BERNOULLI CONVOLUTIONS AND AN INTERMEDIATE VALUE THEOREM FOR ENTROPIES OF K-PARTITIONS

ELON LINDENSTRAUSS, YUVAL PERES, AND WILHELM SCHLAG

ABSTRACT. We establish a strong regularity property for the distributions of the random sums $\sum \pm \lambda^n$, known as “infinite Bernoulli convolutions.” For a.e. $\lambda \in (1/2, 1)$ and any fixed $\ell$, the conditional distribution of $(\omega_{n+1}, \ldots, \omega_{n+\ell})$ given the sum $\sum_{n=0}^{\infty} \omega_n \lambda^n$ tends to the uniform distribution on $\{-1\}^\ell$ as $n \to \infty$. More precise results, where $\ell$ grows linearly in $n$, and extensions to other random sums, are also obtained. As a corollary, we show that a Bernoulli measure-preserving system of entropy $h$, has $K$-partitions of any prescribed conditional entropy in $[0, h]$. This answers a question of Rokhlin and Sinai from the 1960’s, for the case of Bernoulli systems.

1. INTRODUCTION

The impetus for this work was the following question, suggested to us by Ya. G. Sinai:

Given a probability measure preserving system $X = (X, A, \mu, T)$, what are the possible values of the conditional entropy for $K$-partitions in $X$?

In other words, consider all the realizations of $X$ as real-valued stationary processes $\{Y_n\}_{n=\infty}^{\infty}$ such that the left tail $\sigma$-algebra

$$\bigwedge_{n=\infty}^0 \sigma(\ldots, Y_{n-1}, Y_n)$$

is trivial. What are the possible values for the conditional entropy

$$h(Y_n | Y_{n-1}, Y_{n-2}, \ldots) := \mathbb{E} \left[ \log \mathbb{P}(Y_n = \cdot | Y_{n-1}, Y_{n-2}, \ldots) \right] \ ?$$

(1.1)

This question was raised by Sinai and Rokhlin in the early days of entropy theory, and appears along with many other questions (most of which have been transformed into well-known theorems by now) in a survey paper by Rokhlin ([19], paragraph 12.7). When a realization as in the above question exists, we call the $\sigma$-algebra $\xi = \sigma(Y_n; n < 0)$ a $K$-partition and say that $X$ is a $K$-system; see Section 6 for an equivalent definition. If the variables $Y_n$ take values in a countable set, then the conditional entropy (1.1) equals the Kolmogorov-Sinai entropy $h(\mu, T)$, so the question above is interesting only when the range of $Y_n$ is uncountable. (However, if $h(\mu, T) < \infty$, then the conditional distribution of $Y_n$ given $Y_{n-1}, Y_{n-2}, \ldots$ must be countably supported almost surely.) In this case the conditional entropy (1.1) can be lower than $h(\mu, T)$: a classical example due to Rokhlin (see [9]) involving toral automorphisms yield conditional entropies of the form $h(\mu, T) - \log(\gamma)$ where $\gamma > 1$ is an algebraic number, but these examples only yield countably many entropy values.

The authors were partially supported by NSF grants #DMS-9729992 (E. L.), #DMS-980397 (Y. P.), and #DMS-0070588 (W. S.).
In this paper we give a complete answer to Sinai’s question for the case where \( X \) is a Bernoulli system (see Theorem 1.4 below), using certain concrete linear filters that are of great intrinsic interest. Very recently, A. Vershik [27] discovered that results of F. den Hollander and J. E. Steif [7] yield the existence of \( K \)-partitions of arbitrary entropy in an infinite entropy Bernoulli system (see also [25], [26]).

Denote by \( \mu_p \) the biased Bernoulli measure
\[
\mu_p = \prod_{n \in \mathbb{Z}} (p \delta_1 + (1 - p) \delta_{-1}).
\]
on \( X = \{\pm 1\}^\mathbb{Z} \), and write \( H(p) = -p \log p - (1 - p) \log(1 - p) \).

**Theorem 1.1.** For \( \omega = \{\omega_n\}_{n \in \mathbb{Z}} \in X \) and \( \lambda \in [1/2, 1) \), denote
\[
Y_n^{(\lambda)}[\omega] = \sum_{k=0}^{\infty} \omega_{n+k} \lambda^k.
\]
For a.e. \( \lambda \in (1/2, 1) \) and for all \( p \in [\frac{1}{3}, \frac{2}{3}] \) with \( H(p) > |\log \lambda| \), if \( \omega \) has law \( \mu_p \), then the stationary process \( \{Y_n^{(\lambda)}[\omega]\}_{n \in \mathbb{Z}} \) has trivial left tail and conditional entropy
\[
(1.2)
\]
\[
h(Y_n^{(\lambda)} | Y_{n-1}^{(\lambda)} , \ldots ) = H(p) + \log \lambda.
\]
Note that the set of exceptional \( \lambda \) does not depend on \( p \). It follows directly from the definitions that the process in the theorem also has a trivial right tail for all \( \lambda \), but the two sided tail is the full \( \sigma \)-algebra except for certain algebraic \( \lambda \). Also, the conditional entropy given the future satisfies
\[
(1.4)
\]
\[
h(Y_n^{(\lambda)} | Y_{n+1}^{(\lambda)} , \ldots ) = h(\mu_p) = H(p).
\]
In particular, for each \( p \in [\frac{1}{3}, \frac{2}{3}] \) and for a.e. \( \lambda \in (e^{-H(p)}, 1) \), the process \( \{Y_n^{(\lambda)}[\omega]\}_{n \in \mathbb{Z}} \) is more predictable from the past than from the future, although both its left and right tails are trivial.

Let \( \xi_\lambda = \sigma(Y_n^{(\lambda)} ; n < 0) \). Then (1.3) can be written in the form
\[
(1.5)
\]
\[
h(\xi_\lambda | T^{-1} \xi_\lambda ) = H(p) + \log \lambda.
\]
The difficult part of the proof of Theorem 1.1 is establishing tail triviality. We obtain this from the following significantly stronger statement, pertaining to the distribution of the random sums \( \sum_{0}^{\infty} \pm \lambda^k \).

**Notation.** For \( \omega = \{\omega_n\}_{n \in \mathbb{Z}} \in \{\pm 1\}^\mathbb{Z} \) and \( \lambda \in [1/2, 1) \), denote
\[
(1.6)
\]
\[
\Pi_\lambda(\omega) = Y_0^{(\lambda)}[\omega] = \sum_{k=0}^{\infty} \omega_k \lambda^k.
\]
Also, abbreviate \( \omega_n^{\infty+\ell} := (\omega_n, \ldots, \omega_n+\ell) \).

**Theorem 1.2.** For a.e. \( \lambda \in (1/2, 1) \) and for all \( p \in [\frac{1}{3}, \frac{2}{3}] \) with \( H(p) > |\log \lambda| \), if \( \omega \) has law \( \mu_p \), then there exists \( \rho > 0 \) such that almost surely
\[
(1.7)
\]
\[
\sup_{1 \leq \ell \leq n} \sup_{\omega \in \{\pm 1\}^{\ell+1}} \left| \frac{\mu_p[\omega_n^{\infty+\ell} = \theta | \Pi_\lambda(\omega)]}{\mu_p[\omega_n^{\infty+\ell} = \theta]} - 1 \right| \rightarrow 0 \text{ as } n \to \infty.
\]
In particular, for such \( \lambda, p \) and any fixed \( \ell \geq 1 \), every \( \theta \in \{ \pm 1 \}^{\ell+1} \) satisfies
\[
\mu_p[\omega_n^{n+\ell} = \theta | \Pi_\lambda(\omega)] \rightarrow \mu_p[\omega_0^\ell = \theta] \text{ as } n \rightarrow \infty.
\]

The restriction \( p \in \left[ \frac{1}{3}, \frac{2}{5} \right] \) is of a technical nature and some improvements are possible there. However, Theorems 1.1 and 1.2 cannot be strengthened to hold for all \( \lambda \in (1/2, 1) \):

**Theorem 1.3.** If \( \lambda^{-1} \) is a Pisot number then \( \xi_\lambda \) is not a \( K \)-partition, and (1.8) fails. If \( \lambda^{-1} \) is a Salem number then (1.7) fails.

For background on Pisot and Salem numbers, see [2]. We do not know whether \( \xi_\lambda \) is a \( K \)-partition when \( \lambda^{-1} \in (1, 2) \) is a Salem number.

