathbb{Z}^+ \cup \{\infty\}$ denotes the one-point compactification of $\mathbb{Z}^+ = \{1, 2, \dots\}$ and X is compact in the product topology and metrizable. We denote the elements of X by $x = (x_i)$. The continuous mapping $\sigma : X \to X$ defined by $(\sigma x)_i = x_{i+1}$ is called the (left) *shift*. Let k denote a $\mathbb{Z}^+$-valued random variable and put

$$\Sigma_k(\omega) := \{x \in X : x_i \le k(\vartheta^i\omega) \text{ for all } i \in \mathbb{Z}\} = \prod_{i=-\infty}^{\infty} \{1, \dots, k(\vartheta^i\omega)\}.$$

Then $\{\sigma : \Sigma_k(\omega) \to \Sigma_k(\vartheta\omega) : \omega \in \Omega\}$ determines a bundle RDS known as *random k-shift* (cf. [Gun99]). If $\int \log k\, d\mathbb{P} < \infty$, then the partition of X into one-cylinders $\{x \in X : x_i = k\}_{k\in\mathbb{Z}}$ yields a generator and the random version of the Sinai–Kolmogorov Theorem can be applied.

For the random k-shift we consider an invariant measure μ induced by a random probability vector $p = \{p(\omega) = (p_i(\omega)) \in [0,1]^{k(\omega)} : \omega \in \Omega\}$ via disintegrations on cylinder sets as

$$\mu_\omega(\{x \in \Sigma_k(\omega) : x_i = a_i \text{ for } i = -n, \dots, n\}) = \prod_{i=-n}^{n} p_{a_i}(\vartheta^i\omega)$$

for any $n \in \mathbb{N}$ and $a_i \in \{1, \dots, k(\vartheta^i\omega)\}$. The resulting dynamical system (Σ_k, σ, μ) is called the *random p-Bernoulli shift*. Its entropy is given by $h_\mu(\sigma) = -\int \sum_{i=1}^{k(\omega)} p_i(\omega) \log p_i(\omega)\, d\mathbb{P}(\omega)$.

In the following we will be concerned with the classification of random Bernoulli shifts via fibre entropy. We will call two RDS over the same abstract dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, \vartheta)$ *isomorphic* if the induced skew products are isomorphic relative to ϑ.

2.1. THEOREM. *Let (Σ_k, σ) with $\int \log k\, d\mathbb{P} < \infty$ be a random p-Bernoulli shift with corresponding σ-invariant measure μ and entropy $h := h_\mu(\sigma)$. Assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is complete. Then there exist*

• *a probability vector* $q = (q_1, \ldots, q_n)$ *for some* $n \in \mathbb{N}$, *and a corresponding Bernoulli shift* $(\{1, \ldots, n\}^{\mathbb{Z}}, \sigma)$ *with corresponding* σ*-invariant measure* $\nu = q^{\mathbb{Z}}$,

• *a measurable isomorphism* $\Phi : (\Sigma_k, \mu) \to (\Omega \times \{1, \ldots, n\}^{\mathbb{Z}}, \mathbb{P} \times \nu)$ *such that*

$$h_\nu(\sigma) = h, \quad \Phi \circ \Theta = \Theta \circ \Phi, \quad \mathrm{Pr}_\Omega \circ \Phi = \mathrm{Pr}_\Omega.$$

Φ *is a fibrewise homeomorphism in the following sense*: *Write* $\Phi(\omega, x) = (\omega, \phi_\omega(x))$. *There exist* σ*-invariant measurable subsets* $A_1 \subset \Sigma_k$ *and* $A_2 \subset \Omega \times \{1, \ldots, n\}$ *with* $\mu(A_1) = (\mathbb{P} \times \nu)(A_2) = 1$ *such that* ϕ_ω *is a homeomorphism between* $A_1(\omega)$ *and* $A_2(\omega)$, *where* $A_i(\omega) := \{x : (\omega, x) \in A_i\}$.

3. Structure of the proof. In the proof we will allow infinitely many symbols. Every random probability vector $p = (p_i(\omega))_{i \in \mathbb{N}}$ defines a σ-invariant random product measure μ_p on $\Sigma := \Omega \times X$.

We will prove the following: Let p be any random probability vector such that $\mu = \mu_p$ has finite entropy $h_\mu(\sigma) = -\int_\Omega \sum_{i=1}^\infty p_i(\omega) \log p_i(\omega)\, d\mathbb{P}(\omega)$. A sufficient condition for finite entropy is that μ is concentrated on Σ_k for some random variable $k : \Omega \to \mathbb{N}$ with $\int \log k\, d\mathbb{P} < \infty$. We will also assume $h_\mu(\sigma) > 0$. (If $h_\mu(\sigma) = 0$ then for a.e. ω there exists $i = i(\omega)$ with $p_i(\omega) = 1$. In this case a random permutation of symbols yields an isomorphism to a deterministic system.)

Then we will construct an isomorphism Φ between (Σ, μ_p) and (Σ, μ_q), where q is a random probability vector with the following properties:

(i) $h_{\mu_p}(\sigma) = h_{\mu_q}(\sigma)$,

(ii) there exists a deterministic n with $q_i(\omega) = 0$ for $i > n$,

(iii) q_3 is independent of ω and lies in the open interval $]0, 1[$,

(iv) if there exists an i such that $p_i(\omega) \equiv p_i \in \,]0, 1[$ is independent of ω, then all q_j are independent of ω.

Properties (iii) and (iv) show that a twice repeated application of our construction yields an isomorphism between an arbitrary random Bernoulli shift with finite entropy and a deterministic Bernoulli shift.

The construction of the isomorphism Φ consists of several steps. In Section 4 we define certain symbols to be "markers", which will be fixed under the isomorphism. The construction of the isomorphisms depends on the positions of the symbols relative to the markers. This ensures that the construction is shift invariant, which would not be the case if the construction was based on fixed coordinates.

In the next step (Section 5) we construct a decreasing sequence of equivalence relations $R_r(\omega) \subset X \times X$, which partitions X into equivalence classes of comparable measure. This construction relies on the fact that the measure of cylinder sets can asymptotically be estimated using the entropy.

Section 6 contains the construction of the isomorphism. First we define a probability vector q which coincides with p on the set of markers and which has at least one deterministic component. The space X is also partitioned into equivalence classes with respect to μ_q. For every natural number r we construct a correspondence between the equivalence classes with respect to μ_p and the equivalence classes with respect to μ_q. This correspondence defines a relation $R_r(\omega) \subset X \times X$. Some properties of the R_r (listed in Lemma 6.1) will ensure that $\bigcap_{r\geq 1} R_r(\omega)$ is essentially a measure preserving one-to-one relation, which gives the desired isomorphism.

