ON THE CLUSTERING CONJECTURE
FOR BERNOULI FACTORS OF BERNOULI SHIFTS

G. KELLER

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

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

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

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

If $B(q)$ is a continuous factor of $B(p)$, then $q$ is a clustering of $p$, i.e., there is a partition $(I_1, \ldots, I_M)$ of $\{1, \ldots, M\}$ 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 $B(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6})$. 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 $B(p)$ obeys the clustering conjecture.

Suppose now that $p = (\pi_1, \ldots, \pi_1, \pi_2, \ldots, \pi_2, \ldots, \pi_n, \ldots, \pi_n)$, $\sum u_i = M$, and let $P = p^\mathbb{Z}$, $\sigma = \sigma_M$. Set $\lambda_i = \pi_i/\pi_1$, $i = 1, \ldots, n$, and note that $\pi_1^{-1} = u_1 + \sum_{i=2}^n \lambda_i u_i$. Fix $s \in \{1, \ldots, N\}$ (the state space of $B(q)$).
and let \( A = A(s) = \Phi^{-1}(\mathbb{g}[s]_0) \), the set of all bisequences over \( \{1, \ldots, M\} \) whose images under \( \Phi \) have \( s \) as its 0-coordinate. Since \( \Phi \) is continuous, there are \( r_1, r_2 \in \mathbb{N} \) such that \( A \) is the disjoint union of cylinders \( \{ \omega \in \{1, \ldots, M\}^\mathbb{Z}; \omega_i = x_i, -r_1 \leq i \leq r_2 \} \) for certain \( x_i \in \{1, \ldots, M\} \). Let \( r = r_1 + r_2 + 1 \) and \( P_k(A) = P(A \cap \sigma^{-1}A \cap \ldots \cap \sigma^{-(k-1)}A), k \in \mathbb{N} \). \( P_k(A) \) can be written as
\[
P_k(A) = \sum_j a_k(j) \cdot \pi_1^{j_1} \cdots \pi_n^{j_n},
\]
where the sum extends over all \( n \)-tuples \( j = (j_1, \ldots, j_n) \) of nonnegative integers with \( \sum j_i = k + r - 1 \). The coefficients \( a_k(j) \) are nonnegative integers, too.

Extracting \( \pi_1^{k+r-1} \) this can be rewritten as
\[
P_k(A) = \pi_1^{k+r-1} R_k(\lambda_2, \ldots, \lambda_n),
\]
where \( R_k \) is the polynomial given by
\[
R_k(X_2, \ldots, X_n) = \sum_j a_k(j) \cdot X_2^{j_2} \cdots X_n^{j_n},
\]
the sum ranging over the same index set as above.

Since \( \Phi \) is a homomorphism, \( P_k(A) = (P_1(A))^k \) or, equivalently,
\[
\pi_1^{k+r-1} R_k(\lambda_2, \ldots, \lambda_n) = (R_1(\lambda_2, \ldots, \lambda_n))^k,
\]
and we get the following necessary condition for \( B(q) \) being a continuous factor of \( B(p) \):
\[
(1) \quad \frac{R_k(\lambda_2, \ldots, \lambda_n)}{R_1(\lambda_2, \ldots, \lambda_n)} = \left( \frac{R_1(\lambda_2, \ldots, \lambda_n)}{(u_1 + \sum_{i=2}^n \lambda_i u_i)^{r-1}} \right)^{k-1}, \text{ for all } k.
\]

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

**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 \( \mathbb{Q}[X_2, \ldots, X_n] \), it follows that \( (u_1 + \sum_{i=2}^n X_i u_i)^{r-1} \) divides \( R_1(X_2, \ldots, X_n) \). Hence there is a polynomial \( S(X_2, \ldots, X_n) \) over \( \mathbb{Z} \) such that
\[
(2) \quad R_1(X_2, \ldots, X_n) = S(X_2, \ldots, X_n) \cdot \left( u_1 + \sum_{i=2}^n u_i X_i \right)^{r-1}.
\]
Since \( R_1 \) is of degree \( \leq r \), \( S(X_2, \ldots, X_n) = c_1 + \sum_{i=2}^n c_i X_i \) for some \( c_i \in \mathbb{Z} \), and, comparing the coefficients on both sides of \( (2) \), one gets
\[
n_A(i) := a_1(0, \ldots, 0, r, 0, \ldots, 0) = c_i u_i^{r-1}, \quad i = 1, \ldots, n,
\]
BERNOULLI FACTORS

53

(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,

$$P(A) = \pi_1 R_1(\lambda_2, \ldots, \lambda_n)$$

$$= \pi_1 \left( c_1 + \sum_{i=2}^{n} c_i \lambda_i \right) \pi_1^{r-1} \left( u_i + \sum_{i=2}^{n} u_i \lambda_i \right)^{r-1}$$

$$= \sum_{i=1}^{n} c_i \pi_i,$$

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 $p = (1/M, \ldots, 1/M)$, 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


SONDERFORSCHUNGSBEREICH 123, UNIVERSITÄT HEIDELBERG, D-6900 HEIDELBERG, WEST GERMANY

Current address: Mathematisches Institut, Universität Erlangen-Nürnberg, Bismarckstr. 1 1/2, D-8520 Erlangen, West Germany

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use