PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 111, Number 1, January 1991

# ON THE CLUSTERING CONJECTURE FOR BERNOULLI FACTORS OF BERNOULLI SHIFTS

G. KELLER

(Communicated by J. J. Uhl, Jr.)

ABSTRACT. We give sufficient algebraic conditions on the probabilities $p_i$ of a Bernoulli shift $B(\mathbf{p}) = B(p_1, \dots, p_M)$ which imply that if $B(\mathbf{q}) = B(q_1, \dots, q_N)$ is a continuous factor of $B(\mathbf{p})$, then $\mathbf{q}$ is a clustering of $\mathbf{p}$.

Let $\mathbf{p} = (p_1, \dots, p_M)$ and $\mathbf{q} = (q_1, \dots, q_N)$ be probability vectors defining Bernoulli shifts $B(\mathbf{p}) = (\{1, \dots, M\}^{\mathbf{Z}}, \mathbf{p}^{\mathbf{Z}}, \sigma_M)$ and $B(\mathbf{q}) = (\{1, \dots, N\}^{\mathbf{Z}}, \mathbf{q}^{\mathbf{Z}}, \sigma_N)$. We assume throughout that $B(\mathbf{q})$ *is a continuous factor of* $B(\mathbf{p})$, i.e., there is a continuous homomorphism $\Phi$ from $B(\mathbf{p})$ to $B(\mathbf{q})$. (Homomorphism means that $\mathbf{p}^{\mathbf{Z}} \circ \Phi^{-1} = \mathbf{q}^{\mathbf{Z}}$ and $\Phi \circ \sigma_M = \sigma_N \circ \Phi$.) If $h(\mathbf{p}) = -\sum_i p_i \log p_i$ denotes the entropy of $B(\mathbf{p})$, then $h(\mathbf{p}) \geq h(\mathbf{q})$.

Tuncel [4] and, independently, del Junco et al. [2] showed that if $h(\mathbf{p}) = h(\mathbf{q})$, then $\mathbf{q}$ is just a permutation of $\mathbf{p}$, i.e., there is a trivial factor map from $B(\mathbf{p})$ onto $B(\mathbf{q})$. Note, however, that $\Phi$ need not be this trivial map! Del Junco et al. showed that if $\mathbf{p} = (1/M, \dots, 1/M)$ and $h(\mathbf{p}) > h(\mathbf{q})$, then $\mathbf{q} = (i_1/M, \dots, i_N/M)$ with $i_j \in \mathbf{N}$. These results lead to the *clustering conjecture*:

> If $B(\mathbf{q})$ is a continuous factor of $B(\mathbf{p})$, then $\mathbf{q}$ is a clustering of $\mathbf{p}$, i.e., there is a partition $(I_1, \dots, I_N)$ of $\{1, \dots, M\}$ such that $q_k = \sum_{i \in I_k} p_i$.

This conjecture was disproved by Boyle and Tuncel [1]. They showed that $B(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ is a continuous (two-block) factor of $B(\frac{1}{6}, \frac{1}{6}, \frac{2}{9}, \frac{2}{9}, \frac{2}{9})$. Smorodinsky [3] analyzed this example and arrived at a method for producing further counterexamples to the clustering conjecture.

In this note we present an algebraic independence condition on the $p_i$ which assures that $B(\mathbf{p})$ obeys the clustering conjecture.

Suppose now that $\mathbf{p} = (\overbrace{\pi_1, \dots, \pi_1}^{u_1\text{-times}}, \overbrace{\pi_2, \dots, \pi_2}^{u_2\text{-times}}, \dots, \overbrace{\pi_n, \dots, \pi_n}^{u_n\text{-times}})$, $\sum_i u_i = M$, and let $P = \mathbf{p}^{\mathbf{Z}}$, $\sigma = \sigma_M$. Set $\lambda_i = \pi_i/\pi_1$, $i = 1, \dots, n$, and note that $\pi_1^{-1} = u_1 + \sum_{i=2}^n \lambda_i u_i$. Fix $s \in \{1, \dots, N\}$ (the state space of $B(\mathbf{q})$)

Received by the editors February 23, 1987.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 28D99.

and let $A = A(s) = \Phi^{-1}({}_0[s]_0)$, the set of all bisequences over $\{1, \dots, M\}$ whose images under $\Phi$ have $s$ as its 0-coordinate. Since $\Phi$ is continuous, there are $r_1, r_2 \in \mathbf{N}$ such that $A$ is the disjoint union of cylinders $\{\omega \in \{1, \dots, M\}^{\mathbf{Z}}; \omega_i = x_i, -r_1 \le i \le r_2\}$ for certain $x_i \in \{1, \dots, M\}$. Let $r = r_1 + r_2 + 1$ and $P_k(A) = P(A \cap \sigma^{-1} A \cap \dots \cap \sigma^{-(k-1)} A)$, $k \in \mathbf{N}$. $P_k(A)$ can be written as

$$P_k(A) = \sum_{\mathbf{j}} a_k(\mathbf{j}) \cdot \pi_1^{j_1} \cdots \pi_n^{j_n},$$

where the sum extends over all $n$-tuples $\mathbf{j} = (j_1, \dots, j_n)$ of nonnegative integers with $\sum_\nu j_\nu = k + r - 1$. The coefficients $a_k(\mathbf{j})$ are nonnegative integers, too.