In the proof of Theorem 1.2, we use in an essential way the theory of Bernoulli convolutions \( \nu_\lambda^p \) and especially the recent results from [13]. In this paper we go beyond [13] in several ways:

- If \( p \neq \frac{1}{2} \), then the measure \( \mu_p \) has a nondegenerate "multifractal spectrum" in the sense that
  \[
  \dim_H \{ x : -\frac{1}{n} \log \mu_p(\omega_n^p) = \alpha \}
  \]
  considered as a function of \( \alpha \), is nonzero on an interval (namely \( (\log(1-p), \log p) \) for \( p > \frac{1}{2} \)). Peres and Solomyak [16] found that this property is manifested by the existence of an interval (in this case the interval \( (e^{-\log(p)}, p^2 + (1 - p)^2) \) such that for a.e. \( \lambda \) in this interval one has \( \frac{d\nu_\lambda^p}{d\lambda} \notin L^n \), but a.e. \( \frac{d\nu_\lambda^p}{d\lambda} \in L^q \) for some \( q \in (1, 2) \), and in particular \( \nu_\lambda^p \) is absolutely continuous with respect to Lebesgue measure. They also gave a sharp threshold for the existence a.e. of \( L^q \) densities; for \( \mu_p \) this threshold is \( \lambda > [p^2 + (1 - p)^2]^{\frac{1}{2^n}} \). We show here that in fact for a.e. \( \lambda \) in the above range, \( \frac{d\nu_\lambda^p}{d\lambda} \) a.s. has fractional derivatives in \( L^q \). This extends [13] where only the \( L^2 \) case is treated.

- In [13], it is proved that for every \( p \in \left[ \frac{1}{3}, \frac{2}{5} \right] \) the density \( \frac{d\nu_\lambda^p}{d\lambda} \) is in \( L^{2^n} \) for some explicit \( \gamma > 0 \) depending on \( \lambda \) and \( p \) for almost all \( \lambda \in (p^2 + (1 - p)^2, 1) \). Here \( L^q \) is Lebesgue measure on \( \mathbb{R} \) and \( L^{2^n} \) are the standard Sobolev spaces, see (2.1) below. We strengthen this by showing that the set of those \( \lambda \) where this fails for some choice of \( p \in \left[ \frac{1}{3}, \frac{2}{5} \right] \) is also of measure zero. A similar statement holds for the interval of \( \lambda \) for which a.s. \( \frac{d\nu_\lambda^p}{d\lambda} \) has fractional derivatives in \( L^q \).

Using the \( K \)-partitions we constructed from biased Bernoulli convolutions as a basic building block, and applying Ornstein’s isomorphism theorem [12] (see also [20]), we deduce:

**Theorem 1.4.** Any Bernoulli system of entropy \( h \in (0, \infty) \) has \( K \)-partitions of (conditional) entropy \( \eta \) for any \( \eta \in (0, h) \).

It seems that our method of proving this theorem can be extended to general \( K \)-systems, but this will require a more elaborate construction which we defer to a future paper. For Bernoulli systems, we have an alternate proof of Theorem 1.4, using a construction similar to Rokhlin’s, and a theorem of S. G. Dani [4]. This proof is sketched at the end of §8.

2. Regularity for projections of symbolic multifractals

It is an easy consequence of the Kolmogorov zero-one law that for every $\lambda \in (0, 1)$ the distribution $\nu_\lambda$ of the random sum $\sum_0^\infty \pm \lambda^n$ with independently chosen signs is either purely absolutely continuous or purely singular with respect to the Lebesgue measure. For the case of choosing the signs with equal probability $\frac{1}{2}$, it is a classical problem to determine for which $\lambda$ these properties hold. We shall not recount the history of this problem or its many contributors, but rather refer the reader to the survey article [14]. A key idea in this context is the notion of transversality introduced by Pollicott and Simon [18], and Solomyak [22], [23]. For our purposes we shall need the following result of Solomyak [22]:

$$\frac{d\nu_\lambda}{dx} \in L^2(\mathbb{R})$$

for a.e. $\lambda \in (1/2, 1)$ one has $\frac{d\nu_\lambda}{dx} \in L^2(\mathbb{R})$

as well as a stronger statement obtained in [13]:

$$\frac{d\nu_\lambda}{dx} \in L^{2\gamma}(\mathbb{R})$$

Here $L^{2\gamma}(\mathbb{R})$ is the Sobolev space with norm

$$\|f\|_{2\gamma}^2 = \int_{-\infty}^\infty |\hat{f}(\xi)|^2 |\xi|^{2\gamma} d\xi.$$

The existence of some small amount of differentiability of $\frac{d\nu_\lambda}{dx}$ will be essential for showing that the $\xi_\lambda$ defined above are indeed a K-partition for a.e. $\lambda$. In fact, we shall need these properties not just for $\nu_\lambda$, but also for the random sum in which a fixed number of digits are fixed. Since estimates of this type were not considered in [22] or [13], we will derive them here from one of the main results in [13]. We begin by recalling some terminology from [13].

**Definition 2.1.** Let $(\Omega, D)$ be a compact metric space, $J \subset \mathbb{R}$ an open interval, and let $\Pi : J \times \Omega \to \mathbb{R}$ be a continuous map. We assume that for any compact $I \subset J$ and $\ell \geq 0$ there exists a constant $C_{\ell,I}$ such that

$$\left| \frac{d^\ell}{d\lambda^\ell} \Pi(\lambda, \omega) \right| \leq C_{\ell,I}$$

for all $\lambda \in I$ and $\omega \in \Omega$. Given any finite measure $\mu$ on $\Omega$ let $\nu_\lambda = \mu \circ \Pi^{-1}$, where $\Pi(\lambda, \cdot) = (\lambda, \cdot)$. The $\alpha$-energy of $\mu$ is defined as $E_\alpha(\mu, D) = \int_{\Omega} \int_{\Omega} \frac{d\mu(\omega_1) d\mu(\omega_2)}{|\lambda_1 - \lambda_2|^{\alpha}}$. For any distinct $\omega_1, \omega_2 \in \Omega$ and $\lambda \in J$ let

$$\Phi_\lambda(\omega_1, \omega_2) = \frac{\Pi(\lambda, \omega_1) - \Pi(\lambda, \omega_2)}{D(\omega_1, \omega_2)}.$$

For any $\beta \in [0, 1)$ we say that $J$ is an interval of transversality of order $\beta$ for $\Pi$ if there exists a (small) constant $C_\beta$ so that for all $\lambda \in J$ and $\omega_1, \omega_2 \in \Omega$ the condition $|\Phi_\lambda(\omega_1, \omega_2)| \leq C_\beta D(\omega_1, \omega_2)^{\beta}$ implies

$$\left| \frac{d}{d\lambda} \Phi_\lambda(\omega_1, \omega_2) \right| \geq C_\beta D(\omega_1, \omega_2)^{\beta}.$$

For such families of projections it was shown in [13] that for typical $\lambda$ most $\nu_\lambda$ have as much regularity as the original measure $\mu$. This can be stated as follows, see Theorem 2.8 in [13].
Theorem 2.2. Let \( \Omega, J, \Pi \) be as in Definition 2.1 and suppose that \( J \) is an interval of transversality of order \( \beta \) for \( \Pi \). Let \( \mu \) be a positive measure on \( \Omega \) with finite \( \alpha \)-energy for some \( \alpha > 0 \). Then for any compact \( I \subset J \)

\[
\int_I \|r^2 d\lambda\|_{2\gamma}^{2\gamma} d\lambda \leq C_\gamma E_\alpha(\mu, D) \quad \text{if} \quad 0 < (1 + 2\gamma)(1 + a_0 \beta) \leq \alpha,
\]

where \( a_0 \) is some absolute constant.

We shall apply this theorem to \( \Omega = \Sigma^{2+} \) where \( \Sigma \) is an alphabet with \( m \) symbols, and the map \( \Pi_\lambda : \Omega \to \mathbb{R} \) given by

\[
\Pi_\lambda(\omega) = \sum_{j=0}^{\infty} d_{2j}(\lambda)\lambda^j.
\]

For the applications in this paper we will always take \( d_k(\lambda) \) as constants, but in general they are just required to be \( C^\infty \) functions. In fact, a finite number of derivatives suffices depending on \( m \) and \( J \).

