The Annals of Probability
1992, Vol. 20, No. 1, 397–402

EXPLICIT CODES FOR SOME INFINITE ENTROPY BERNOULLI SHIFTS

By Steven Kalikow and Benjamin Weiss

University of Southern California and Hebrew University of Jerusalem

Some explicit isomorphisms are constructed between Bernoulli shifts with infinite entropy.

The basic distribution with infinite entropy is the uniform distribution on [0, 1]. Conceptually simpler distributions are discrete distributions $\{p_1, p_2, \dots, p_k, \dots\}$ with $\Sigma - p_k \log p_k = +\infty$. Ornstein (1970) has shown that the processes obtained by having independent identically distributed (i.i.d.) random variables with these distributions or any other having infinite entropy are isomorphic. Represent the discrete-valued process by $\{x_n\}$, $n \in \mathbb{Z}$, which are independent $\mathbb{N}$-valued random variables with distribution $\{p_1, p_2, \dots\}$. The isomorphism is given by a function which commutes with the shift

$$f\colon \mathbb{N}^{\mathbb{Z}} \to [0,1]^{\mathbb{Z}},$$

whose coordinates $\{f_n\}$, $n \in \mathbb{Z}$ satisfy the following:

1. f_0 is uniformly distributed on $[0, 1]$;
2. $\{f_n\}$, $n \in \mathbb{Z}$, are independent;
3. the map $\{x_n\} \to \{f_n\}$ is almost surely $1:1$.

The proof exhibits f_0 as some complicated limit of finite codes and gives no easy method of explicitly calculating f_0. Our main purpose in this note is to give an explicit description of such an f_0 for some family of distributions $\{p_k\}$. The interest here is in the existence of some explicit coding between continuous i.i.d. random variables and discrete ones. There are two ideas involved. The first suffices to give f with properties 1 and 2 above. In order to get invertibility, property 3, we elaborate slightly on the classic example of Meshalkin (1959).

In the second part of the paper, we will give an explicit isomorphism between the time one map of a Poisson point process on the line and $[0,1]^{\mathbb{Z}}$ with the shift. Although it does not directly give an imbedding of a finite entropy Bernoulli shift in a flow, nonetheless, when combined with the first part of the paper, it gives an imbedding of a discrete Bernoulli shift in a Bernoulli flow.

Received November 1989; revised December 1990.
AMS 1980 *subject classifications*. Primary 28D05; secondary 28D10, 60G10.
Key words and phrases. Bernoulli shifts, infinite entropy, isomorphisms.

1. From discrete to continuous. The distributions $\{p_k\}_1^\infty$ that we deal with here have the following special form. There is a random length L with distribution of

$$\mathrm{Prob}(L = l_i) = \pi_i, \qquad 1 \le i < \infty,$$

and then conditioned on the length l_i, we choose a binary sequence of length l_i with all binary sequences of length l_i equally likely with probability 2^{-l_i}. Explicitly the p_k's may be defined by

$$p_k = 2^{-l_i}\pi_i \quad \text{for} \sum_{j=1}^{i-1} 2^{l_j} < k \le \sum_{j=1}^{i} 2^{l_j}.$$

Furthermore, we assume that

$$\sum_1^\infty l_i \pi_i = +\infty, \tag{1}$$

which is equivalent to infinite entropy.

A random variable x drawn from such a distribution will be written

$$x = x(1)x(2)\cdots x(m)\cdots,$$

where $x(i)$ is either a 0 or a 1 for $i \le L$ and *undefined* otherwise and L is one of the integers $l_1, l_2, \ldots$ and so on.

