Normal numbers with given limits of multiple ergodic averages

Lingmin Liao  
LAMA UMR 8050, CNRS, Université Paris-Est Créteil,  
61 Avenue du Général de Gaulle, 94010 Créteil Cedex, France  
e-mail: lingmin.liao@u-pec.fr  
Michał Rams  
Institute of Mathematics, Polish Academy of Sciences  
ul. Śniadeckich 8, 00-656 Warszawa, Poland  
e-mail: rams@impan.pl

Abstract

We are interested in the set of normal sequences in the space \(\{0,1\}^\mathbb{N}\) with a given frequency of the pattern 11 in the positions \(k,2k\). The topological entropy of such sets is determined.

1 Introduction and statement of results

Let \(\Sigma = \{0,1\}^\mathbb{N}\). In [K12, FLM12], the authors proposed to calculate the topological entropy spectrum of level sets of multiple ergodic averages. Here, the topological entropy means Bowen’s topological entropy (in the sense of [B73]) which can be defined for any subset, not necessarily invariant. Among other questions, they asked for the topological entropy of

\[
A_\alpha := \{ (\omega_k)_{\mathbb{N}}^\infty \in \Sigma : \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \omega_k \omega_{2k} = \alpha \} \quad (\alpha \in [0,1]).
\]

As a first step to solve the question, they also suggested to study a subset of \(A_0\):

\[
A := \{ (\omega_k)_{\mathbb{N}}^\infty \in \Sigma : \omega_k \omega_{2k} = 0 \text{ for all } k \geq 1 \}.
\]

The topological entropy of \(A\) was later given by Kenyon, Peres and Solomyak [KPS12].

2010 Mathematics Subject Classification: Primary 28A80, Secondary 11K16, 37B40
Theorem 1.1 (Kenyon-Peres-Solomyak). We have
\[ h_{\text{top}}(A) = -\log(1 - p) = 0.562399..., \]
where \( p \in [0, 1] \) is the unique solution of
\[ p^2 = (1 - p)^3. \]

Enlightened by the idea of [KPS12], the question about \( A_\alpha \) was finally answered by Peres and Solomyak [PS12], and then in higher generality by Fan, Schmeling and Wu [FSW16].

Theorem 1.2 (Peres-Solomyak, Fan-Schmeling-Wu). For any \( \alpha \in [0, 1] \), we have
\[ h_{\text{top}}(A_\alpha) = -\log(1 - p) - \frac{\alpha}{2} \log \frac{q(1 - p)}{p(1 - q)}, \]
where \( (p, q) \in [0, 1]^2 \) is the unique solution of the system
\[
\begin{align*}
    p^2(1 - q) &= (1 - p)^3, \\
    2pq &= \alpha(2 + p - q).
\end{align*}
\]
In particular, \( h_{\text{top}}(A_0) = h_{\text{top}}(A) \).

Another, interesting, related set is
\[
B := \left\{ (\omega_k)^\infty_{k=1} \in \Sigma : \omega_k =\omega_{2k} \text{ for all } k \geq 1 \right\}.
\]

The sequence \( x \in \{0, 1\}^\mathbb{N} \) is said to be simple normal if the frequency of the digit 0 in the sequence is 1/2. It is said to be normal if for all \( n \in \mathbb{N} \), each word in \( \{0, 1\}^n \) of length \( n \) has frequency 1/2^n. We denote the set of normal sequences by \( \mathcal{N} \).

We are interested in the intersection of \( \mathcal{N} \) with the set \( A_\alpha \) of given frequency of the pattern 11 in \( w_kw_{2k} \). For the usual ergodic (Birkhoff) averages the normal numbers all belong to one set in the multifractal decomposition – the situation for multiple ergodic averages turns out to be very different.

Our results are as follows:

Theorem 1.3. For \( \alpha \leq 1/2 \) we have
\[ h_{\text{top}}(\mathcal{N} \cap A_\alpha) = \frac{1}{2} \log 2 + \frac{1}{2} H(2\alpha), \]
where \( H(x) = -x \log x - (1 - x) \log(1 - x) \). For \( \alpha > 1/2 \) the set \( \mathcal{N} \cap A_\alpha \) is empty.