The idea of the proof follows a deterministic proof of Keane and Smorodinsky (see [CFS82, Chapter 10, §7]). In particular, in the investigation of the combinatorial properties of the correspondences between equivalence classes we can use most of the arguments of the deterministic proof. However, the random situation requires some new ideas. The main difficulty here are nonuniformities which arise when the p_i are not bounded away from 0 or 1. Most of the work which is needed to handle these nonuniformities is done in Section 4.

4. Markers and skeletons

4.1. *Definition of the markers.* We start with a random probability vector $p = (p_i(\omega))_{i\in\mathbb{N}}$ and set $\mu = \mu_p$. Without loss of generality we can assume (due to a permutation of symbols) that

$$\int p_1(\omega)\log p_1(\omega)\,d\mathbb{P}(\omega) < 0.$$

Hence there exists $\varepsilon > 0$ such that

$$\mathbb{P}\{\omega \in \Omega : p_1(\omega) \in [\varepsilon, 1-\varepsilon]\} > 0.$$

In the case where $p_i(\omega) \equiv p_i \in \,]0,1[$ is independent of ω for some i we assume $i = 1$. The symbol 1 will play the special role of the marker for the symbols in the infinite sequence. We pick a "good" set $A \in \mathcal{F}$ such that

$$\mathbb{P}(A) > 0 \quad \text{and} \quad \varepsilon \le p_1(\omega) \le 1-\varepsilon \quad \text{for every } \omega \in A.$$

In case p_1 is deterministic choose $A = \Omega$, else choose A such that $-\int_{A^c} p_1(\omega)\log p_1(\omega)d\mathbb{P} > 0$.

For $\omega \in A$ the marker 1 will not be changed under our isomorphism Φ, i.e. if $x_k = 1$ and $\vartheta^k\omega \in A$, then the kth coordinate of $\phi_\omega(x)$ will be equal to 1.

4.2. *Entropy "relative to the markers".* Here we consider the "dynamics relative to the fixed markers". Set

$$\widetilde{p_i}(\omega) := \begin{cases} p_i(\omega) & \text{if } \omega \notin A,\\ p_i(\omega)/(1-p_1(\omega)) & \text{if } \omega \in A,\ i \neq 1,\\ 0 & \text{if } \omega \in A,\ i = 1,\end{cases}$$

and

$$F(\omega,x) := \begin{cases} -\log \widetilde{p}_{x_0}(\omega) & \text{if } \widetilde{p}_{x_0}(\omega) > 0,\\ 0 & \text{if } \widetilde{p}_{x_0}(\omega) = 0. \end{cases}$$

Then $h_0 := \int F d\mu_p$ can be interpreted as the relative entropy produced by the symbols $\notin A \times \{1\}$. We have $h_0 < h < \infty$. Since

$$h_0 \geq -\int_{A^c} p_1(\omega)\log p_1(\omega)\, d\mathbb{P}(\omega) - \sum_{i=2}^{\infty}\int_A \frac{p_i(\omega)}{1-p_1(\omega)}\log\frac{p_i(\omega)}{1-p_1(\omega)}\, d\mathbb{P}(\omega),$$

$h_0 = 0$ implies by the choice of A that $p_1(\omega) = p_1$ is deterministic and for almost every ω there exists some $i = i(\omega) > 1$ with $p_i(\omega) = 1 - p_1$. But in this case there exists a trivial isomorphism to a deterministic Bernoulli shift. Hence we can assume $h_0 > 0$.

Our next goal is to write F as a pointwise limit of an increasing sequence of bounded measurable functions. For this purpose set

$$J_m(\omega) := \{i \in \mathbb{N} : \widetilde{p}_i(\omega) \geq 1/m\}$$

for $m \geq 1$, $\omega \in \Omega$ and choose $i_m = i_m(\omega)$ minimal with

$$\widetilde{p}_{i_m}(\omega) := \min_{i\in J_m(\omega)} \widetilde{p}_i(\omega).$$

Define

$$\widetilde{p}_i(m,\omega) := \begin{cases} \widetilde{p}_i(\omega) & \text{for } i \in J_m(\omega)\setminus\{i_m\},\\ 1-\sum_{i\in J_m(\omega)\setminus\{i_m\}}\widetilde{p}_i(\omega) & \text{for } i = i_m,\\ 0 & \text{otherwise,} \end{cases}$$

$$\widetilde{F}_m(\omega,x) := \begin{cases} -\log \widetilde{p}_{x_0}(m,\omega) & \text{for } x_0 \in J_m(\omega),\\ 0 & \text{otherwise.} \end{cases}$$

Then $\lim_{m\to\infty}\widetilde{p}_i(m,\omega) = \widetilde{p}_i(\omega)$ for all $\omega \in \Omega$, $i \in \mathbb{N}$ and $\widetilde{F}_{m+1}(\omega,x) \geq \widetilde{F}_m(\omega,x)$, i.e. $\widetilde{F}_m \nearrow F$ for $m \to \infty$, and

$$\lim_{m\to\infty}\int \widetilde{F}_m\, d\mu = \int F\, d\mu = h_0.$$

For $r \in \mathbb{N}$ choose M_r such that

$$\int \widetilde{F}_{M_r}\, d\mu \geq (1-2^{-(r+1)})h_0$$

and set $F_r := \widetilde{F}_{M_r}$. We note that

$$F_r \leq \log M_r \text{ and } F = \sup_r F_r = \lim_{r\to\infty} F_r. \tag{1}$$

4.3. *Construction of "skeletons".* Define a random variable $Z : \Sigma \to \mathbb{N}$ and random subsets Z_n, $n \in \mathbb{N}$, of Σ as follows: for $(\omega,x) \in \Sigma$ we choose $k \geq 0$ minimal such that $\vartheta^k\omega \in A$, $x_k = 1$ and set

$$Z(\omega,x) = \begin{cases} \operatorname{card}\{j : 0 < j < k \text{ and } \vartheta^j\omega \in A\ (\Rightarrow x_j \neq 1)\} & \text{if } \omega \in A,\ x_0 = 1,\\ 0 & \text{otherwise,} \end{cases}$$

$$Z_n := \{(\omega,x) : Z(\omega,x) \geq n\}.$$

Thus we have determined sets of (ω,x) with $\omega \in A$, $x_0 = 1$, and the next n symbols corresponding to a noise realization in the good set A are different from 1. We have $Z_{n+1} \subset Z_n$, $\mu(\bigcap_{n\geq 0} Z_n) = 0$ and $\mu(Z_n) \geq \varepsilon^{n+1}\mathbb{P}(A) > 0$ by the choice of A.