The condition (1) is equivalent to the condition that if $\{L_1, L_2, \ldots\}$ are i.i.d. random variables with distribution $\{\pi_k\}_1^\infty$, then with probability 1, $L_n \ge n$ infinitely many times. This is an immediate consequence of the Borel–Cantelli lemma and a standard summation by parts. If we are presented with a typical sequence of values $\{x_n\}_{-\infty}^\infty$ drawn independently from the distribution $\{p_k\}$ as above, then with probability 1 for each index n the preceding condition is valid for the associated $L_n, L_{n+1}, \ldots$. Keeping this in mind, the definition of f_0 is as follows. It actually depends only on $\{x_1, x_2, \ldots\}$.

Form the infinite sequence

$$x_1(1), x_2(2), x_3(3), \ldots, x_n(n), \ldots.$$

By the above, with probability 1, there are infinitely many terms in this sequence that are actually defined with a value 0 or 1. Delete all undefined terms, that is, whenever $L_i < i$, delete $x_i(i)$ from the above sequence. Call the new sequence

$$\bar{x}(1)\bar{x}(2)\cdots\bar{x}(i)\cdots$$

and define

$$f_0(x_1, x_2, \ldots) = \sum_1^\infty \bar{x}(i) 2^{-i}.$$

In case there are only finitely many terms that are defined set $f = 0$. When f_0

is applied to a shift of the sequence $(x_n, x_{n+1}, \ldots)$ the preliminary sequence

$$x_n(1)x_{n+1}(2)x_{n+2}(3) \cdots x_{n+i}(i+1) \cdots$$

is totally disjoint from all other shifts. This means that even conditioned on a fixed sequence of values for the lengths L_n, the resulting f_n's will be independent random variables uniformly distributed on the unit interval.

This completes the first part of our project and gives a stationary coding from independent discrete random variables to independent continuous random variables. The mapping that we have defined is clearly onto but is not invertible. Indeed without knowing the lengths, the problem of inverting the mapping is completely hopeless. We get around this problem by doing a separate coding for the lengths.

2. Coding the lengths. For the rest of this discussion we will fix the length probabilities π_k to be 2^{-k} for all $k \geq 1$. Meshalkin (1959) showed how to give an explicit coding between the Bernoulli shifts $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ and $(\frac{1}{2}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8})$. A brief description of his code may be found in Weiss (1972). Iterating that construction gives a simple explicit code between the Bernoulli shifts $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ and $(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots)$. Choose for the lengths $l_1, l_2, \ldots$ any strictly increasing sequence of integers such that

$$\sum_{1}^{\infty} l_k 2^{-k} = +\infty.$$

For this the length probabilities will have to be of a special form. The first two digits of the final f_0 will now be given by the code that maps the length sequence to two binary digits. Then the rest of the digits will be given by the map f_0 previously described. In this way the coding defined by f_0 will be invertible, since given $f_0(x_n, x_{n+1}, \ldots)$ for all n will enable us to recover, from the first two digits, the entire length distribution L_n. Then, given the length distribution, the map is trivially invertible.

To keep the paper self-contained, we now give in more detail the code between $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ and $(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots)$. For this we represent the integers $1, 2, 3, \ldots$ as follows: $1, 01, 001, 0001, 00001, \ldots$. If an integer is chosen according to the latter distribution, then the first digit is equiprobably 1 or 0, and the second digit, *given that it exists*, is again equiprobably 1 or zero and so on. In general, if we condition on the event that the nth digit is there, then it is equiprobably equal to 0 or 1. It is a little easier to explain the mapping going from the $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ shift to the (2^{-k}) shift, so we do so. Given a bi-infinite sequence of random variables $\{u_n(1)u_n(2)\}_{-\infty}^{\infty}$ that are each 0 or 1 independently with probability $\frac{1}{2}$, we wish to form a bi-infinite sequence of integers as above. To do so we first group the $u_n(1)$'s in pairs as follows:

Each $u_n(1)$ whose value is 1 is paired with the first $u_m(1)$ with $m > n$ that is equal to a zero such that the block $u_n(1)u_{n+1}(1) \cdots u_m(1)$ contains an equal number of 0's and 1's. Thus each $u_n(1)$ that equals 1 with $u_{n+1}(1) = 0$ forms a pair, then ignoring these pairs each $u_n(1)$ that is a 1 and has a $u_m(1)$ equal to