Further,
\[ h_{\text{top}}(\mathcal{N} \cap A) = h_{\text{top}}(\mathcal{N} \cap A_0) = \frac{1}{2} \log 2. \]

Moreover, \( \mathcal{N} \cap B \subset A_{1/2} \) and
\[ h_{\text{top}}(\mathcal{N} \cap B) = h_{\text{top}}(\mathcal{N} \cap A_{1/2}) = h_{\text{top}}(B) = \frac{1}{2} \log 2. \]
The last statement of Theorem 1.3 was recently proved, in higher generality, in [ABC].

Let us now define the set of sequences with prescribed frequency of 0’s and 1’s:

\[ E_\theta := \{ x \in [0, 1] : \lim_{n \to \infty} \frac{\omega_1(x) + \cdots + \omega_n(x)}{n} = \theta \}. \]

In particular, \( E_{1/2} \) is the set of simple normal sequences.

**Theorem 1.4.** We have

\[ h_{\text{top}}(E_\theta \cap A_\alpha) = (1 - \frac{\theta}{2})H(\frac{2\theta - \alpha}{2 - \theta}) + \frac{\theta}{2} H(\frac{\theta - \alpha}{\theta}) \]

for \( \alpha \leq \theta \leq (2 + \alpha)/3 \), otherwise \( E_\theta \cap A_\alpha = \emptyset \). Further,

\[ h_{\text{top}}(E_\theta \cap A) = h_{\text{top}}(E_\theta \cap A_0) = \frac{2 - \theta}{2} H(\frac{2\theta}{2 - \theta}). \]

Note that

\[ h_{\text{top}}(E_{1/2} \cap A) = \frac{3}{4} H(\frac{2}{3}) > h_{\text{top}}(\mathcal{N} \cap A). \]

**Remark.** Applying the results of [PS12] one can show that

\[ h_{\text{top}}(E_\theta \cap A_\alpha) = h_{\text{top}}(A_\alpha) \]

if and only if \( \alpha, \theta \) satisfy the relation

\[ (2\theta - \alpha)^2(\theta - \alpha)(2 - \theta) = \theta(2 - 3\theta + \alpha)^3. \]

In particular, when

\[ \theta = \frac{2}{3} \left(1 + \left(\frac{2}{23}\right)^{2/3} \sqrt[3]{3\sqrt{69} - 23} - \left(\frac{2}{23}\right)^{2/3} \sqrt[3]{3\sqrt{69} + 23}\right) = 0.354..., \]

i.e., the unique real solution of the equation \( 4\theta^2(2 - \theta) = (2 - 3\theta)^3 \), we have

\[ \dim_H E_\theta \cap A = \dim_H A. \]

We omit the details.
2 Proof of Theorem 1.3

Given \( \alpha \in [0, 1] \), let \( \mu_\alpha \) be a probability measure on \( \Sigma \) given by

- if \( k \) is odd then \( \omega_k = 1 \) with probability \( \frac{1}{2} \),
- if \( k \) is even and \( \omega_{k/2} = 1 \) then \( \omega_k = 1 \) with probability \( 2\alpha \),
- if \( k \) is even and \( \omega_{k/2} = 0 \) then \( \omega_k = 1 \) with probability \( 1 - 2\alpha \),

with the events \( \{\omega_k = 1\} \) and \( \{\omega_\ell = 1\} \) independent except when \( k/\ell \) is a power of 2. Precisely, let \((p_0, p_1) := (1/2, 1/2)\) and let

\[
\begin{pmatrix}
p_{00} & p_{01} \\
p_{10} & p_{11}
\end{pmatrix} := \begin{pmatrix} 2\alpha & 1 - 2\alpha \\ 1 - 2\alpha & 2\alpha \end{pmatrix}.
\]

Let \( C_n(\omega_1, \ldots, \omega_n) \) be the set of sequences beginning with the word \( \omega_1 \cdots \omega_n \in \{0, 1\}^n \). Such sets are called cylinders of order \( n \). The measure \( \mu_\alpha \) of a cylinder is given by

\[
\mu_\alpha(\omega_1 \cdots \omega_n]) = \prod_{k=1}^{\lceil n/2 \rceil} p_{\omega_{2k-1}} \cdot \prod_{k=1}^{\lfloor n/2 \rfloor} p_{\omega_k \omega_{2k}} = \frac{1}{2^{\lceil n/2 \rceil}} \cdot \prod_{k=1}^{\lfloor n/2 \rfloor} p_{\omega_k \omega_{2k}}.
\]

where \( \lceil \cdot \rceil, \lfloor \cdot \rfloor \) denote the ceiling function and the integer part function correspondingly.