Extracting $\pi_1^{k+r-1}$ this can be rewritten as

$$P_k(A) = \pi_1^{k+r-1} R_k(\lambda_2, \dots, \lambda_n),$$

where $R_k$ is the polynomial given by

$$R_k(X_2, \dots, X_n) = \sum_{\mathbf{j}} a_k(\mathbf{j}) \cdot X_2^{j_2} \cdots X_n^{j_n},$$

the sum ranging over the same index set as above.

Since $\Phi$ is a homomorphism, $P_k(A) = (P_1(A))^k$ or, equivalently,

$$\pi_1^{k+r-1} R_k(\lambda_2, \dots, \lambda_n) = \pi_1^{kr} (R_1(\lambda_2, \dots, \lambda_n))^k,$$

and we get the following necessary condition for $B(\mathbf{q})$ being a continuous factor of $B(\mathbf{p})$:

$$(1) \qquad \frac{R_k(\lambda_2, \dots, \lambda_n)}{R_1(\lambda_2, \dots, \lambda_n)} = \left( \frac{R_1(\lambda_2, \dots, \lambda_n)}{(u_1 + \sum_{i=2}^n \lambda_i u_i)^{r-1}} \right)^{k-1}, \quad \text{for all } k.$$

**Proposition 1.** *Suppose that $\lambda_2, \dots, \lambda_n$ are algebraically independent. If $B(\mathbf{q})$ is a continuous factor of $B(\mathbf{p})$, then $\mathbf{q}$ is a clustering of $\mathbf{p}$.*

*Proof.* In this case (1) is equivalent to the corresponding polynomial identities with $\lambda_i$ replaced by $X_i$. As the polynomials in $n - 1$ variables over the integers constitute a unique factorization domain and as the linear polynomial $(u_1 + \sum_{i=2}^n X_i u_i)$ is irreducible over $\mathbf{Q}[X_2, \dots, X_n]$, it follows that $(u_1 + \sum_{i=2}^n X_i u_i)^{r-1}$ divides $R_1(X_2, \dots, X_n)$. Hence there is a polynomial $S(X_2, \dots, X_n)$ over $\mathbf{Z}$ such that

$$(2) \qquad R_1(X_2, \dots, X_n) = S(X_2, \dots, X_n) \cdot \left( u_1 + \sum_{i=2}^n u_i X_i \right)^{r-1}.$$

Since $R_1$ is of degree $\le r$, $S(X_2, \dots, X_n) = c_1 + \sum_{i=2}^n c_i X_i$ for some $c_i \in \mathbf{Z}$, and, comparing the coefficients on both sides of (2), one gets

$$n_A(i) := a_1(0, \dots, 0, r, 0, \dots, 0) = c_i u_i^{r-1}, \qquad i = 1, \dots, n,$$

(the $r$ occurs at the $i$th position). $n_A(i)$ is just the number of blocks over the $r$ coordinates $-r_1, \ldots, r_2$ which are contained in $A$ and all of whose entries have probability $\pi_i$. Hence $c_i = c_i(A) = n_A(i)u_i^{-(r-1)}$ is a nonnegative integer,

$$\begin{aligned} P(A) &= \pi_1^r R_1(\lambda_2, \ldots, \lambda_n) \\ &= \pi_1 \left(c_1 + \sum_{i=2}^{n} c_i\lambda_i\right) \pi_1^{r-1} \left(u_i + \sum_{i=2}^{n} u_i\lambda_i\right)^{r-1} \\ &= \sum_{i=1}^{n} c_i\pi_i, \end{aligned}$$

and, for each fixed $i = 1, \ldots, n$,

$$\sum_{s=1}^{N} c_i(A(s)) = u_i^{-(r-1)} \sum_{s=1}^{N} n_{A(s)}(i) = u_i. \quad \square$$

*Remark.* If $\mathbf{p} = (1/M, \ldots, 1/M)$, then (1) reduces to the key observation in the proof of the clustering conjecture for this case as given by del Junco et al. [2].

*Note added in proof.* Recently, S. Tuncel [Ergodic Theory Dynamical Systems **9** (1989), 561–570] proved a considerable extension of this result to the case where only the transcendental elements from $\{\lambda_2, \ldots, \lambda_n\}$ are assumed to be algebraically independent. This includes the case where all $\lambda_i$ are algebraic.

## References


1. M. Boyle and S. Tuncel, *Infinite-to-one codes and Markov measures*, Trans. Amer. Math. Soc. **285** (1984), 657–684.
2. A. del Junco, M. Keane, B. Kitchens, B. Marcus, and L. Swanson, *Continuous homomorphisms of Bernoulli schemes, Ergodic theory and dynamical systems* I, Progress in Mathematics, vol. 10, Birkhäuser, Boston, Basel, and Stuttgart, 1981, pp. 91–111.
3. M. Smorodinsky, *Block codes for Bernoulli shifts*, Israel J. Math. **49** (1984), 325–330.
4. S. Tuncel, *Conditional pressure and coding*, Israel J. Math. **39** (1981), 101–112.



Sonderforschungsbereich 123, Universität Heidelberg, D-6900 Heidelberg, West Germany

*Current address*: Mathematisches Institut, Universität Erlangen-Nürnberg, Bismarckstr. 1 $\frac{1}{2}$, D-8520 Erlangen, West Germany