For any distinct \( \omega, \tau \in \Omega \) let

\[
|\omega \wedge \tau| = \min\{i \geq 0 \mid \omega_i \neq \tau_i\}.
\]

Fix some interval \( J = [\lambda_1, \lambda_2] \subset (0, 1) \). Theorem 2.2 was applied in [13] with the metric \( D(\omega, \tau) = \lambda_2^{\omega \wedge \tau} \). Here, it will be more convenient to rewrite the preceding theorem in terms of a slightly different metric, \( d(\omega, \tau) = \lambda_2^{\omega \wedge \tau} \). We will abbreviate \( E_\alpha(\mu, d) \) as \( E_\alpha(\mu) \). In order to apply Theorem 2.2 we need transversality of the family \( \Pi_\lambda \). First we recall the notion of transversality most useful for maps from symbolic spaces. Say that the transversality condition holds on an interval \( J \subset (0, 1) \) if

\[
\lambda \mapsto \Pi_\lambda(\omega) - \Pi_\lambda(\tau) \quad \text{has no double zeros on} \quad J \quad \text{for} \quad \omega \neq \tau.
\]

Here \( \lambda_0 \) is a “double zero” for \( f \) if and only if \( f(\lambda_0) = f'(\lambda_0) = 0 \) (so it includes zeros of higher order as well). The following lemma establishes the connection between Definitions 2.1 and this definition.

Lemma 2.3. Suppose the transversality condition is satisfied on the interval \( J = [\lambda_1, \lambda_2] \). Then \( J \) is an interval of transversality of order \( \beta \) for \( \Pi \) if \( \lambda_1 > \lambda_2^{1+\beta} \).

We omit the proof of this simple lemma, see [13, Lemma 5.3]. Thus for the projections from symbolic spaces to \( \mathbb{R} \), Theorem 2.2 implies the following result.

Theorem 2.4. Let \( \mu \) be a measure on \( \Omega = \Sigma^{2+} \) with finite \( \alpha \)-energy for the metric \( d(\omega, \tau) = \lambda_2^{\omega \wedge \tau} \) where \( \alpha > 1 + 2\gamma \) for some \( \gamma > 0 \). Suppose that \( [\lambda_1, \lambda_2] \) is an interval of transversality for \( \Pi_\lambda \). Then

\[
\int_{\lambda_1}^{\lambda_2} \|d\lambda\|_{2\gamma}^{2\gamma} d\lambda \leq C E_\alpha(\mu)
\]

where \( C \) does not depend on \( \mu \).

Proof. Fix some sufficiently small \( \beta > 0 \). Partition \( J = [\lambda_1, \lambda_2] \) into subintervals \( J_i = [\lambda_i^+, \lambda_{i+1}^+] \) for \( i = 0, \ldots, m \) so that \( \lambda_i^+ \geq (1 + \beta)(\lambda_{i+1}^+) \). By Lemma 2.3 all \( J_i \) are intervals of transversality of order \( \beta \) for \( \Pi \). Notice that the metric on \( \Omega \) depends on \( i \), in fact \( d_i(\omega, \tau) = (\lambda_{i+1})^{\omega \wedge \tau} \).
In view of our assumption, \( \mu \) has finite \( \alpha \)-energy with respect to \( d_i \) for each \( i \). Theorem 2.4 therefore follows from Theorem 2.2 provided \( \beta > 0 \) is small enough (depending on \( \alpha \) and \( \gamma \)).

Our next goal is to prove \( L^q \)-smoothness properties of the densities \( \frac{d\mu}{dx} \) for \( 1 < q \leq 2 \). This improves upon results in [16] and [13]. Before formulating our new result, we need a few definitions. Let \( \mathcal{W}_k \) denote cylinder sets with prescribed coordinates \( \omega_j^{k-1} \), i.e., all sets of the form

\[
W(\omega_0^{k-1}) := \{ x \mid \tau_j = \omega_j, \; 0 \leq j < k \}.
\]

For an arbitrary measure \( \mu \) on \( m \) symbols and \( q > 1 \) let

\[
D_q(\mu) = \frac{1}{q-1} \liminf_{k \to \infty} \frac{1}{k} \log \sum_{W \in \mathcal{W}_k} [\mu(W)]^q
\]

be the \( L^q \)-dimension of \( \mu \) (note that we deviate from the notations of [16] by a factor of \( \log m \)). A related notion is the \( q \)-energy \( \mathcal{E}_{q0}(\mu) \) of \( \mu \) defined as follows:

\[
\mathcal{E}_{q0}(\mu) = \int_0^\infty \int \frac{\mu(B(x, r))^{q-1}}{r^{\alpha(q-1)}} \frac{dr}{r} d\mu(x)
\]

Up to a bounded factor,

\[
\mathcal{E}_{q0}(\mu) \asymp \mathcal{E}_{q0}(\mu) := \sum_{k=0}^\infty \lambda_1^{-k(q-1)\alpha} \sum_{W \in \mathcal{W}_k} (\mu(W))^q.
\]

Note that if \( \lambda_1 > e^{-D_0(\mu)} \), then \( \mathcal{E}_{q0}(\mu) < \infty \). In Theorem 4.1 of [15] Peres and Solomyak prove that for any \( q \in (1, 2] \) and a.e. \( \lambda \) in an interval of transversality, one has

\[
\lambda > e^{-D_0(\mu)} \implies \frac{d\lambda}{dx} \in L^q(\mathbb{R}).
\]

Moreover, the condition on the left-hand side of (2.7) is optimal. For the Bernoulli measures \( \mu \) on \( m \) symbols \( \{b_1, \ldots, b_m\} \) with \( \mu(b_j) = p_j \) and \( \sum_j p_j = 1 \) one easily obtains that

\[
e^{-D_0(\mu)} = \left[ \sum_{j=1}^m p_j q \right]^{-\frac{1}{q-1}}.
\]

Of course, \( D_q(\mu) \) increases as \( q \) decreases and \( D_1(\mu) = \lim_{q \to 1+} D_q(\mu) = h(\mu) = \sum_{j=1}^m -p_j \log p_j \).

In order to formulate the following result, we need to introduce the Sobolev spaces \( L^{q, \gamma} \). They are defined via the norm

\[
\|f\|_{q, \gamma} = \left\| (I - \nabla)^{\frac{\gamma}{2}} f \right\|_q = \left\| \left( 1 + 4\pi^2 |x|^2 \right)^{\frac{\gamma}{2}} f \right\|_q.
\]

There is a natural connection between these Sobolev spaces and the modulus of continuity. More precisely, let \( \omega_q(f, t) = \|f(\cdot + t) - f(\cdot)\|_q \) be the \( L^q \)-modulus of continuity. Then, for any \( 1 < q \leq 2 \)

\[
\|f\|_{q, \gamma}^q \leq C \int_\mathbb{R} \frac{[\omega_q(f, t)]^q}{|t|^{1+\gamma}} dt.
\]

We prove the easier converse, i.e.,

\[
\omega_q(f, t) \leq C |t|^{-\gamma} \|f\|_{q, \gamma}
\]
in Lemma 2.9 below. For these facts, as well as more background, see Chapter V, §3 in Stein [24]. We can now state the $L^2$-smoothness theorem.

**Theorem 2.5.** Let $\mu$ be a positive measure on $\Omega = \Sigma^\mathbb{Z}_+$. Let $1 < q \leq 2$ and assume that $E_{q,\alpha}(\mu) < \infty$ for some $\alpha > 1 + \frac{2q}{q-1}$ where $\gamma > 0$, with the metric being $d(\omega, \zeta) = |\omega - \zeta|$. If $[\lambda_1, \lambda_2]$ is an interval of transversality, then

$$\int_{\lambda_1}^{\lambda_2} \left\| \frac{d\lambda}{dx} \right\|_{q,\gamma} \, d\lambda \leq C E_{q,\alpha}(\mu).$$

**Remark:** By (2.11), Theorem 2.5 implies

$$\int_{\lambda_1}^{\lambda_2} \left\| \frac{d\lambda}{dx} - \frac{d\lambda}{dx}(\cdot + h) \right\|_{q,\gamma} \, d\lambda \leq C|h|^{\gamma} E_{q,\alpha}(\mu).$$

In order to prove this result we need to recall some of the details from [13], see also [14]. Firstly, let $\psi \in C^\infty(\mathbb{R})$ have the property that $\hat{\psi} \geq 0$ and

$$\text{supp}(\hat{\psi}) \subset \{ \xi \in \mathbb{R} \mid 1 \leq |\xi| \leq 4 \}$$

$$1 = \sum_{j=-\infty}^{\infty} \hat{\psi}(2^{-j}\xi) \text{ for all } \xi \neq 0.$$

This is referred to as a Littlewood-Paley decomposition, and the existence of such a function $\psi$ is standard, see [24]. Moreover, $\psi$ can be chosen so that $\psi = \phi * \phi$ for some $\phi \in C^\infty$ with $\hat{\phi} \geq 0$ and with both $\psi$ and $\phi$ being real-valued. Notice that there is some absolute constant $C > 0$ such that

$$C^{-1} \leq \sum_{j=-\infty}^{\infty} \hat{\phi}(2^{-j}\xi) \leq C \text{ for all } \xi \neq 0.$$

Let $\psi_j(x) = 2^j \psi(2^j x)$ and $\phi_j(x) = 2^j \phi(2^j x)$. The two important features of $\psi$ are its rapid decay and the fact that all its moments vanish. These two properties lead to the following basic estimate.