We will prove that the measure \( \mu_\alpha \) is supported on the set \( \mathcal{N} \cap A_\alpha \).

**Lemma 2.1.** We have

\[
\mu_\alpha(\mathcal{N} \cap A_\alpha) = 1.
\]

**Proof.** Denote

\[
x_n(\omega) = \frac{2}{n} \sum_{k=n/2+1}^{n} \omega_k.
\]

For a \( \mu_\alpha \)-typical \( \omega \), the Law of Large Numbers implies

\[
x_{2n}(\omega) = \frac{1}{4} + \frac{x_n(\omega)}{2} 2\alpha + \frac{1 - x_n(\omega)}{2} (1 - 2\alpha) + o(1).
\]

Noting that \( |\frac{4n - 1}{2}| < 1 \), we have as \( k \to \infty \),

\[
x_{2k}(\omega) \to \frac{1}{2}.
\]

By [PS12, Lemma 5], this implies that \( \mu_\alpha \)-almost surely

\[
\lim_{n \to \infty} x_n(\omega) = \frac{1}{2}.
\]

(2.1)
Then, for $\mu_\alpha$-a.e. $\omega$,

$$\frac{2}{n} \sum_{k=n/2+1}^{n} \omega_k \omega_{2k} = x_n(\omega)(2\alpha + o(1)) \to \alpha.$$  

Thus $\mu_\alpha(A_n) = 1$.

Now, we show $\mu(N) = 1$. We can divide the set of natural numbers into infinitely many subsets of the form $A_k = \{2k - 1, 4k - 2, \ldots, 2^\ell (2k - 1), \ldots \}$ ($k \geq 1$). Let $B_k$ be the $\sigma$-field generated by events $\{\omega_{2^\ell(2k-1)} = 1\}$, $\ell \in \mathbb{N}$.

Observe that for the measure $\mu$ the $\sigma$-fields $B_k$ are independent.

Observe further that $\mu(\omega_{2^\ell(2k-1)} = 1) = 1/2$ for every $k, \ell$. Indeed, for $\ell = 0$ it follows from the definition of $\mu$, and then it is proved by induction:

$$\begin{align*}
\mu(\omega_{2^\ell(2k-1)} = 1) &= \mu(\omega_{2^\ell(2k-1)} = 1 \land \omega_{2^{\ell-1}(2k-1)} = 1) + \mu(\omega_{2^\ell(2k-1)} = 1 \land \omega_{2^{\ell-1}(2k-1)} = 0) \\
&= 2\alpha \cdot 1/2 + (1 - 2\alpha) \cdot 1/2 = 1/2.
\end{align*}$$

Consider now, for any $n$, the sequence $\omega_{m+1}, \ldots, \omega_{m+n}$. If $m \geq n$ then positions $m+1, \ldots, m+n$ come all from different $A_k$'s, thus $\omega_{m+1}, \ldots, \omega_{m+n}$ are independent and each of them takes values 0, 1 with probability 1/2 respectively. That is, the measure $\mu$ restricted to such subset of positions is $(\frac{1}{2}, \frac{1}{2})$-Bernoulli, and for any word $\eta \in \{0, 1\}^n$ with $n \leq m$, the probability that we have $\omega_{m+i} = \eta_i$ for $i = 1, \ldots, n$ equals $2^{-n}$. Thus, for a given word $\eta \in \{0, 1\}^n$ we can divide $\mathbb{N}$ into intervals $[2^j + 1, 2^{j+1}]$, inside all except initial finitely many of them (with $j < \log_2 n$) for any $\mu$-generic sequence $\omega$ the frequency of appearance of $\eta$ equals $2^{-n} + O(2^{-j/2} \log j)$, and this means that the $\mu$-generic sequence $\omega$ is normal.