For $n \in \mathbb{N}$ set

$$\begin{aligned}\widetilde{a}_n(\omega,x) &= \max\{j \leq 0 : Z(\Theta^j(\omega,x)) \geq n\},\\ \widetilde{b}_n(\omega,x) &= \min\{j > 0 : Z(\Theta^j(\omega,x)) \geq n\},\\ \widetilde{\ell}_n(\omega,x) &= \widetilde{b}_n(\omega,x) - \widetilde{a}_n(\omega,x),\end{aligned}$$

which are $\mathbb{P}$-a.s. well defined and yield random variables describing length of words which contain a word corresponding to Z_n. We have

$$\widetilde{a}_n(\omega,x) \searrow -\infty \quad \text{and} \quad \widetilde{b}_n(\omega,x), \widetilde{\ell}_n(\omega,x) \nearrow +\infty. \tag{2}$$

If $\widetilde{a}_n(\omega,x) \leq j < \widetilde{b}_n(\omega,x)$, then $\widetilde{\ell}_n(\Theta^j(\omega,x)) = \widetilde{\ell}_n(\omega,x)$, $\widetilde{a}_n(\Theta^j(\omega,x)) = \widetilde{a}_n(\omega,x) - j$, and $\widetilde{b}_n(\Theta^j(\omega,x)) = \widetilde{b}_n(\omega,x) - j$.

By Birkhoff's Ergodic Theorem, for $r \in \mathbb{N}$,

$$\lim_{n\to\infty} \frac{1}{\widetilde{\ell}_n(\omega,x)} \sum_{j=\widetilde{a}_n(\omega,x)}^{\widetilde{b}_n(\omega,x)-1} F_r(\Theta^j(\omega,x)) = \inf F_r d\mu \geq (1-2^{-r})h_0 \qquad \mu\text{-a.s.}$$

Hence we can find $0 < N_1 < N_2 < \ldots$ such that

$$\frac{N_r}{r2^r} h_0 \geq \log M_r \tag{3}$$

and

$$\mu\left\{(\omega,x) : \frac{1}{\ell_r(\omega,x)} \sum_{j=a_r(\omega,x)}^{b_r(\omega,x)-1} F_r(\Theta^j(\omega,x)) \leq (1-2^{-r})h_0\right\} \leq 2^{-r}, \tag{4}$$

where $a_r(\omega,x) := \widetilde{a}_{N_r}(\omega,x)$, $b_r(\omega,x) := \widetilde{b}_{N_r}(\omega,x)$ and $\ell_r(\omega,x) := \widetilde{\ell}_{N_r}(\omega,x)$.

Let S be the set of all finite words from the alphabet $\{*, 1, \star\}$. We define a so-called *r-skeleton* for (ω,x) by

$$s_r : \Sigma \to \Omega \times S, \quad (\omega,x) \mapsto (\vartheta^{a_r(\omega,x)}\omega, \tau_{a_r}(\omega,x), \tau_{a_r+1}(\omega,x), \ldots, \tau_{b_r-1}(\omega,x)),$$

where

$$\tau_j(\omega,x) = \begin{cases} * & \text{if } \vartheta^j\omega \in A^{\mathrm{c}},\\ 1 & \text{if } \vartheta^j\omega \in A,\ x_j = 1,\\ \star & \text{if } \vartheta^j\omega \in A,\ x_j \neq 1.\end{cases}$$

An r-skeleton is characterized by the property that it can be turned into an allowable word by replacing the symbols $*$ and $\star$ by some symbols from $\{2, 3, \ldots\}$ or $\{1, 2, \ldots\}$, respectively. Such a word is called a *filler* for the skeleton. Note that r-skeletons consist of consecutive $(r-1)$-skeletons.

5. Partitions into equivalence classes

5.1. *Equivalence relations.* For $(\omega, x) \in \Sigma$ and $r \in \mathbb{N}$ set

$$D_r(\omega, x) := \{(\omega, y) : s_r(\omega, x) = s_r(\omega, y)\}.$$

We will now construct equivalence relations on the sets of fillers for $s_r(\omega, x)$, which define partitions of $D_r(\omega, x)$ into equivalence classes.

For $(\omega, x) \in \Sigma$ choose $k = k(\omega, x) < b_1(\omega, x)$ maximal with

$$\sum_{j=a_1(\omega,x)}^{k} F_1(\omega, x) \le -\log M_1 + \frac{1}{2} h_0 \ell_1(\omega, x)$$

which is possible because $\frac{1}{2} h_0 \ell_1(\omega, x) \ge \frac{1}{2} h_0 N_1 \ge \log M_1$ by (3). Now define two relations $\stackrel{1}{\sim}$ and $\stackrel{1}{\smile}$ by

$$\begin{aligned}(\omega, x) \stackrel{1}{\sim} (\omega, y) \Leftrightarrow s_1(\omega, x) &= s_1(\omega, y),\ k(\omega, x) = k(\omega, y), \\ x_j &\stackrel{1}{\leftrightarrow} y_j \text{ for } a_1(\omega, x) \le j \le k(\omega, x), \\ (\omega, x) \stackrel{1}{\smile} (\omega, y) \Leftrightarrow (\omega, x) &\stackrel{1}{\sim} (\omega, y),\end{aligned}$$

where

$$x_i \stackrel{r}{\leftrightarrow} y_i \;\Leftrightarrow\; x_i = y_i \text{ or } \{x_i, y_i\} \cap (J_{M_r}(\vartheta^i \omega) \setminus \{i_{M_r}(\vartheta^i \omega)\}) = \emptyset$$

for $r \in \mathbb{N}$. For $a_1(\omega, x) \le j < b_1(\omega, x)$ we have

$$(\omega, x) \stackrel{1}{\sim} (\omega, y) \;\Leftrightarrow\; \Theta^j(\omega, x) \stackrel{1}{\sim} \Theta^j(\omega, y).$$

Define

$$G_1(\omega, x) := \begin{cases} F_1(\omega, x) & \text{for } k(\omega, x) \ge 0, \\ 0 & \text{for } k(\omega, x) < 0. \end{cases}$$