**Lemma 2.6 (Lemma 4.6, [13]).** Assume that $J$ is an interval of transversality of order $\beta$ for $\Pi_\lambda$. Suppose $\rho \in C^\infty(\mathbb{R})$ is supported on $J$. For any distinct $\omega, \zeta \in \Omega$, any integer $j$, and any positive integer $N$,

$$\left| \int_{\lambda_1}^{\lambda_2} \rho(\lambda) \psi(2^j[\Pi_{\lambda}(\omega) - \Pi_{\lambda}(\zeta)]) \, d\lambda \right| \leq C_N (1 + 2^j d(\omega, \zeta)^{1+\alpha \beta})^{-N}$$

where $C_N$ depends only on $N, \rho, \beta$, and $\alpha$ and $\alpha_0$ is some absolute constant.

The intuition underlying (2.15) is the following: If $|\Pi_{\lambda}(\omega) - \Pi_{\lambda}(\zeta)|$ is large, then the integrand in (2.15) is small. If, on the other hand, $|\Pi_{\lambda}(\omega) - \Pi_{\lambda}(\zeta)|$ is small, then by transversality the derivative of $\Pi_{\lambda}(\omega) - \Pi_{\lambda}(\zeta)$ has to be large. A change of variables therefore linearizes the argument of $\psi$ as a function of $\lambda$, and the large amount of cancellation in $\psi$ (vanishing moments) then shows that the integral is small. See [13] or [14] for details.
Lemma 2.7. Let \( \mu \) be a complex measure on the symbol space \( \Omega = \Sigma^{\mathbb{Z}+} \). Let \( \rho \in C^\infty(\mathbb{R}) \) be a nonnegative function supported on the interval of \( \beta \)-transversality \([\lambda_1, \lambda_2] \). Then for any \( 1 \leq q \leq 2 \)

\[
\int \left\| \phi_j * \Pi_\lambda(\mu) \right\|_q^q \rho(\lambda) \, d\lambda \leq C 2^{j(q-1)} \sum_{W \in \mathcal{W}_k} |\mu|(W)^q
\]

where \( k \) is given by \( 2^j \asymp \lambda_1^{-k(1+\alpha_0 \beta)} \).

Proof. Define a linear operator \( T \) from \( \mathbb{R}^N \) to the continuous functions on \( \mathbb{R}^2 \), which we view as a subspace of \( L^1(\mathbb{R}^2; dx \rho(\lambda) \, d\lambda) \), by

\[
T(\bar{a}) = \phi_j * \Pi_\lambda(\mu^\alpha)
\]

where \( \bar{a} = \{a_W\}_{W \in \mathcal{W}_k} \) and \( \mu^\alpha \) is the measure

\[
\mu^\alpha(A) = \sum_{W \in \mathcal{W}_k} a_W \frac{\mu(A \cap W)}{|\mu|(W)}.
\]

Note that \( \mu = \mu^\alpha \) for \( \bar{a} \) given by \( a_W = |\mu|(W) \) for all \( W \in \mathcal{W}_k \). To obtain (2.16) we shall interpolate between \( L^1 \) and \( L^2 \). We start with the simple \( L^1 \)-bound: Since \( \|\phi_j\|_1 = \|\phi\|_1 \)

\[
\left\| T(\bar{a}) \right\|_1 = \int \left\| \phi_j * \Pi_\lambda(\mu^\alpha) \right\|_1 \rho(\lambda) \, d\lambda \leq \int \|\phi\|_1 \left\| \Pi_\lambda(\mu^\alpha) \right\| \rho(\lambda) \, d\lambda
\]

\[
\leq C \|\mu^\alpha\|_1 \leq C \sum_{W \in \mathcal{W}_k} |a_W| = C \|\bar{a}\|_{l^1}.
\]

For the \( L^2 \)-bound, observe firstly that

\[
\left\| T(\bar{a}) \right\|_2^2 = \int \int \left| \phi_j * \Pi_\lambda(\mu^\alpha)(x) \right|^2 \, dx \, \rho(\lambda) \, d\lambda
\]

\[
= \int \int 2^j (\phi * \phi) \left( 2^j (\Pi_\lambda(\omega) - \Pi_\lambda(\bar{a})) \right) \rho(\lambda) \, d\lambda \, d\mu^\alpha(\omega) \, d\mu^\alpha(\bar{a}).
\]

Since \( \phi * \phi = \psi \), Lemma 2.6 implies that

\[
\left\| T(\bar{a}) \right\|_2^2 \leq C_N 2^j \int \int (1 + 2^j d(\omega, \bar{a})^{1+\alpha_0})^{-N} d\mu^\alpha(\omega) \, d\mu^\alpha(\bar{a})
\]

\[
\leq C_N 2^j \sum_{\ell=0}^\infty (1 + 2^j \lambda_1^{-\ell(1+\alpha_0 \beta)})^{-N} \, |\mu^\alpha| \times |\mu^\alpha| ([\omega, \bar{a}] \in \Omega^2 | |\omega \wedge \bar{a}| = \ell]
\]

\[
\leq C_N 2^j \sum_{\ell=0}^k 2^{-\ell N} \lambda_1^{-\ell(1+\alpha_0 \beta)} \sum_{W \in \mathcal{W}_k} |\mu^\alpha|(W)^2
\]

\[
\leq C_N 2^j \sum_{\ell=0}^k 2^{-\ell N} \lambda_1^{-\ell(1+\alpha_0 \beta)} \sum_{W \in \mathcal{W}_k} |\mu^\alpha|(W)^2 |\Sigma|^{k-\ell}
\]

\[
(2.17)
\]

\[
\leq C_N 2^j (1+\bar{a})^{-\ell(1+\alpha_0 \beta)} \sum_{W \in \mathcal{W}_k} |a_W|^2 \leq 2^j \|\bar{a}\|_{l^2}^2.
\]

\[
(2.18)
\]
To pass from (2.17) to (2.18) requires \( \lambda_1^{-1(1+\alpha_0\beta))N} > |\Sigma| \), which holds for sufficiently large \( N \) that does not depend on \( \mu \). The lemma now follows by the Riesz-Thorin interpolation theorem. \( \square \)

Lemmas 2.8 and 2.9 below are standard facts whose proofs are given for the reader’s convenience.

**Lemma 2.8.** Let \( 1 < q \leq 2 \). Then for any \( \gamma \geq 0 \)

\[
\|f\|_{q,\gamma} \leq C_q \left( \sum_{j=0}^{\infty} 2^{j\gamma q} \|\phi_j * f\|_q^q + \|f\|_q \right)
\]

for any \( f \in L^1(\mathbb{R}) \) for which the right-hand side is finite.

**Proof.** Let \( \chi \in C^\infty(\mathbb{R}) \) such that \( \hat{\chi} \in C^\infty(\mathbb{R}), \hat{\chi} \) is compactly supported, and such that \( \hat{\chi}(\xi) = 1 \) for \(-4 \leq \xi \leq 4\). By Young’s inequality

\[
\left\| (I - \Delta)^{\frac{\gamma}{2}} \hat{\chi} * f \right\|_q \leq \left\| (1 + 4\pi^2 |\xi|^2)^{\gamma/2} \hat{\chi} \right\|_q \|f\|_1 \leq C \|f\|_1.
\]

Moreover, with \( g = [(1 - \hat{\chi})f]^{\gamma} \), by the Littlewood-Paley theorem (cf. (2.14)),

\[
\left\| (I - \Delta)^{\frac{\gamma}{2}} g \right\|_q = \left\| \left( \sum_{j \in \mathbb{Z}_+} |\phi_j * (I - \Delta^{\frac{\gamma}{2}} g)|^2 \right)^{\frac{1}{2}} \right\|_q \\
\leq \left\| \left( \sum_{j \in \mathbb{Z}_+} |\phi_j * (I - \Delta^{\frac{\gamma}{2}} g)|^q \right)^{\frac{1}{q}} \right\|_q \\
= \int \left( \sum_{j=0}^{\infty} |\phi_j * (I - \Delta^{\frac{\gamma}{2}} g(x)|^q \right) \ dx \leq C \sum_{j=0}^{\infty} 2^{j\gamma q} \|\phi_j * f\|_q^q.
\]