Next, we will calculate the local dimension of the measure $\mu_\alpha$ with the help of Mass Distribution Principle, [?, ?]. We denote for $x \in [0, 1]$

$$H(x) = -x \log x - (1 - x) \log(1 - x)$$

with convention $H(0) = H(1) = 0$.

**Lemma 2.2.** We have

$$h_{\mu_\alpha} = \frac{1}{2} \log 2 + H(2\alpha).$$

**Proof.** For $\omega \in \Sigma$ denote

$$C_n(\omega) = \{\tau \in \Sigma; \tau_k = \omega_k \forall k \leq n\}.$$  

Let

$$h_n(\omega) := \log \mu_\alpha(C_{2n}(\omega)) - \log \mu_\alpha(C_n(\omega)).$$
By the Law of Large Numbers, for \( \mu_\alpha \)-typical \( \omega \) and for big enough \( n \) we have
\[
\frac{2}{n} h_n(\omega) = -\log 2 + (1 - x_n(\omega))((2\alpha \log(2\alpha) + (1 - 2\alpha) \log(2\alpha)) + x_n(\omega)((1 - 2\alpha) \log(1 - 2\alpha)) + (2\alpha) \log(2\alpha)) + o(1).
\]
Thus,
\[
\lim_{n \to \infty} -\frac{1}{n} h_n(\omega) = \frac{1}{2} \log 2 + \frac{1}{2} H(2\alpha) \quad \mu_\alpha - a.e..
\]
Note that for all \( k, n \in \mathbb{N} \)
\[
\frac{1}{k 2^n} \log \mu_\alpha(C_{k2^n}(\omega)) = \frac{1}{k 2^n} \sum_{i=1}^{n-1} h_{2i}.
\]
Then for all \( k \in \mathbb{N} \)
\[
\lim_{n \to \infty} -\frac{1}{k 2^n} \log \mu_\alpha(C_{k2^n}(\omega)) = \frac{1}{2} \log 2 + \frac{1}{2} H(2\alpha) \quad \mu_\alpha - a.e..
\]
Hence, by [PS12, Lemma 5],
\[
h_{\mu_\alpha} = \lim_{n \to \infty} -\frac{1}{n} \mu_\alpha(C_n(\omega)) = \frac{1}{2} \log 2 + \frac{1}{2} H(2\alpha) \quad \mu_\alpha - a.e..
\]
Applying the Mass Distribution Principle ends the proof.

To finish the proof of the lower bound we note that \( A \subset A_0 \) but the measure \( \mu_0 \) is actually supported on \( A \), that the measure \( \mu_{1/2} \) is supported on \( B \), and that the relation \( \mathcal{N} \cap B \subset A_{1/2} \) follows from
\[
\frac{1}{n} \{n + 1 \leq j \leq 2n : \omega_j = \omega_{2j} = 1\} = \frac{1}{n} \{n + 1 \leq j \leq 2n : \omega_j = 1\} \to \frac{1}{2}
\]
being satisfied for every \( \omega \in \mathcal{N} \cap B \).

For the upper bound, let us first observe that
\[
\frac{1}{n} \sum_{k=1}^{n} \omega_k \omega_{2k} \leq \frac{1}{n} \sum_{k=1}^{n} \omega_k
\]
and the right hand side converges to \( 1/2 \) for every normal sequence \( \omega \). Thus, the set \( \mathcal{N} \cap A_\alpha \) is empty for all \( \alpha > 1/2 \).

We will now need the following lemma

**Lemma 2.3.** Let \( \omega \) be a normal sequence and let \( (n_k = \ell_1 + k \ell_2) \) be an arithmetic subsequence of \( \mathbb{N} \). Then \( \omega \) restricted to the positions \( (n_k) \) is normal.