Note that $G_1(\Theta^j(\omega, x)) = 0$ for $k(\omega, x) < j < b(\omega, x)$. Inductively we will now construct suitable random variables $G_r(\omega, x)$ satisfying $G_{r-1} \le G_r \le F_r$ and

$$\sum_{j=a_r(\omega,x)}^{b_r(\omega,x)-1} G_r(\Theta^j(\omega, x)) \le (1 - 2^{-r}) h_0 \ell_r(\omega, x) - \log M_r. \tag{5}$$

For $r = 1$ this is clear, as

$$\begin{aligned}\sum_{j=a_1(\omega,x)}^{b_1(\omega,x)-1} G_1(\Theta^j(\omega, x)) &= \sum_{j=a_1(\omega,x)}^{k(\omega,x)} F_1(\Theta^j(\omega, x)) \\ &\le -\log M_1 + \frac{h_0}{2} \ell_1(\omega, x) \\ &= h_0 \ell_1(\omega, x) \left(1 - \frac{1}{2}\right) - \log M_1.\end{aligned}$$

Now consider $r > 1$ and assume that G_{r-1} has been constructed. Then choose $k_r = k_r(\omega, x) < b_r(\omega, x)$ maximal with

$$\sum_{j=a_r(\omega,x)}^{k_r} F_r(\Theta^j(\omega, x)) + \sum_{j=k_r}^{b_r(\omega,x)-1} G_{r-1}(\Theta^j(\omega, x)) \le -\log M_r + h_0(1 - 2^{-r})\ell_r(\omega, x)$$

and define

$$G_r(\omega, x) = \begin{cases} F_r(\omega, x) & \text{for } k_r(\omega, x) \ge 0, \\ G_{r-1}(\omega, x) & \text{for } k_r(\omega, x) < 0. \end{cases}$$

Then

$$\sum_{j=a_r(\omega,x)}^{b_r(\omega,x)-1} G_r(\Theta^j(\omega, x)) = \sum_{j=a_r(\omega,x)}^{k_r(\omega,x)} F_r(\Theta^j(\omega, x)) + \sum_{j=k_r(\omega,x)+1}^{b_r(\omega,x)-1} G_{r-1}(\Theta^j(\omega, x)) \le -\log M_r + h_0(1 - 2^{-r})\ell_r(\omega, x). \tag{6}$$

Note that this construction is possible, as with the help of (3),

$$\sum_{j=a_r(\omega,x)}^{b_r(\omega,x)-1} G_{r-1}(\Theta^j(\omega, x)) \le (1 - 2^{-(r-1)})h_0\ell_r(\omega, x) \le (1 - 2^{-r})h_0\ell_r(\omega, x) - \log M_r.$$

Define

$$G(\omega, x) := \sup_r G_r(\omega, x) = \lim_{r\to\infty} G_r(\omega, x).$$

By (1), (4) and the choice of $k_r(\omega, x)$,

$$\mu\bigg\{(\omega, x) : \frac{1}{\ell_r(\omega, x)} \sum_{j=a_r(\omega,x)}^{b_r(\omega,x)-1} G_r(\Theta^j(\omega, x)) \ge (1 - 2^{-r})h_0 - \frac{2\log M_r}{\ell_r(\omega, x)}\bigg\} \ge 1 - 2^{-r}.$$

We can deduce that

$$\int G_r\, d\mu = \int \frac{1}{\ell_r(\omega, x)} \sum_{j=a_r(\omega,x)}^{b_r(\omega,x)-1} G_r(\Theta^j(\omega, x))\, d\mu(\omega, x),$$

as $\sum G_r$ is constant for $\Theta^{-a_r(\omega,x)}(\omega, x), \ldots, \Theta^{b_r(\omega,x)-1}(\omega, x)$. Therefore

$$\int G_r\, d\mu \ge (1 - 2^{-r})\bigg[(1 - 2^{-r})h_0 - \frac{2\log M_r}{N_r}\bigg].$$

The right-hand side of this inequality tends to $h_0 = \int F\, d\mu$ as $r \to \infty$, since

$$\frac{2\log M_r}{N_r} \le \frac{h_0}{r2^{r-1}}$$

by (3). This implies

$$G = \sup G_r = F \quad \mu\text{-a.s.} \tag{7}$$

We introduce the following relations inductively for $r > 1$:

$$\begin{aligned}(\omega, x) \overset{r}{\smile} (\omega, y) :\Leftrightarrow\ & s_r(\omega, x) = s_r(\omega, y),\\ & \Theta^j(\omega, x_j) \overset{r-1}{\sim} \Theta^j(\omega, y_j) \text{ for } a_r(\omega, x) \le j < b_r(\omega, x),\\ (\omega, x) \overset{r}{\sim} (\omega, y) :\Leftrightarrow\ & (\omega, x) \overset{r}{\smile} (\omega, y),\\ & x_j \overset{r}{\leftrightarrow} y_j \text{ for } a_r(\omega, x) \le j \le k_r(\omega, x).\end{aligned}$$

Clearly the relation $\overset{r}{\sim}$ is finer than the relation $\overset{r}{\smile}$ and thus the equivalence classes of $\overset{r}{\smile}$ are unions of equivalence classes of $\overset{r}{\sim}$.

5.2. *Measure of the equivalence classes.* Define

$$\begin{aligned}E_r(\omega, x) &:= \{y \in X : (\omega, x) \overset{r}{\sim} (\omega, y)\},\\ E'_r(\omega, x) &:= \{y \in X : (\omega, x) \overset{r}{\smile} (\omega, y)\}.\end{aligned}$$

5.1. LEMMA. *We have*

$$\mu_\omega(E_r(\omega, x)) \ge \mu_\omega(D_r(\omega, x)) \exp(-h_0(1 - 2^{-r})\ell_r(\omega, x)), \tag{8}$$

$$\mu_\omega(E'_r(\omega, x)) \ge \mu_\omega(D(\omega, x)) \exp(-h_0(1 - 2^{-r+1})\ell_r(\omega, x)). \tag{9}$$

Furthermore, there exists a set $B_r \subset \Sigma$ with $\mu(B_r) \ge 1 - 2^{-r}$ such that for all $(\omega, x) \in B_r$,

$$\mu_\omega(E_r(\omega, x)) \le \mu_\omega(D_r(\omega, x)) M_r^2 \exp(-h_0(1 - 2^{-r})\ell_r(\omega, x)). \tag{10}$$

Proof. By definition of G_r and $\overset{r}{\sim}$,

$$\mu_\omega(E_r(\omega, x)) = \mu_\omega(D_r(\omega, x)) \exp\Big(-\sum_{j=a_r(\omega,x)}^{b_r(\omega,x)-1} G_r(\Theta^j(\omega, x))\Big). \tag{11}$$