To obtain the last inequality we used that

\[
(2.19) \quad \|\phi_j * (I - \Delta^{\frac{\gamma}{2}} g\|_q \leq C 2^{j\gamma q} \|\phi_j * f\|_q
\]

for all \( j \geq 0 \). To verify this, let \( \varphi \) be such that \( \text{supp}(\hat{\varphi}) \subset \{ \frac{1}{2} < |\xi| < 8 \}, \hat{\varphi} \hat{\phi} = \hat{\phi} \), and with both \( \varphi \) and \( \hat{\varphi} \) smooth. Then with

\[
m_j(\xi) := (1 + 4\pi^2 |\xi|^2)^{\gamma/2}(1 - \hat{\chi})(\xi) \hat{\varphi}(2^{-j} \xi)
\]

the Fourier transform of the function on the left-hand side of (2.19) can be written as

\[
(2.20) \quad m_j \hat{\phi} j f = m_j \hat{\phi} j * f.
\]

It is now easy to see that

\[
\sup_{j \geq 0} 2^{-j\gamma} \|m_j \| \leq C
\]

so that (2.19) follows from Young’s inequality. Alternatively, it is clear that

\[
\|m_j\|_\infty + \|\xi m_j'(\xi)\|_\infty \leq C 2^{j\gamma}
\]

for all \( j \geq 0 \). Therefore, (2.19) follows from (2.20) and Marcinkiewicz’s multiplier theorem, see Stein [24], chapter IV Theorem 6. \( \square \)
Proof of Theorem 2.5. Choose \( \beta > 0 \) such that \( \alpha \geq (1 + a_0 \beta) \left[ 1 + \frac{\alpha}{q-1} \right] \). As in the proof of Theorem 2.4 we can restrict attention to the case where \([\lambda_1, \lambda_2]\) is an interval of \( \beta \)-transversality. Choose \( \rho \in C^\infty_0 \); \( \rho \geq 0 \) so that \( \rho = 1 \) on \([\lambda_1, \lambda_2]\) and such that \( \text{supp}(\rho) \) is an interval of \( \beta \)-transversality. By Lemmas 2.7 and 2.8

\[
\int \left\| \frac{d\lambda}{dx} \right\|_{q, \gamma}^q \rho(\lambda) \, d\lambda \leq C \int \left[ \sum_{j \geq 0} 2^{j+\gamma} \left\| \varphi_j * \Pi_\lambda(\mu) \right\|_{q, \gamma}^q + \left\| \Pi_\lambda(\mu) \right\|_{q}^q \right] \rho(\lambda) \, d\lambda
\]

\[
\leq C \sum_{k \geq 0} \lambda_1^{-k(1 + a_0 \beta)(\gamma+q-1)} \sum_{W \in \mathcal{W}_k} \mu(W)^q + C \|\mu\|_q^q
\]

(2.21)

To obtain (2.21) one uses that \( \|\mu\|_q \leq \mathcal{E}_{q, \lambda}(\mu) \) for any \( \alpha \). \( \Box \)

Lemma 2.9. For any \( 1 < q < \infty \) and \( 0 \leq \gamma \leq 1 \) there exist constants \( C_{q, \gamma} \) so that

\[
\|f - f(\cdot + h)\|_q \leq C_{q, \gamma} |h|^{\gamma} \|f\|_{q, \gamma}
\]

for any \( |h| \leq 1 \) and \( f \in L_{q, \gamma}^\lambda \).

Proof. It suffices to consider \( f \in C^\infty_0(\mathbb{R}) \cap L_{q, \gamma}^\lambda \). Fix such an \( f \). Let \( \chi \in C^\infty_0 \) be such that \( \chi = 1 \) on \([-1, 1]\) and \( \text{supp} \chi \subset [-2, 2] \). Let \( f = f_1 + f_2 \) where \( \hat{f}_1 = \hat{\chi} \hat{f} \). Clearly,

(2.22)

\[
\|f - f(\cdot + h)\|_q \leq \|f_1 - f_1(\cdot + h)\|_q + \|f_2 - f_2(\cdot + h)\|_q.
\]

Let

\[
\varphi_h(\xi) = (1 - \chi(h \xi))(1 + |\xi|^2)^{-\frac{\gamma}{2}}.
\]

Observe that

\[
\|\varphi_h\|_\infty + \|\xi \varphi_h(\xi)\|_\infty \leq C |h|^{\gamma}
\]

where the constant \( C \) does not depend on \( |h| \). Hence the Marcinkiewicz multiplier theorem applies to this multiplier \( \varphi_h \). For the second term on the right-hand side of (2.22) one therefore has

(2.23)

\[
\|f_2 - f_2(\cdot + h)\|_q \leq 2\|f_2\|_q = 2 \left\| \left[ \varphi_h(1 + 4\pi^2|\xi|^2)^{-\frac{\gamma}{2}} \hat{f} \right] \right\|_q
\]

The term involving \( f_1 \) in (2.22) is estimated as follows:

(2.24)

\[
\|f_1 - f_1(\cdot + h)\|_q \leq \left\| \int_0^h f_1'(x + t) \, dt \right\|_{L_{q, \gamma}(\mathbb{R})} \leq |h| \|f_1'\|_q
\]

To obtain (2.24), observe that

(2.25)

\[
\hat{f}_1(\xi) = 2\pi i \xi \hat{\chi}(h \xi)|\xi|^{-\gamma} \xi^{-\gamma} \hat{\varphi}_h(\xi).
\]

Let \( m(\xi) = \xi |\xi|^{-\gamma} \hat{\chi}(h \xi) \). Since \( \gamma \leq 1 \), \( m \) is a Marcinkiewicz multiplier and \( \left\| [m(h \cdot) \varphi]^{\gamma} \right\|_q \leq C_q \|g\|_q \) for all \( g \in L^q(\mathbb{R}) \). Since

\[
m(h \xi) = h |h|^{\gamma} \xi |\xi|^{-\gamma} \hat{\chi}(h \xi),
\]

(2.24) follows from (2.25). Combining (2.24), (2.23), and (2.22) proves the lemma. \( \Box \)
3. Uniformity of the Sobolev estimates in the probability for biased Bernoulli convolutions

In Section 2 we recalled the main result from [13]: If $\Pi_\lambda : X \to \mathbb{R}$ is a family of projections that satisfies the transversality condition on some interval $J$, then for any given $\mu$, for a.e. $\lambda \in J$ (possibly depending on $\mu$) the projected measures $\nu_\lambda = \Pi_\lambda(\mu)$ have nice regularity properties; the degree of regularity of $\nu_\lambda$ captures a.e. the dimension of $\mu$ -- a precise formulation is given by the notion of Sobolev dimension introduced in [13]. In all interesting examples it is known that there exist exceptional parameters $\lambda$ where the degree of smoothness of the projection drops. In [13] as well as in Theorem 2.5, this set of “bad” $\lambda$ values, however, might depend on $\mu$. In the case of Erdős projections (4.1) (unlike the Euclidean projections) it is reasonable to believe that the class of exceptional $\lambda$, for which the generic behavior fails, should be to some extent universal, i.e., independent of the measure $\mu$ under consideration. For example, one might conjecture that special algebraic numbers are the only possible exceptions for Bernoulli convolutions. In this section we provide some indication for the universality of the set of bad $\lambda$-values. In fact, for $0 < p < 1$ let

$$\mu_p = \prod_{j=0}^{\infty} (p \lambda_j + (1-p) \delta_1)$$

be the biased (or asymmetric) Bernoulli measure and denote by $\nu_\lambda^p$ the distribution of $\sum_{j=0}^{\infty} \omega_j \lambda_j^p$ under this measure.

**Proposition 3.1.** Let $0 < \lambda_1 < \lambda_2 < 0.649$, $\gamma > 0$, and the interval $I \subset (0,1)$ be such that

$$\sup_{p \in I} \mathcal{E}_{\alpha,\gamma}(\mu_p) < \infty$$

for $\alpha > 1 + \frac{\gamma}{q-1}$. Then

$$\int_{\lambda_1}^{\lambda_2} \sup_{p \in I} \left\| \frac{d\mu_p}{dx} \right\|^{q \gamma} d\lambda < \infty. \quad (3.2)$$

This is a corollary of the following more general statement. Let $\Omega = \Sigma^\mathbb{Z}_+$ be the one-sided shift space on $m$ symbols $\Sigma$ and $\Pi_\lambda$ be as in (2.5).