**Proof.** This is a well-known result of Kamae [K73].
Let us fix some \( m > 0 \). For \( N > m \) and \( i = 0, 1, \ldots, m \) denote by \( R(N, i) \) the set \( \{ 2^i(2k - 1), k \leq 2^{N-i-1} \} \) (for example, \( R(N, 0) \) is the set of odd numbers smaller than \( 2^N \)). Further, let \( R(N, i, I) = R(N - 2, i), \ R(N, i, II) = R(N - 1, i) \setminus R(N - 2, i) \), and \( R(N, i, III) = R(N, i) \setminus R(N - 1, i) \). Note here obvious relations
\[
2R(N, i, I) = R(N, i + 1, I) \cup R(N, i + 1, II),
2R(N, i, II) = R(N, i + 1, III),
2R(N, i, III) \cap R(N, i + 1) = \emptyset.
\]

We denote by \( \mathcal{N}(N, m, \varepsilon) \) the set of sequences \( \omega \) such that for all \( n \geq N \) in each \( R(n, i, *) \), \( i = 0, \ldots, m \), \( * \in \{ I, II, III \} \) the frequency of 1’s is between \( 1/2 - \varepsilon \) and \( 1/2 + \varepsilon \). By Lemma 2.3,
\[
\mathcal{N} \subset \bigcap_{\varepsilon > 0} \bigcap_{m = 1}^{\infty} \bigcup_{N = m + 1}^{\infty} \mathcal{N}(N, m, \varepsilon).
\]

Similarly, let us denote by \( A(\alpha, N, \varepsilon) \) the set of sequences \( \omega \) such that for all \( n \geq N \) we have
\[
\alpha - \varepsilon < 2^{-n+1} \sum_{j=1}^{2^{n-1}} \omega_j \omega_{2j} < \alpha + \varepsilon.
\]
We have
\[
A_{\alpha} = \bigcap_{\varepsilon > 0} \bigcup_{N = 1}^{\infty} A(\alpha, N, \varepsilon).
\]

To obtain the upper bound, we will estimate from above the number of cylinders \( [\omega_1, \ldots, \omega_{2^N}] \) needed to cover the set \( \mathcal{N}(N, m, \varepsilon) \cap A(\alpha, N, \varepsilon) \). Let us fix \( N, m, \varepsilon \). For \( i = 1, \ldots, m \), \( k_1, k_2 \in \{0, 1\} \), and \( * \in \{ I, II \} \) we denote
\[
X_{k_1 k_2 *, i}^1(\omega) = \sharp \{ n \in R(N, i - 1, *); \omega_n = k_1, \omega_{2n} = k_2 \}.
\]

For example, \( X_{01, I}^1(\omega) \) denotes the number of odd positions smaller than \( 2^{N-2} \) such that \( \omega_n = 0, \omega_{2n} = 1 \). Similarly, let
\[
X_{k_1, i}^1(\omega) = \sharp \{ n \in R(N, i, *); \omega_n = k_1 \}.
\]

We have obvious relations: for any \( i \)
\[
X_{10, I}^i + X_{11, I}^i = X_{1, I}^{i-1}
X_{00, I}^i + X_{01, I}^i = X_{0, I}^{i-1}
X_{10, II}^i + X_{11, II}^i = X_{1, II}^{i-1}
\]
\[X_{00,II}^i + X_{01,II}^i = X_{0,II}^{i-1}\]
\[X_{01,I}^i + X_{11,I}^i = X_{1,I}^i + X_{1,II}^i\]
\[X_{00,I}^i + X_{10,I}^i = X_{0,I}^i + X_{0,II}^i\]
\[X_{01,II}^i + X_{11,II}^i = X_{1,III}^i\]
\[X_{00,II}^i + X_{10,II}^i = X_{0,III}^i.\]

Note that for a sequence \(\omega \in N(N,m,\varepsilon)\) the right hand sides in all those relations is in range \(2^{N-3-i} \cdot (1 - \varepsilon, 1 + \varepsilon)\). In particular,
\[|X_{11,I}^i - X_{00,I}^i| \leq \varepsilon \cdot 2^{N-2-i}.\]