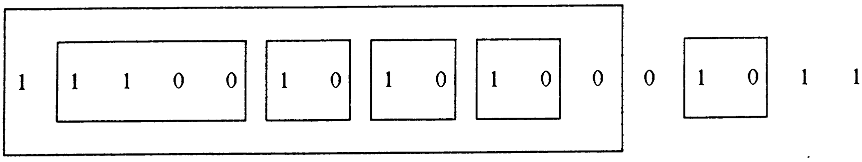

FIG. 1.

a 0 immediately to its right is paired with it and so on. See Figure 1 for an example.

The recurrence of the simple random walk guarantees that with probability 1 all symbols are paired. At this point for each pair $(u_n(1)u_m(1))$ we combine $u_n(2)$ with $u_m(2)$ to form a new pair of independent random variables. It is best to picture the original pairs as columns and then we get a new sequence of such columns by ignoring the lower coordinate [which is $u_n(1)$] and placing the upper one over $u_m(2)$ when n and m were paired above. The infinite sequence of $u_m(2)u_n(2)$ that is formed in this way (deleting all indices that correspond to the ones) is a new sequence of i.i.d. random variables with distribution $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$.

This map, which is easily seen to be invertible, was essentially what Meshalkin did to exhibit an isomorphism between $(\frac{1}{2}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8})$ and the 4-shift. We do not stop here but apply the very same process again to the sequence of $u_m(2)u_n(2)$'s. Repeating the procedure infinitely many times will leave us with a bi-infinite sequence of random variables of the form 1, 01, 001,... as described above. To invert the procedure we simply reverse the steps above. The independence of the resulting random variables is clear as is the fact that their distribution is (2^{-k}) and this completes the construction.

3. An isomorphism from the time one map of Poisson point process to $(0, 1)^{\mathbb{Z}}$. Let Ω be the set of countable subsets of the real line. For any $\omega \in \Omega$ and any interval $(a, b) \subset \mathbb{R}$, we say that the configuration on (a, b) (when ω is understood) is $\omega \cap (a, b)$. When $\omega \cap (a, b) = \varnothing$, we say that (a, b) is empty (ω is understood). We put a measure μ on Ω, which will be called the Poisson point process. μ is defined uniquely by:

1. For any interval $I = (a, b)$, the probability that I is empty depends only on $b - a$.
2. Given $r \in \mathbb{R}$, let x be the first point of ω such that $x > r$. Then $x - r$ is exponentially distributed with parameter 1, independent of the distribution of $\omega \cap [-\infty, r]$.

If these two properties hold, then the following properties are forced.

1. Let $s \subset \mathbb{R}$ have Lebesgue measure m. Then $\#\omega \cap S$ is Poisson distributed with parameter m.

2. If $S_1, S_2, S_3 \cdots$ are disjoint sets in $\mathbb{R}$, then $\omega \cap S_1, \omega \cap S_2, \omega \cap S_3 \cdots$ are all independent of each other.

Let $T: \Omega \to \Omega$ be defined by $T(\omega) = \omega_1$, where for all $r \in \mathbb{R}$, $r \in \omega_1$, iff $r - 1 \in \omega$. We now exhibit an explicit isomorphism from (Ω, T) to $(I^{\mathbb{Z}}, \hat{T})$ (from here on in $I = [0,1]$), where $\hat{T}$ is the shift map on $I^{\mathbb{Z}}$. This together with the results of the previous section defines an isomorphism from (Ω, T) to $(Q^{\mathbb{Z}}, \hat{\hat{T}})$, where Q is a particular countable partition of infinite entropy. A trivial coding cannot be carried out because with positive probability no point of the process is in I.

We now define an isomorphism $\phi: \Omega \to I^{\mathbb{Z}}$ such that ϕ is a bijection, is measure preserving and commutes with $\hat{T}$ and T. Let $\hat{I}$ be an interval.