**Definition 3.2.** A one-parameter family $\{\mu_t\}_{t \in I}$ of measures on $\Omega$ is called analytically extendable if there exists a complex neighborhood $U \supset I$ such that for every cylinder $W \in \mathcal{W}_k$ the functions $\mu_t(W)$ can be extended to an analytic function $\mu_z(W)$ on $U$ such that

$$\forall V \in \mathcal{W}_k, k > \ell, z \in U : \sum_{W \subseteq V \cap \mathcal{W}_k} |\mu_z(W)| \leq C (1 + C|z|)^{k-\ell} |\mu_z(V)|. \quad (3.3)$$

For example, the biased Bernoulli measures $\mu_p$ are such a family. Observe that while in general $\mu_z$ are not measures for $z \not\in \mathbb{R}$, they do give rise to measures $\mu_z^k$ on each of the spaces $\Sigma^k$, since

$$\sum_{W \subseteq V \cap \mathcal{W}_k} \mu_z(W) = \mu_z(V) \quad \text{for all} \quad V \in \mathcal{W}_k, k > \ell, z \in U.$$
Define
\[ E_{q,\alpha}(\mu_z) := \sum_{k \geq 0} \lambda_1^{-(q-1)\alpha k} \sum_{W \in W_k} |\mu_z(W)|^q \]
for \( z \in U \). A final remark is that the exact dependence in (3.3) on \( 3z \) is not important, only that it is uniform. In particular, it can be of the form \( |3z|^\beta \).

**Theorem 3.3.** Suppose \( \{\mu_t\}_{t \in I} \) is an analytically extendable family of measures on \( \Omega \). If \( J = [\lambda_1, \lambda_2] \) is an interval of transversality for \( \Pi_{\lambda} \), then
\[ (3.4) \quad \int \sup_{t \in I} \left\| \frac{d\Pi_{\lambda}(\mu_t)}{dx} \right\|_{q,\gamma}^q d\lambda \leq C \sup_{z \in U} E_{q,\alpha}(\mu_z) \]
where \( \alpha > 1 + \frac{\gamma}{q - 1} \Gamma \), and \( C \) depends on the family \( \{\mu_t\}_{t \in I} \) only through the constant in Definition 3.2.

**Proof.** As above we can assume that \( J \) is an interval of \( \beta \)-transversality with \( \beta > 0 \) very small. In particular, \( \beta < 1 \). First we note that
\[ \left| \phi_j \ast \Pi_{\lambda}(\mu_t)(x) - \phi_j \ast \Pi_{\lambda}(\mu_t^k)(x) \right| \leq \int \left| \phi_j(x - \Pi_{\lambda}(\omega)) - \phi_j(x - \Pi_{\lambda}(\omega)) \right| d\mu_t(\omega) \leq 2\|\phi'\|_\infty C \lambda^k \|\mu_t\| = C2^j \lambda^k \|\mu_t\|. \]
Here \( \Pi_{\lambda}(\omega) := \sum_{j=1}^{k-1} d_{\omega_j}(\lambda)^j \). Therefore, for any positive integer \( k_1 \)
\[ (3.5) \quad \int \sup_{t \in I} \left\| \phi_j \ast \Pi_{\lambda}(\mu_t) \right\|_{q,\gamma}^q d\lambda \leq C \int \sup_{t \in I} \left\| \phi_j \ast \Pi_{\lambda}(\mu_t^k) \right\|_{q,\gamma}^q d\lambda + C2^j \lambda^k \|\mu_t\|^q. \]
Take \( k \) so that \( 2^j \lambda^k \gg (1+\alpha_0) k \), see Lemma 2.7, and \( k_1 = a_1 k \) where \( a_1 \) is some fixed large constant that does not depend on \( \varepsilon \) and \( \beta \). The first term on the right-hand side can be estimated as follows. For every \( x \in \mathbb{R} \) the function \( z \mapsto [\phi_j \ast \Pi_{\lambda}(\mu_t^k)](x) \) is analytic on \( U \). Hence, \( z \mapsto \left[ \phi_j \ast \Pi_{\lambda}(\mu_t^k) \right](x) \) is subharmonic, and thus so is
\[ M_j(z, \lambda) := \left\| \phi_j \ast \Pi_{\lambda}(\mu_t^k) \right\|_{q,\gamma}^q. \]
The dependence of \( M_j \) on \( k_1 \) is implicit. Fix a loop \( \Gamma \subset U \) around the interval \( I \) at distance \( \varepsilon \), where \( I \) is as in the statement of the proposition. Let \( K_{\Gamma}(p, z) \) denote the Poisson kernel of the domain enclosed by \( \Gamma \). By subharmonicity of \( M_j \),
\[ M_j(p, \lambda) \leq \int_{\Gamma} K_{\Gamma}(p, z) M_j(z, \lambda) d\sigma(z) \]
for any \( p \in I \). Recall the well-known property
\[ 0 \leq K_{\Gamma}(p, z) \leq \frac{C}{\text{dist}(p, \Gamma)}. \]
Therefore, by Lemma 2.7
\[ \int \sup_{t \in I} M_j(t, \lambda) d\lambda \leq \frac{C}{\text{dist}(I, \Gamma)} \int_{\Gamma} \int \sup_{t \in I} M_j(z, \lambda) d\lambda d\sigma(z) \]
\[ \leq \frac{C}{\varepsilon} \int_{\Gamma} 2^{(q-1)j} \sum_{W \in W_k} |\mu_t^k(W)|^q d\sigma(z) \]
(3.6)
Now by condition (3.3),
\begin{equation}
\|\mu_z^{k}k\|_q(W) \leq C(1 + C|z|)^{k-1} |\mu_z(W)| \leq C(1 + C\varepsilon)^{(q-1)k} |\mu_z(W)| \quad \text{for all } z \in \Gamma.
\end{equation}
By our choice of \( k \) one has \( 2^{\frac{1}{j}q_j(q-1)} = \lambda_1^{-(q-1)0(q-1)q_j}. \) Summing (3.5) and (3.6) with weights \( 2^{\frac{1}{j}q_j} \) over \( j \) therefore yields, with a suitable choice of \( a_1, \)
\begin{align*}
\int_{\mathbb{T}} & \left( \sum_{j \geq 0} 2^{\frac{1}{j}q_j} \left\| \phi_j * \Pi_\alpha(\mu) \right\|_q \right) d\lambda \\
& \leq \frac{C}{\varepsilon} \int_{\mathbb{T}} \sum_{k \geq 0} \lambda_1^{-(q-1)0(q-1)q_j} (1 + C\varepsilon)^{q_1k} \sum_{W \in W_k} |\mu_z(W)|^q ds(z) \\
& \quad + C \sum_{k \geq 0} \lambda_1^{-(q-1)(q-1)q_j} \l_2^{\alpha_1k} \sup_{t \in T} \|\mu_t\|^q \\
& \leq \frac{C}{\varepsilon} \int_{\mathbb{T}} \sum_{k \geq 0} \lambda_1^{-(q-1)(q-1)q_j} \sum_{W \in W_k} |\mu_z(W)|^q ds(z) + C \sup_{t \in T} \|\mu_t\|^q \\
& \leq \frac{C}{\varepsilon} \sup_{z \in \Gamma} \mathcal{E}_{\alpha_1}(\mu_z) \leq \frac{C}{\varepsilon} \sup_{z \in U} \mathcal{E}_{\alpha_1}(\mu_z).
\end{align*}
In (3.8) we used (3.7) and to pass to line (3.9) we assumed that \( \varepsilon \) and \( \beta \) are small enough so that
\begin{equation}
\lambda_1^{-(q-1)(q-1)q_j} (1 + C\varepsilon)^{q_1} \leq \lambda_1^{(q-1)0(q-1)0}.
\end{equation}
Since \( \alpha > 1 + \frac{\varepsilon}{\varepsilon_1} \), this is always possible. \( \square \)

4. Regularity for biased Bernoulli convolutions

In this section we take \( \Omega = \{ \pm 1 \}^\mathbb{Z}_+ \) equipped with the product measure
\begin{equation}
\mu_p = \prod_{i=0}^\infty (p \delta_1 + (1 - p) \delta_{-1})
\end{equation}
and the Erdős projection
\begin{equation}
\Pi_\lambda : \omega \mapsto \sum_{i=0}^\infty \omega_i \lambda^i, \quad \tau_A^p = \mu_p \circ \Pi_\lambda^{-1}.
\end{equation}
As an auxiliary tool, we will apply Theorem 2.5 to the symbol class
\( \Omega_b = \{-b, -b+2, \ldots, b-2, b\}^\mathbb{Z}_+ \)
equipped with the Erdős projection and certain Bernoulli measures. We cite the following results from [16], see Corollary 5.2 there. See also Solomyak [22].