We can now start the counting. The values of \(\{\omega_n; n \in R(N,0)\}\) can be chosen in no more than \(2^{2N-1}\) ways. After we have chosen \(\{\omega_n; n \in R(N,i - 1)\}\), we can choose \(\{\omega_n; n \in R(N,i)\}\) in no more than
\[
\left(\frac{X_{1-I}^{i-1}}{X_{1-I,I}^{i}}\right) \cdot \left(\frac{X_{0-0}^{i-1}}{X_{0-I,II}^{i}}\right) \cdot \left(\frac{X_{1-I,II}^{i-1}}{X_{1-I,III}^{i}}\right) \cdot \left(\frac{X_{0-0,II}^{i-1}}{X_{0-0,III}^{i}}\right)
\]
ways. Finally, after we have chosen \(\{\omega_n; n \in R(N,i)\}\) for all \(i \leq m\), we will still have \(2^{N-m-1}\) positions left, which we can cover in no more than \(2^{2N-m-1}\) ways. Thus, for any choice of \((X_{00,I}^i, X_{11,I}^i, X_{00,II}^i, X_{11,II}^i)\), the logarithm of total number of cylinders needed \(Z\) is not larger than
\[
\log Z((X_{00,I}^i, X_{11,I}^i, X_{00,II}^i, X_{11,II}^i)) \leq (2^{N-1} + 2^{N-m-1}) \log 2 + \sum_{i=1}^{m} 2 \log \left(\frac{2^{N-3-i}}{X_{1-I,I}^{i}}\right) + 2 \log \left(\frac{2^{N-3-i}}{X_{1-I,II}^{i}}\right) + 2^{N-3-i} O(\varepsilon)
\]
and there are no more than \(\prod_{i=1}^{m} 2^{4(N-i-3)} < 2^{4mN} \ll 2^{2N}\) such choices.

We estimate
\[
\log \left(\frac{n}{k}\right) \approx n \left(-\frac{k}{n} \log \frac{k}{n} - \frac{n-k}{n} \log \frac{n-k}{n}\right) = nH\left(\frac{k}{n}\right)
\]
and observe that \(H\) is a concave function, thus we can apply Jensen inequality. We get
\[
\log Z((X_{00,I}^i, X_{11,I}^i, X_{00,II}^i, X_{11,II}^i)) \leq (2^{N-1} + 2^{N-m-1}) \log 2 + \sum_{i=1}^{m} 2^{N-i-1} \cdot H\left(\frac{1}{\sum_{i=1}^{m} 2^{N-i-1}} \cdot \sum_{i=1}^{m} 2^{N-i-2} \frac{X_{1-I,II}^{i}}{2^{N-i-3}}\right) + \sum_{i=1}^{m} 2^{N-i-3} \cdot O(\varepsilon).
\]
Hence,
\[
\log Z((X_{00,I}^i, X_{11,I}^i, X_{00,II}^i, X_{11,II}^i)) \leq 2^{N-1} \log 2 + 2^{N-1} H \left( \sum_{i=1}^{m} (X_{11,I}^i + X_{11,II}^i) \right) + 2^N \cdot (O(\varepsilon + 2^{-m})).
\]

On the other hand, for all \( \omega \in A(\alpha, N, \varepsilon) \),
\[
\left| 2^{-N+1} \sum_{i=1}^{m} (X_{11,I}^i + X_{11,II}^i) - 2\alpha \right| < \varepsilon.
\]
Passing with \( m, N \) to infinity and with \( \varepsilon \) to 0, we finish the proof of the upper bound.

3 Proof of Theorem 1.4

Given \( p, q \in [0, 1] \), let \( \mu_{p,q} \) be a probability measure on \( S \) given by
- if \( k \) is odd then \( \omega_k = 1 \) with probability \( p \),
- if \( k \) is even and \( \omega_k/2 = 0 \) then \( \omega_k = 1 \) with probability \( p \),
- if \( k \) is even and \( \omega_k/2 = 1 \) then \( \omega_k = 1 \) with probability \( q \),
with events \( (\omega_k = 1) \) and \( (\omega_k = 1) \) independent except when \( k/\ell \) is a power of 2. Precisely, let \( (p_0, p_1) := (1-p, p) \) and let
\[
\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} := \begin{pmatrix} 1-p & p \\ 1-q & q \end{pmatrix}.
\]
Then the measure \( \mu_{p,q} \) of a cylinder is given by
\[
\mu_{p,q}([\omega_1 \cdots \omega_n]) = \prod_{k=1}^{[n/2]} p_{\omega_{2k-1}} \cdot \prod_{k=1}^{[n/2]} p_{\omega_{2k}},
\]
where \([\cdot], [\cdot]\) denote the ceiling function and the integer part function correspondingly.