Let $\hat{\mu}$ be the measure μ restricted to the condition that $\hat{I}$ is not empty. Make the further restriction that $\hat{\mu}$ is to be regarded as a measure on configurations on $\hat{I}$, that is, $\hat{\Omega}$ is the set of all finite nonempty subsets of the unit interval and $\hat{\mu}$ is the above defined measure on $\hat{\Omega}$.

LEMMA. $\hat{\Omega}$, $\hat{\mu}$ *is measure-theoretically isomorphic to the unit interval with Lebesgue measure.*

The lemma is a special case of the well-known theorem that any separable nonatomic regular measure space is isomorphic to the unit interval with Lebesgue measure. In the spirit of this note we should really write out a proof of this fact for this case, making the isomorphism explicit. To do this we would proceed as follows. Divide $\hat{\Omega}$ into $\hat{\Omega}_n$'s, where $\hat{\Omega}_n$ is the space of configurations of size n. For each $\hat{\Omega}_n$, we shall have an interval of I of size $\hat{\mu}(\hat{\Omega}_n)$. Thus the problem is merely to get $\hat{\Omega}_n$ to correspond to Lebesgue measure. Now $\hat{\Omega}_n$ is essentially n independent random variables, drawn from I and arranged in order. For n independent variables, the standard correspondence between I^n and I will work (see below).

Let I_n be the interval from n to $n + 1$. Fix $\omega \in \Omega$. Suppose $\omega \cap I_n \neq \varnothing$, $\omega \cap I_{n+1} = \varnothing, \omega \cap I_{n+2} = \varnothing, \omega \cap I_{n+3} = \phi, \dots, \omega \cap I_{n+k-1} = \varnothing, \omega \cap I_{n+k} \neq \varnothing$. This analysis includes the case where $k = 1$. Let ϕ be the isomorphism of the lemma (where the domain of ϕ is assumed to be configurations on I_{n+k} rather than configurations of I_0). Expand $\phi(\omega \cap I_{n+k})$ into binary expansion, that is, $\phi(\omega \cap I_{n+k}) = \sum_{j=1}^{\infty} a_j 2^{-j}$, where each a_j is either 0 or 1. The reason that a completely trivial invertible coding cannot be carried out is that for any given interval, there is a positive probability that no point of the process exists. Let $r_{n+k} = (\sum_{j=1}^{\infty} a_{jk} 2^{-j})(1 - p) + p$, where p is the probability that the unit interval is empty.

For $1 \leq i < k$, let $r_{n+i} = \sum_{j=0}^{\infty} a_{i+jk} 2^{-(j+1)} p$. This defines $r_{n+1}, r_{n+2} \cdots r_{n+k}$ and it should be noted that each integer M is in a unique set of the form $\{n+1, n+2 \cdots n+k\}$ such that $\omega \cap I_n \neq \phi$, $\omega \cap I_{n+k} \neq \phi$ and $\omega \cap I_{n+i} = \phi$ for all i, $1 \leq i < k$. Thus we have defined r_M uniquely for every integer M. The reader can verify that the map $\omega \mapsto \{r_M\}_{M \in \mathbb{Z}}$ can serve as the desired isomorphism.

REFERENCES

MESHALKIN, L. D. (1959). A case of isomorphism of Bernoulli schemes. *Dokl. Akad. Nauk SSSR* **128** 41–44.

ORNSTEIN, D. (1970). Two Bernoulli shifts with infinite entropy are isomorphic. *Adv. in Math.* **5** 339–348.

WEISS, B. (1972). The isomorphism problem in ergodic theory. *Bull. Amer. Math. Soc.* (*N.S.*) **78** 668–684.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTHERN CALIFORNIA
LOS ANGELES, CALIFORNIA 90089

INSTITUTE OF MATHEMATICS
HEBREW UNIVERSITY OF JERUSALEM
JERUSALEM
ISRAEL