Together with (5) this implies (8). For B_r we choose the complement of the set appearing in (4). By (6) we have $k_r(\omega, x) \le b_r(\omega, x) - 3$ on this set, hence by (1) and the definition of G_r,

$$\begin{aligned}&\sum_{j=a_r(\omega,x)}^{b_r(\omega,x)-1} G_r(\Theta^j(\omega, x))\\ &\quad= \sum_{j=a_r(\omega,x)}^{b_r(\omega,x)-1} F_r(\Theta^j(\omega, x)) + \sum_{j=k_r(\omega,x)+1}^{b_r(\omega,x)-1} (G_{r-1}(\Theta^j(\omega, x)) - F_r(\Theta^j(\omega, x)))\\ &\quad\ge \sum_{j=a_r(\omega,x)}^{b_r(\omega,x)-1} F_r(\Theta^j(\omega, x)) - 2\log M_r > (1 - 2^{-r}) h_0^{\ell_r(\omega,x)} - 2\log M_r\end{aligned}$$

by (4). Thus we obtain (10) with the help of (11).

To prove (9) observe that $E'_r(\omega, x)$ (if $r > 1$) is the intersection of $\overset{r-1}{\sim}$ equivalence classes E'_i of length ℓ_i such that $\sum_i \ell_i = \ell_r(\omega, x)$. It follows from

(8) that

$$\mu_\omega(E'_r(\omega,x)) = \prod_i \mu_\omega(E'_i) \ge \mu_\omega(D_r(\omega,x)) \prod_i e^{-h_0(1-2^{-r+1})\ell_i}$$
$$= \mu_\omega(D_r(\omega,x)) e^{-h_0(1-2^{-r+1})\ell_r(\omega,x)}. \ \blacksquare$$

6. Construction of the isomorphism

6.1. *The probability vector q.* With $a \in [0,1]$ define

$$q_1(\omega) = p_1(\omega) \text{ for } \omega \in A, \quad q_1(\omega) = 0 \text{ for } \omega \in A^c,$$
$$q'_3(a,\omega) = a\varepsilon, \quad q'_2(a,\omega) = 1 - q_1(\omega) - q'_3(a,\omega).$$

Note that

$$a \mapsto a\varepsilon \log a\varepsilon + \mathbb{P}(A)(1-a\varepsilon)\log(1-a\varepsilon) + \int_{A^c} (1-q_1(\omega)-a\varepsilon) \log \frac{1-q_1(\omega)-a\varepsilon}{1-q_1(\omega)}\, d\mathbb{P}(\omega) = \widetilde{h}^{(1)}(a)$$

is continuous in a and $\lim_{a\to 0} \widetilde{h}^{(1)}(a) = 0$. If $\widetilde{h}^{(1)}(1) \ge h_0$, then by the intermediate value theorem there exists an $a_0 \in (0,1]$ such that $\widetilde{h}^{(1)}(a_0) = h_0$ and we can consider

$$q_2(\omega) = q'_2(a_0,\omega), \quad q_3(\omega) := q'_3(a_0,\omega) = a_0\varepsilon.$$

Otherwise we set $q_2 := q_2(1)$, $\widehat{h} = \widetilde{h}^{(1)}(1) - \varepsilon\log\varepsilon$ and choose $q_3, \ldots, q_n$ such that $q_3 + \ldots + q_n = \varepsilon$, q_i does not depend on ω, and

$$\sum_{i=3}^n q_i \log q_i = h_0 - \widehat{h}.$$

For $\nu := \mu_q$ we construct skeletons and equivalence classes of fillers in the same way as for μ above. For a.e. $(\omega,x) \in \Sigma$ we have

$$\nu_\omega(D_r(\omega,x)) = \mu_\omega(D_r(\omega,x)) = p_1(\vartheta^{a_r(\omega,x)}\omega) \prod_{j\in J}(1 - p_1(\vartheta^j\omega))$$

with $J = \{j : a_r(\omega,x) < j < b_r(\omega,x),\ \vartheta^j\omega \in A\}$.

$D_r(\omega,x)$ splits into finitely many $\overset{r}{\sim}$ equivalence classes $E_1^{(r)},\ldots,E_s^{(r)}$ corresponding to μ and analogously into equivalence classes $\widetilde{E}_1^{(r)},\ldots,\widetilde{E}_t^{(r)}$ corresponding to ν instead of μ in the construction of the equivalence relation $\overset{r}{\sim}$.

6.2. *Relations on $X \times X$*

6.1. Lemma. *For $\mathbb{P}$-almost all $\omega \in \Omega$ there exists a sequence $(R_r(\omega))_{r\in\mathbb{N}}$ of subsets $R_r(\omega) \subset X \times X$ which is decreasing in the sense that $R_{r+1}(\omega) \subset R_r(\omega)$ and has the following properties*:

(i) $(x,y) \in R_r(\omega) \Rightarrow s_r(\omega,x) = s_r(\omega,y)$.

(ii) *If r is even, then $R_r(\omega)\cap(D_r(\omega,x)\times D_r(\omega,x))$ is a* (*finite*) *union of products of equivalence classes, i.e. if $(\omega,x)\overset{r}{\sim}(\omega,x')$ and $(\omega,y)\overset{r}{\smile}(\omega,y')$, then $(x,y)\in R_r(\omega)\Leftrightarrow(x',y')\in R_r(\omega)$. If r is odd, the roles of $\overset{r}{\sim}$ and $\overset{r}{\smile}$ are interchanged.*

(iii) *For fixed $x\in X$ and every n-cylinder $C\subset X$ the set $\{\omega : R_r(\omega)\cap(\{x\}\times C)\neq\emptyset\}$ is measurable.*

(iv) *$R_r(\omega)$ is* correct *in the sense that $\nu_\omega(R_r(\omega,B))\geq\mu_\omega(B)$ for every measurable subset B of X, where $R_r(\omega,B)=\{y\in X:\exists x\in B$ such that $(x,y)\in R_r(\omega)\}$.*

(v) $(x,y)\in R_r(\omega)\Leftrightarrow(\sigma^j x,\sigma^j y)\in R_r(\vartheta^j\omega)$ *for* $a_r(\omega,x)\leq j<b_r(\omega,x)$.