For positive integers \( m < n \), write \( \omega_m^n \) for the word \( \omega_m \omega_{m+1} \cdots \omega_n \). For \( i, j \in \{0, 1\} \) and \( \omega \in \Sigma \), denote
\[
N_i(\omega_m^n) = \sharp \{ m \leq k \leq n : \omega_k = i \},
\]
and
\[
N_{ij}(\omega_m^n) = \sharp \{ m \leq k \leq n : \omega_k\omega_{2k} = ij \}.
\]
We also denote
\[
N_{i,\text{odd}}(\omega_m^n) = \sharp \{ m \leq k \leq n : k \text{ odd}, \ \omega_k = i \}.
\]
Then we have

\[ \mu_{p,q}(C_n(\omega)) = (1 - p)^{N_{0,\text{odd}}}{p^{N_{1,\text{odd}}}}(1 - p)^{N_{00}}p^{N_{01}}(1 - q)^{N_{10}}q^{N_{11}}, \]

with \( N_{i,\text{odd}} = N_{i,\text{odd}}(\omega^n) \), and \( N_{ij} = N_{ij}(\omega_n^{n/2}) \). Thus

\[ -\log \mu_{p,q}(C_n(\omega)) = -\frac{1}{2}\left( \frac{N_{0,\text{odd}}}{n/2} \log(1 - p) + \frac{N_{1,\text{odd}}}{n/2} \log p \right. \]

\[ \left. + \frac{N_{00}}{n/2} \log(1 - p) + \frac{N_{01}}{n/2} \log p \right) + \frac{N_{10}}{n/2} \log(1 - q) + \frac{N_{11}}{n/2} \log q \right). \]

(3.1)

Lemma 3.1. If \( p = (2\theta - \alpha)/(2 - \theta) \) and \( q = \alpha/\theta \), then

\( \mu_{p,q}(E_\theta \cap A_\alpha) = 1. \)

Proof. Denote

\[ x_n(\omega) = \frac{N_1(\omega_n^{n/2} + 1)}{n/2} = \frac{2}{n} \sum_{k=n/2+1}^{n} \omega_k. \]

By the Law of Large Numbers, for \( \mu_{p,q} \)-almost all \( \omega \)

\[ x_{2n}(\omega) = \frac{p}{2} + \frac{x_n(\omega)}{2}q + \frac{1 - x_n(\omega)}{2}p + o(1) = p + \frac{x_n(\omega)}{2}, \frac{q - p}{2} + o(1). \]

Note that \(|\frac{x_n^2}{2}| < 1\). Then, as \( k \to \infty \),

\[ x_{2n}(\omega) \to \frac{2p}{2 + p - q}. \]

By [PS12, Lemma 5], it implies that \( \mu_{p,q} \)-almost surely

\[ \lim_{n \to \infty} x_n(\omega) = \frac{2p}{2 + p - q} = \theta, \]

(3.2)

where the last equality comes from the choices of \( p \) and \( q \). Thus \( \mu_{p,q}(E_\theta) = 1. \)

On the other hand, by applying the Law of Large Numbers again, for \( \mu_{p,q} \)-a.e. \( \omega \),

\[ \frac{2}{n} \sum_{k=n/2+1}^{n} \omega_k = x_n(\omega)(1 + o(1)) \to q\theta = \alpha. \]

By [PS12, Lemma 5], we conclude \( \mu_{p,q}(A_\alpha) = 1. \)

\[ \square \]