**Lemma 4.1.** The transversality condition holds for the pair \((\Omega_b, \Pi_\lambda)\) on \((0, y(b)], \) where
\begin{equation*}
y(1) > 0.649, \quad y(2) = 0.5, \quad y(3) > 0.415
\end{equation*}
Observe that the estimate for \( y(1) \) given above states that \([0, 0.649] \) is an interval of transversality for the series \( \sum_{i=0}^\infty \pm \lambda^i. \) In particular, Theorem 2.5 applies for \( \lambda \in (0, 0.649]. \) In order to establish the existence of \( L^2 \) -densities beyond this point, Peres and Solomyak [16] used the values of \( y(2) \) and \( y(3) \), see also the proof of Proposition 4.2 below.
Fix some $0 < p < 1$. For any integers $n, \ell \geq 0$ and $\vec{\theta} = (\theta_0, \ldots, \theta_\ell) \in \{\pm 1\}^{\ell+1}$ let 
\[ \mu_{n, \ell, \vec{\theta}}^p \] 
where $\vec{\omega}^{n+\ell} := (\omega_n, \omega_{n+1}, \ldots, \omega_{n+\ell})$. In other words, $\mu_{n, \ell, \vec{\theta}}^p$ is the Bernoulli measure $\mu_p$ conditioned on $\omega^{n+\ell} = \vec{\theta}$. The main result of this section is the following proposition.

**Proposition 4.2.** There exists a set $\mathcal{N} \subset (0, 1)$ of measure zero with the following property: For every $p \in \left[ \frac{1}{4}, \frac{2}{3} \right]$ and $\lambda \in (e^{-h(\mu_p)}, 1) \setminus \mathcal{N}$ the measures $\Pi_\lambda(\mu_{n, \ell, \vec{\theta}}^p)$ are absolutely continuous with respect to Lebesgue measure for every $n, \ell \geq 0$ and $\theta \in \{\pm 1\}^{\ell+1}$. Furthermore, the densities $f_{n, \ell, \vec{\theta}}^{(\lambda, p)}$ of these measures satisfy a Hölder condition in the $L^1$-norm uniformly in $n, \ell, \vec{\theta}$, i.e.,

\[ \left\| f_{n, \ell, \vec{\theta}}^{(\lambda, p)} - f_{n, \ell, \vec{\theta}}^{(\lambda, p)}(\cdot + h) \right\|_{L^1} \leq C_\lambda n 2^{\ell} h^\sigma \quad \text{for all} \quad 0 \leq h \leq 1 \]

where $\sigma$ is a positive constant which is explicitly computable in terms of $\lambda$ and $p$.

**Proof.** We first fix some $p \in \left[ \frac{1}{4}, \frac{2}{3} \right]$ and show that there is a set $\mathcal{N}_p$ of measure zero with the stated properties. Let $\Omega, \Pi$ be as above. By Lemma 2.9 it is clearly sufficient to show that for almost every $\lambda \in (e^{-h(\mu_p)}, 1)$ there is some $q > 1$ and $\gamma > 0$ such that

\[ \| f_{n, \ell, \vec{\theta}}^{(\lambda, p)} \|_{L^q} \leq C_\lambda 2^{\ell} n \]

for every $n, \ell \geq 0$ and $\theta \in \{\pm 1\}^{\ell+1}$, see (2.12). This gives us (4.2) with $\sigma = \gamma$. Note that since

\[ f_{n, \ell, \vec{\theta}}^{(\lambda, p)}(x + \sum_{i=0}^{\ell} (\theta_i - \tau_i) \lambda^{n+i}) = f_{n, \ell, \vec{\theta}}^{(\lambda, p)}(x) \]

the norms on the left-hand side of (4.3) do not depend on the choice of $\vec{\theta}$.

Let $\lambda_1 \in (e^{-h(\mu_p)}, 0.649)$ and choose $\gamma > 0$ and $1 < q < 2$ satisfying

\[ -D_q(\mu_p) < \log \lambda_1 \left( 1 + \frac{q^\gamma}{q - 1} \right). \]

This can always be achieved since $D_q(\mu_p) \to h(\mu_p)$ as $q \to 1+$. Let $J = [\lambda_1, 0.649]$. Theorem 2.5 therefore implies that there is a constant $C$ that does not depend on $n, \ell, \vec{\theta}$ such that

\[ \int_J \| f_{n, \ell, \vec{\theta}}^{(\lambda, p)} \|_{L^q}^q \ d\lambda \leq C E_{q, \alpha}(\mu_{n, \ell, \vec{\theta}}^p) \leq C \sum_{k=0}^{\infty} \lambda_1^{-\alpha(q-1)k} \sum_{W \in \mathcal{W}_k} \mu_{n, \ell, \vec{\theta}}^p(W)^q \]

\[ \leq C \sum_{k=0}^{\infty} \lambda_1^{-\alpha(q-1)k} \left[ \frac{p^q}{1 - p^q} \right]^{k-\ell} \leq C \left[ \frac{p^q}{1 - p^q} \right]^{-\ell}. \]

Let $q_1 := -\log(p^n + (1 - p)^q)$. Observe that $c_1 = \log 2$ has this property. This gives rise to the choice of $2^\ell$ in (4.2). Consequently,

\[ \int_J \sum_{n, \ell \geq 1} n^{-q} e^{-q \gamma \ell} \| f_{n, \ell, \vec{\theta}}^{(\lambda, p)} \|_{L^q}^q \ d\lambda \leq C \]

and thus for a.e. $\lambda \in J$

\[ \| f_{n, \ell, \vec{\theta}}^{(\lambda, p)} \|_{L^q} \leq C \lambda n 2^{\ell} \]

for all $n, \ell \geq 1$ and all $\theta \in \{\pm 1\}^{\ell+1}$. 


To pass beyond the point 0.649 we use the method from Peres and Solomyak [16]. First, note that the main conclusion (4.7) of the preceding paragraph remains valid with $q = 2$ provided

$$
\lambda > e^{-D_2(x_0)} = p^2 + (1-p)^2.
$$

In the range $p \in \left[ \frac{1}{3}, \frac{2}{3} \right]$ one has $p^2 + (1-p)^2 \leq \frac{2}{3} < 0.6$. Splitting the sum $\sum_{n=0}^{\infty} \pm \lambda^n$ into odd and even indices one obtains the relation

$$
(4.8) \quad \frac{\langle \lambda \rangle_{n, \ell_{-2}}}{\langle \lambda_2 \rangle_{n, \ell_1}} = \frac{\langle \lambda \rangle_{n_1, \ell_{1,2}}}{\langle \lambda \rangle_{n_2, \ell_{2,2}}},
$$

where $n_{i, \ell_i}$ are defined by

$$
\{n, n+1, \ldots, n+\ell\} = \{2n_1, 2n_1+2, \ldots, 2n_1+2\ell\} \cup \{2n_2+1, 2n_2+3, \ldots, 2n_2+2\ell+1\}
$$

and $\ell_1$, $\ell_2$ are the subsequences of $\ell$ with even and odd indices, respectively. Therefore,

$$
\left\| \frac{\langle \lambda \rangle_{n, \ell_{-2}}}{\langle \lambda \rangle_{n, \ell_1}} \right\|_{2, \gamma} \leq \left\| \frac{\langle \lambda \rangle_{n, \ell_{1,2}}}{\langle \lambda \rangle_{n_2, \ell_{2,2}}} \right\|_{2, \gamma}
$$

for any choice of $\gamma$. In particular, if $\lambda$ is such that $\langle \lambda \rangle_{n, \ell_{1,2}}$ satisfies (4.7) with $q = 2$ for all choices of $n_1, \ell_1$ and $\ell_1$, and some fixed $\gamma$, then $\langle \lambda \rangle_{n, \ell_{-2}}$ will have the same property. This implies that it suffices to show that almost every $\lambda \in \left[ p^2 + (1-p)^2, \sqrt{p^2 + (1-p)^2} \right]$ satisfies (4.7) with $q = 2$. Proceeding inductively, repeated squaring then allows one to cover the entire interval $(p^2 + (1-p)^2, 1)$. Since $p \in \left[ \frac{1}{3}, \frac{2}{3} \right]$, it furthermore suffices to cover $(0.649, \sqrt{5/9})$. The convolution $\langle \lambda \rangle_{n, \ell_{-2}} \ast \langle \lambda \rangle_{n, \ell_{1,2}}$ is the distribution of the sum $\sum_{n=0}^{\infty} a_n \lambda^n$ with digit-set $\{-2,0,2\}$ and probabilities $p^2, 2p(1-p), (1-p)^2$, and the digits with indices between $n, n+\ell$ fixed. In view of Lemma 4.1 with $b = 2$, $(0, \frac{4}{3})$ is a transversality interval in this case. Assume therefore that