(vi) *$R_r(\omega)$ is minimal in the sense that if $S\subset R_r(\omega)$ satisfies the conditions* (i)–(v), *then $S=R_r(\omega)$.*

Proof. We construct $(R_r(\omega))$ explicitly. Start with $R_1^{(0)}(\omega):=\{(x,y): s_1(\omega,x)=s_1(\omega,y)\}$. There is a natural ordering on the set of fillers for $s_r(\omega,x)$, which induces an ordering on the equivalence classes. We go through all pairs $(E_i^{(1)},\widetilde{E}_j^{(1)})$ in order and put

$$R_1^{(k+1)}(\omega):=\begin{cases}R_1^{(k)}(\omega)\setminus(E_i^{(1)}\times\widetilde{E}_j^{(1)}) & \text{if this defines a correct subset,}\\ R_1^{(k)}(\omega) & \text{otherwise.}\end{cases}$$

Then $R_1(\omega):=R_1^{(pq)}(\omega)$ is a correct and minimal subset satisfying all the properties (i)–(vi).

We are now in an ω-wise situation described in [CFS82, Chapter 10, §7] so that we can take over some combinatorial results from there.

Suppose $r>1$ is even. Recall that $s_r(\omega,x)$ is a sequence of $(r-1)$-skeletons

$$s_{r-1}(\Theta^{j_1}(\omega,x))s_{r-1}(\Theta^{j_2}(\omega,x))\dots s_{r-1}(\Theta^{j_k}(\omega,x)).$$

Each of these skeletons defines a partition of X in equivalence classes, where the $\overset{r}{\smile}$ equivalence classes are exactly the sections of $\overset{r-1}{\sim}$ equivalence classes with respect to $s_{r-1}(\Theta^{j_\ell}(\omega,x))$. Now assume that $R_{r-1}(\omega)$ with the properties (i)–(vi) exists. We can define a new subset

$$\begin{aligned}R_r^{(0)}(\omega):=\{(x,y): s_r(\omega,x)=s_r(\omega,y),\\ (\sigma^{j_l}x,\sigma^{j_l}y)\in R_{r-1}(\vartheta^{j_l}\omega)\text{ for }1\leq l\leq k\}\\ \subset R_{r-1}(\omega).\end{aligned}$$

This $R_r^{(0)}(\omega)$ has all the properties (i)–(v):

- (i) and (v) follow from the definition.
- (ii) holds, as $(\omega,x)\overset{r}{\sim}(\omega,y)\Rightarrow(\omega,x)\overset{r}{\smile}(\omega,y)\Rightarrow(\omega,x)\overset{r-1}{\sim}(\omega,y)$.
- (iv) follows from [CFS82, Chapter 10, §7, Lemma 7].

• (iii) is true, as $R_r(\omega)$ is a countable union of products of cylinder sets depending measurably on ω.

$R_r^{(0)}(\omega)$ is not necessarily minimal, but by the same procedure we have already applied to obtain $R_1(\omega)$ from $R_1^{(0)}(\omega)$ we can construct a subset $R_r(\omega)$ which keeps all properties (i)–(v) and is minimal in addition. Namely we can construct $R_r(\omega)$ by going in order through all pairs of $\overset{r}{\sim}$ (with respect to μ) and $\overset{r}{\smile}$ (with respect to ν) equivalence classes respectively. We remove all pairs which are not needed for correctness. If r is odd, we can proceed analogously, interchanging the roles of $\overset{r}{\sim}$ and $\overset{r}{\smile}$. ∎

6.3. *A one-to-one relation.* For $r \in \mathbb{N}$ and $E \subset \Sigma$ put

$$R_r(E) := \{(\omega, y) : \exists (\omega, x) \in E \text{ with } (x, y) \in R_r(\omega)\},$$
$$R_r^{-1}(E) := \{(\omega, x) : \exists (\omega, y) \in E \text{ with } (x, y) \in R_r(\omega)\}$$

and set

$$R := \bigcap_{r=1}^{\infty} R_r.$$

Using the completeness of $(\Omega, \mathcal{F}, \mathbb{P})$ it is possible to deduce from Lemma 6.1(iii) the measurability of $R_r(E)$ and $R_r^{-1}(E)$ and thus of $R(E)$ and $R^{-1}(E)$ for any measurable set $E \subset \Sigma$.

By (2) and Lemma 6.1(v) there exist subsets $S_1, T_1 \subset \Sigma$ with $\mu(S_1) = \nu(T_1) = 1$ such that

$$((\omega, x), (\omega, y)) \in R \Leftrightarrow (\Theta^j(\omega, x), \Theta^j(\omega, y)) \in R$$

whenever $j \in \mathbb{Z}$ and $(\omega, x) \in S_1$, $(\omega, y) \in T_1$.

6.2. Lemma. *There exist shift invariant measurable sets $S \subset S_1$, $T \subset T_1$ with $\mu(S) = \nu(T) = 1$ such that R restricted to $S \times T$ is a one-to-one relation.*

Proof. For even $r \in \mathbb{N}$ set

$$K_r := \{(\omega, x) : R_r(\omega, x) \text{ consists of one } \overset{r}{\smile} \text{ equivalence class}\},$$

and

$$L_r := \{(\omega, x) \in K_r : y_0 \text{ is uniquely determined in the equivalence class } R_r(\omega, x)\}.$$

Fix $D_r(\omega, x)$. Let $n_r(\omega, x)$ be the number of $\overset{r}{\sim}$ equivalence classes (with respect to μ) $E_r \subset D_r(\omega, x)$ for which there exist at least two $\overset{r}{\smile}$ equivalence classes $\widetilde{E}'_r \neq \widetilde{E}''_r$ (with respect to ν) with $E_r \times \widetilde{E}'_r, E_r \times \widetilde{E}''_r \subset R_r(\omega)$. By the construction of $R_r(\omega)$ and [CFS82, Chapter 10, §7, Lemma 6], $n_r(\omega, x)$ is bounded from above by the total number of $\overset{r}{\sim}$ equivalence classes in $D_r(\omega, x)$

with respect to ν_ω. From (9), $n_r(\omega,x) \le \exp(h_0(1-2^{-(r-1)})\ell_r(\omega,x))$. Defining

$$K_r' := \{(\omega,x) \in \Sigma \setminus G_r : \mu_\omega(E_r(\omega,x)) \le M_r^2 e^{-h_0(1-2^{-r})\ell_r(\omega,x)}\mu_\omega(D_r(\omega,x))\}$$

we thus have

$$\begin{aligned}
&\mu_\omega(D_r(\omega,x)\cap K_r') \\
&\quad \le \mu_\omega(D_r(\omega,x))\exp(h_0(1-2^{-(r-1)})\ell_r(\omega,x))M_r^2\exp(-h_0(1-2^{-r})\ell_r(\omega,x)) \\
&\quad = \mu_\omega(D_r(\omega,x))M_r^2\exp(-h_0 2^{-r}\ell_r(\omega,x)) \\
&\quad \le \mu_\omega(D_r(\omega,x))M_r^2\exp(-h_0 2^{-r}N_r),
\end{aligned}$$

which shows that

$$\mu(K_r') \le M_r^2\exp(-h_0 2^{-r}N_r) \le M_r^2 M_r^{-r} \to 0 \quad \text{as } r\to\infty.$$