Lemma 3.2. For \( p = (2\theta - \alpha)/(2 - \theta) \) and \( q = \alpha/\theta \), we have

\[ h_{\mu_{p,q}} = (1 - \frac{\theta}{2})H(\frac{2\theta - \alpha}{2 - \theta}) + \frac{\theta}{2}H(\frac{\theta - \alpha}{\theta}). \]
Proof. By (3.1), we have for $\mu_{p,q}$ almost all $w \in \Sigma$,

$$h_{\mu_{p,q}} = \lim_{n \to \infty} \frac{-\log \mu_{p,q}(C_n(w))}{n}$$

$$= -\frac{1}{2} \left( (1 - p) \log(1 - p) + p \log p + (1 - \theta)(1 - p) \log(1 - p) \right.$$ 

$$+ (1 - \theta)p \log p + \theta(1 - q) \log(1 - q) + \theta q \log q \right)$$

$$= \frac{1}{2} \left( (2 - \theta)H(p) + \theta H(q) \right)$$

$$= (1 - \frac{\theta}{2})H(\frac{2\theta - \alpha}{2 - \theta}) + \frac{\theta}{2} H(\frac{\theta - \alpha}{\theta}).$$

\[ \square \]

**Lemma 3.3.** If $\theta \notin [\alpha, (2 + \alpha)/3]$ we have $E_\emptyset \cap A_\alpha = \emptyset$, otherwise for $p = (2\theta - \alpha)/(2 - \theta)$ and $q = \alpha/\theta$, we have for all $x \in E_\emptyset \cap A_\alpha$,

$$\lim_{n \to \infty} \frac{-\log \mu_{p,q}(C_n(w))}{n} = (1 - \frac{\theta}{2})H(\frac{2\theta - \alpha}{2 - \theta}) + \frac{\theta}{2} H(\frac{\theta - \alpha}{\theta}).$$

**Proof.** Observe that for any $x \in E_\emptyset \cap A_\alpha$, for any small $\varepsilon > 0$, for $n$ large enough, we have

$$N_1(\omega_{n/2}^n) \in [\frac{\varepsilon n}{2}(1 - \varepsilon), \frac{\varepsilon n}{2}(1 + \varepsilon)]$$

$$N_1(\omega_n^{2n}) \in [\varepsilon n(1 - \varepsilon), \varepsilon n(1 + \varepsilon)]$$

$$N_{11}(\omega_n^{2n}) \in [\frac{\alpha n}{2}(1 - \varepsilon), \frac{\alpha n}{2}(1 + \varepsilon)].$$

The obvious inequalities $N_{11}(\omega_n^{2n}) \leq N_1(\omega_n^{n/2})$ and $N_1(\omega_n^{n/2}) - N_{11}(\omega_n^{2n}) \leq n/2 + N_0(\omega_{n/2}^n) = n - N_1(\omega_{n/2}^n)$ imply $\theta \in [\alpha, (2 + \alpha)/3]$. Furthermore, we have

$$\log \mu_{p,q}(C_{2n}(\omega)) - \log \mu_{p,q}(C_n(\omega))$$

$$= N_{11}(\omega_n^{2n}) \log q + (N_1(\omega_n^{n/2}) - N_{11}(\omega_n^{2n})) \log(1 - q)$$

$$+ (N_1(\omega_n^{2n}) + N_{11}(\omega_n^{2n})) \log p$$

$$+ (n - N_1(\omega_n^{n/2}) - N_1(\omega_n^{2n}) + N_{11}(\omega_n^{2n})) \log(1 - p)$$

$$= n \left( (1 - \frac{\theta}{2})H(\frac{2\theta - \alpha}{2 - \theta}) + \frac{\theta}{2} H(\frac{\theta - \alpha}{\theta}) + \varepsilon \cdot O(1) \right).$$

Hence by the same argument of the proof of Lemma 2.2, we have for all $x \in E_\emptyset \cap A_\alpha$,

$$\lim_{n \to \infty} \frac{-\log \mu_{p,q}(C_n(w))}{n} = (1 - \frac{\theta}{2})H(\frac{2\theta - \alpha}{2 - \theta}) + \frac{\theta}{2} H(\frac{\theta - \alpha}{\theta}).$$

\[ \square \]
References