On the other hand, we consider

$$K_r'' := \{(\omega,x)\in\Sigma : \mu_\omega(E_r(\omega,x)) > M_r^2\exp(-h_0(1-2^{-r})\ell_r(\omega,x))\}.$$

By (10), $\mu(K_r'') \le 2^{-r} \to 0$ as $r\to\infty$. Therefore we deduce from $K_r \supset \Sigma\setminus(K_r'\cup K_r'')$ that

$$\lim_{r\to\infty}\mu(K_r) = 1.$$

In order to deduce an analogous result for L_r we consider

$$\begin{aligned}
\widetilde{L}_r := \{(\omega,y) : {}& y_0 \text{ uniquely defined} \\
&\text{in the } \overset{r}{\smile} \text{ equivalence class with respect to } \nu\}.
\end{aligned}$$

If we define functions $\widetilde{F}$ and $\widetilde{G}_r$ analogously to F and G_r with p_i replaced by q_i, then

$$(\omega,y)\in\widetilde{L}_r \Leftrightarrow \widetilde{G}_{r-1}(\omega,y) = \widetilde{F}(\omega,y) \ne 0,$$

which implies via (7) that $\lim_{r\to\infty}\nu(\widetilde{L}_r) = 1$. From Lemma 6.1(iv), for $L_r' := R_r^{-1}(\widetilde{L}_r)$ it follows that $\lim_{r\to\infty}\mu(L_r') \ge \lim_{r\to\infty}\nu(\widetilde{L}_r) = 1$. Since $L_r = K_r \cap L_r'$ we have

$$\lim_{r\to\infty}\mu(L_r) = 1.$$

Define $L := \bigcup_{r=1}^\infty L_{2r}\cap S_1$. The set $S_2 := \bigcap_{j=-\infty}^\infty \Theta^j L$ is shift invariant with $\mu(S_2) = 1$. For $(\omega,x)\in S_2$ there exists at most one $y\in X$ with $((\omega,x),(\omega,y))\in R$.

Analogously we can find a shift invariant set $T_2\subset T_1$ with $\nu(T_2) = 1$ such that for $(\omega,y)\in T_2$ there exists at most one $x\in X$ with $((\omega,x),(\omega,y))\in R$. Set

$$\begin{aligned}
S &:= \{(\omega,x)\in S_2 : \exists(\omega,y)\in T_2 \text{ with } ((\omega,x),(\omega,y))\in R\},\\
T &:= \{(\omega,y)\in T_2 : \exists(\omega,x)\in S_2 \text{ with } ((\omega,x),(\omega,y))\in R\}.
\end{aligned}$$

Clearly $R(S) \subset T$ and $R^{-1}(T) \subset S$. Therefore R is a one-to-one relation on $S \times T$. We have

$$1 = \mu(S_2) = \nu(T_2) \le \lim_{r\to\infty} \mu(R_r^{-1}(T_2)) = \mu(R_r^{-1}(T_2)),$$

which implies

$$\mu(S) = \mu(S_2 \cap R_r^{-1}(T_2)) = 1.$$

Analogously $\nu(T) = 1$. ∎

6.4. *The isomorphism Φ.* Now we are prepared to finish the proof. Define $\Phi : S \to T$ via

$$\{\Phi(\omega, x)\} = R(\{(\omega, x)\}) \quad \text{and} \quad y = \phi_\omega(x) \Leftrightarrow (\omega, y) = \Phi(\omega, x).$$

From 6.3 it is clear that Φ is measurable and commutes with Θ, i.e.

$$\phi_{\theta\omega} \circ \sigma = \sigma \circ \phi_\omega.$$

For $E \subset S$ we have (using $\nu(T^{\mathrm{c}}) = 1$)

$$\begin{aligned}\nu(\Phi(E)) &= \nu(T \cap R(E)) = \nu(R(E)) \\ &= \nu\Big(\bigcap_{r=1}^{\infty} R_r(E)\Big) = \lim_{r\to\infty} \nu(R_r(E)) \ge \mu(E),\end{aligned}$$

and analogously $\mu(\Phi^{-1}(E)) \ge \nu(E)$ for $E \subset T$, i.e. Φ is measure preserving.

It remains to show the continuity of

$$\phi_\omega : S(\omega) := \{x : (\omega, x) \in S\} \to T(\omega) := \{y : (\omega, y) \in T\}$$

and of ϕ_ω^{-1}. Let

$$C := \{x \in X : x_{-n} = a_{-n}, \ldots, x_n = a_n\}$$

be a cylinder set and $x \in X$ with $\phi_\omega(x) \in C$. There exists an even $r \in \mathbb{N}$ with $\Theta^i(\omega, x) \in L_r$ for $|i| \le n$ (L_r from the proof of Lemma 6.2), i.e. the coordinates $(\phi_\omega(x))_i$ for $|i| \le n$ are uniquely determined by the $\overset{r}{\smile}$ equivalence classes $E'_r(\Theta^i(\omega, x))$ for $-n \le i \le n$. Hence

$$S(\omega) \cap \Big(\bigcap_{i=-n}^{n} \Theta^{-i} E_r(\Theta^i(\omega, x))\Big) \subset \phi_\omega^{-1}(C),$$

which means that $\phi_\omega^{-1}(C)$ contains an open neighbourhood of x in $S(\omega)$. The continuity of ϕ_ω^{-1} follows in the same way.

7. Remarks. Our main result states in particular that one can find relative isomorphisms which are able to decouple the noise from the shift in random shifts. This means that qualitatively a random Bernoulli shift shows the same behaviour as a "deterministic" Bernoulli shift, at least in the ergodic theoretical sense. Though random Bernoulli shifts do not generate any new dynamical phenomena, they are useful in investigations of RDS, e.g. those generating random fractal structures (cf. [Kif96]), as they draw a real-time picture of the dynamics in contrast to other symbolic descriptions which might be obtained via relative isomorphisms.

If we recall from Ornstein's theory (see for example [Orn70], [Pet83] or [CFS82]) that all two-sided Bernoulli shifts of the same entropy are isomorphic, we have the following amendment to Theorem 2.1.

7.1. COROLLARY. *Any two random Bernoulli shifts* $(\Sigma_{k_1}, \sigma, \mu_1)$ *and* $(\Sigma_{k_2}, \sigma, \mu_2)$ *of the same entropy and with* $\log k_i \in \mathbb{L}^1(\Omega, \mathbb{P})$, $i = 1, 2$, *are relatively isomorphic.*

This result can even be extended to a wider class of random shifts and RDS. As pointed out by Thouvenot such a work can make use of an extension in [Lin77, Appendix] of the relative isomorphism theory of [Th75a]. There, the same notion of relativized or fibre entropy is used together with the assumption that $\mathcal{F}$ is countably generated by a partition H. Then it is proved that if there exists a finite partition P of the bundle E or $\Omega \times X$ that has the property known as H-conditionally finitely determined, then the corresponding RDS, if it has finite fibre entropy, is relatively isomorphic to a random Bernoulli shift of the same entropy. In fact, the condition of a finite partition P can be weakened to countable partitions with finite entropy, as it is only used to obtain a version of the Shannon–McMillan–Breiman Theorem, which in our situation is provided by [Bog93, Theorem 2.2.5].

The notion of a partition to be H-conditionally finitely determined is rather abstract and difficult to check. Fortunately it could be shown by [Th75b] that it is implied by the so-called very weak Bernoulli property of a partition; these notions are in fact equivalent, as shown in [Rah78] (see also [Kie84]). We will use only a restriction of that notion, the weak Bernoulli property of a partition (see e.g. [Shi77]) with an ω-wise representation according to [Kie84].

7.2. DEFINITION. Let $\mathcal{P}$, $\mathcal{Q}$ be two partitions of a measurable bundle E such that $\mathcal{P}(\omega)$ and $\mathcal{Q}(\omega)$ are finite partitions of E_ω $\mathbb{P}$-a.s. They are called (ω, ε)-*independent* with respect to a probability measure μ for some $\varepsilon > 0$ if

$$\sum_{P \in \mathcal{P}(\omega),\ Q \in \mathcal{Q}(\omega)} |\mu_\omega(P \cap Q) - \mu_\omega(P)\mu_\omega(Q)| < \varepsilon.$$

The partition $\mathcal{P}$ is called *weak Bernoulli* with respect to a bundle RDS φ and its φ-invariant measure μ if for every $\varepsilon > 0$ and $\mathbb{P}$-a.a. $\omega \in \Omega$ there exists an $N = N(\omega, \varepsilon)$ such that the partitions $\bigvee_{i=0}^{n-1} \varphi(i, \omega)^{-1} \mathcal{P}(\vartheta^i \omega)$ and $\bigvee_{i=0}^{r-1} \varphi(i + t, \omega)^{-1} \mathcal{P}(\vartheta^{i+t} \omega)$ are (ω, ε)-independent for all $n \geq 0$, $r \geq 0$, $t \geq s + N$.

Obviously random Bernoulli shifts and also random Markov shifts with aperiodic matrices P (cf. [Gun99]) have the weak Bernoulli property with respect to the partition in one-cylinders. Thus the already mentioned result of Thouvenot and Lind, which we can cite as follows for our situation, applies to them.

THEOREM 7.3 (Thouvenot–Lind). *Assume that $\mathcal{F}$ is countably generated. Then any two random k-shifts $(\Sigma_{k_1}, \sigma, \mu_1)$, $(\Sigma_{k_2}, \sigma, \mu_2)$ which satisfy the weak Bernoulli property and have the same entropy are (relatively) isomorphic. In particular they are isomorphic to a random Bernoulli shift with this entropy.*

Let us mention once more that such a classification result allows the decoupling of the noise process described by the abstract dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, \vartheta)$ and the shift dynamics. The price one has to pay for this decoupling is the loss of information, e.g. on time scales, due to the relative isomorphism. Moreover this result shows implicitly that an interesting class of random shifts does not generate any new dynamical features. Nevertheless it is useful in analyzing smooth RDS with hyperbolic properties.

REFERENCES

[Bog93] T. Bogenschütz, *Equilibrium states for random dynamical systems*, PhD thesis, Universität Bremen, 1993.

[BG92] T. Bogenschütz and V. M. Gundlach, *Symbolic dynamics for expanding random dynamical systems*, Random Comput. Dynamics 1 (1992), 219–227.

[Con97] N. D. Cong, *Topological Dynamics of Random Dynamical Systems*, Clarendon Press, Oxford, 1997.

[CFS82] I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinai, *Ergodic Theory*, Grundlehren Math. Wiss. 245, Springer, New York, 1982.

[Gun99] V. M. Gundlach, *Random shifts and matrix products*, Habilitationsschrift, 1999.

[GK99] V. M. Gundlach and Y. Kifer, *Random hyperbolic systems*, in: H. Crauel and V. M. Gundlach (eds.), Stochastic Dynamics, Berlin, Springer, 1999, 117–145.

[Kie84] J. Kieffer, *A simple development of the Thouvenot relative isomorphism theory*, Ann. Probab. 12 (1984), 204–211.

[Kif86] Y. Kifer, *Ergodic Theory of Random Transformations*, Birkhäuser, Boston, 1986.

[Kif96] —, *Fractal dimensions and random transformations*, Trans. Amer. Math. Soc. 348 (1996), 2003–2038.

[Lin77] D. A. Lind, *The structure of skew products with ergodic group automorphisms*, Israel J. Math. 28 (1977), 205–248.

[Orn70] D. Ornstein, *Bernoulli shifts with the same entropy are isomorphic*, Adv. Math. 4 (1970), 337–352.

[Pet83] K. Petersen, *Ergodic Theory*, Cambridge Univ. Press, Cambridge, 1983.

[Rah78] M. Rahe, *Relatively finitely determined implies relatively very weak Bernoulli*, Canad. J. Math. 30 (1978), 531–548.

[Shi77] P. Shields, *Weak and very weak Bernoulli partitions*, Monatsh. Math. 84 (1977), 133–142.

[Th75a] J.-P. Thouvenot, *Quelques propriétés des systèmes dynamiques qui se décomposent en un produit de deux systèmes dont l'un est un schéma de Bernoulli*, Israel J. Math. 21 (1975), 177–207.

[Th75b] —, *Remarques sur les systèmes dynamiques donnés avec plusieurs facteurs*, ibid., 215–232.

Institut für Dynamische Systeme
Universität Bremen, Postfach 330 440
28334 Bremen, Germany
E-mail: gundlach@math.uni-bremen.de
gunter@math.uni-bremen.de

Received 11 August 1999;
revised 1 February 2000 (3813)