$$
(4.9) \quad p^4 + 4p^2(1-p)^2 + (1-p)^4 < \lambda_1 \lambda_2 = \frac{1}{2},
$$

and choose $\gamma > 0$ so that $\lambda_{1,2,2,\gamma} > p^4 + 4p^2(1-p)^2 + (1-p)^4$. Then, by (4.8)

$$
\int_{\lambda_1}^{\lambda_2} \left\| \frac{\langle \lambda \rangle_{n, \ell_{1,2}}}{\langle \lambda \rangle_{n_2, \ell_{2,2}}} \right\|_{2, \gamma}^2 d\lambda \leq \int_{\lambda_1}^{\lambda_2} \left( \int_{\mathbb{R}} \left\| \frac{\langle \lambda \rangle_{n_1, \ell_{1,2}}}{\langle \lambda \rangle_{n_2, \ell_{2,2}}} \lambda \right\|_{2, \gamma}^2 d\lambda \right)^{\frac{1}{2}} \left( \int_{\mathbb{R}} \left\| \frac{\langle \lambda \rangle_{n_1, \ell_{1,2}}}{\langle \lambda \rangle_{n_2, \ell_{2,2}}} \right\|_{2, \gamma}^2 d\lambda \right)^{\frac{1}{2}} d\lambda
$$

and

$$
(4.10) \quad \leq C \int_{\lambda_1}^{\lambda_2} \left( \left( \int_{\mathbb{R}} \left\| \frac{\langle \lambda \rangle_{n_1, \ell_{1,2}}}{\langle \lambda \rangle_{n_2, \ell_{2,2}}} \right\|_{2, \gamma}^2 d\lambda \right)^{\frac{1}{2}} \left( \int_{\mathbb{R}} \left\| \frac{\langle \lambda \rangle_{n_1, \ell_{1,2}}}{\langle \lambda \rangle_{n_2, \ell_{2,2}}} \right\|_{2, \gamma}^2 d\lambda \right)^{\frac{1}{2}} \right)^{\frac{1}{2}} d\lambda.
$$

The two factors in (4.10) are now estimated by means of Theorem 2.4. A straightforward calculation analogous to (4.5) yields that

$$
\int_{\lambda_1}^{\lambda_2} \left\| \frac{\langle \lambda \rangle_{n_1, \ell_{1,2}}}{\langle \lambda \rangle_{n_1, \ell_{1,2}}} \right\|_{2, \gamma}^2 d\lambda \leq C[p^4 + 4p^2(1-p)^2 + (1-p)^4]^{-\ell_1}
$$

and similarly for the second factor in (4.10). Let $c_2$ satisfy

$$
4c_2 > \max_{p \in \left[ \frac{1}{3}, \frac{2}{3} \right]} -\log[p^4 + 4p^2(1-p)^2 + (1-p)^4].
$$
Observe that \( c_2 = \log 2 \) is an admissible choice. As in (4.6) it follows that
\[
\int_{\sqrt{\lambda_1}}^{\sqrt{\lambda_2}} \sum_{n \geq 1} n^{-2 \gamma} e^{-2 \gamma n^2} \| f_{n, \ell}^{(\lambda, p)} \|_{2, \gamma}^2 \, d\lambda < \infty.
\]
Hence, for a.e. \( \lambda \in [\sqrt{\lambda_1}, \sqrt{\lambda_2}] \) one obtains
\[
\| f_{n, \ell}^{(\lambda, p)} \|_{2, \gamma} \leq C \lambda n^{2 \gamma} \leq C \lambda n^{2 \ell}
\]
for all \( n, \ell \geq 1 \) and all \( \theta \in \{ \pm 1 \}^{\ell+1} \). Maximizing over \( p \in [\frac{1}{3}, \frac{2}{3}] \) one obtains that the interval in (4.9) contains \( (11/27, 1/2) \). Since \( \sqrt{11/27} < 0.649 \), we have covered \((0, 1/\sqrt{2})\). In view of (4.10) it thus remains to cover the interval \((1/\sqrt{2}, \sqrt{5/9})\). To this end consider the triple convolution \( \nu_{\lambda}^{(p)} \ast \nu_{\lambda}^{(q)} \ast \nu_{\lambda}^{(r)} \). This is the distribution of the sum \( \sum_{n=0}^{\infty} a_n \lambda^n \) with digit set \( \{-3, -1, 0, 1, 3\} \) and probabilities \( p^3, 3p^2(1-p), 3p(1-p)^2, (1-p)^3 \). We shall also need the following analogue of (4.8) involving three terms:
\[
\hat{f}_{n, \ell}^{(\lambda, p)}(\xi) = \hat{f}_{n, \ell}^{(\lambda, p)}(\xi) \hat{f}_{n, \ell}^{(\lambda, q)}(\xi) \hat{f}_{n, \ell}^{(\lambda, r)}(\xi)
\]
where \( n, \ell, \xi \) are obtained by splitting the sum \( \sum_{k=0}^{\infty} \pm \lambda^k \) modulo 3. Applying Lemma 4.1 with \( b = 3 \) and Theorem 2.4 yields
\[
(4.12)
\]
for small \( \gamma > 0 \) and with \( \lambda_2 = (245/729)^{\frac{1}{3}}, \lambda_3 = 0.427^3 \). Here \( c_3 \) can be taken to be \( \leq \log 2 \). Since \( (245/729)^{\frac{1}{3}} < 1/\sqrt{2} \) and \( 0.427^3 > \sqrt{5/9} \), we are done.

We now show how to obtain uniformity in \( p \), i.e., having one null set \( \mathcal{N} \) so that (4.2) holds for all \( p \in [\frac{1}{3}, \frac{2}{3}] \) and all \( \lambda > e^{-h(p)} \) with \( \lambda \notin \mathcal{N} \). We begin by establishing this under the additional assumption that \( \lambda \) is below the transversality limit 0.649. Partition the parameter set
\[
\left\{ (p, \lambda) \mid \frac{1}{3} < p < \frac{2}{3}, \ e^{-h(p)} < \lambda < 0.649 \right\}
\]
into a countable number of closed rectangles \( R_{ij} = [p_i, p_j] \times [\lambda_i, \lambda_j] \) on each of which we can apply Theorem 3.3 with \( U \) a small complex neighborhood of \([p_i, p_j]\). In order to do so, we take \( q_i > 1 \), \( \gamma_i > 0 \), so that
\[
-D_{q_i}(\mu_p) < \log \lambda_i \left( 1 + \frac{q_i \gamma_i}{q_i - 1} \right)
\]
for all \( p \in [p_i, p_j] \). A calculation similar to (4.5) shows that
\[
\int_{\Lambda_i} \sum_{n \geq 1} n^{-q_i} e^{-n q_i \gamma_i} \sup_{p \in [p_i, p_j]} \| f_{n, \ell}^{(\lambda, p)} \|_{2, \gamma} \, d\lambda < \infty
\]
where we may take \( c_i = \log 2 \) (by taking \( U \) to be small enough one may use any \( c_i \) satisfying \( q_i c_i > -\min_{p \in [p_i, p_j]} \log [p^{p_i} (1-p)^{p_j}] \)).

Next consider the region
\[
(p, \lambda) \in \left[ \frac{11}{27}, \frac{2}{3} \right] \times \left[ \sqrt{\frac{11}{27}} + \varepsilon_0, \frac{1}{\sqrt{2}} \right].
\]
Let $U$ be a complex neighborhood of the interval $[\frac{1}{3}, \frac{2}{3}]$ depending on $\varepsilon_0$ (which is some fixed small number, for example we can take $\varepsilon_0 = \frac{1}{201}$). Apply Theorem 3.3 with $q = 2$ and this region $U$ to the integrals in equation (4.10). More precisely, introducing suprema into the calculation leading up to (4.10) shows that

$$
\int_{\Lambda_1} \sup_{p \in [\frac{1}{3}, \frac{2}{3}]} \left\| f_{\lambda,p}^{(\lambda)} \right\|_{L^2}^2 \lambda \leq C \left( \int_{\Lambda_1} \sup_{p \in [\frac{1}{3}, \frac{2}{3}]} \left\| f_{n_1^1,\ell}^{(\lambda)} \right\|_{L^2}^2 \lambda \right)^{\frac{1}{2}} \left( \int_{\Lambda_1} \sup_{p \in [\frac{1}{3}, \frac{2}{3}]} \left\| f_{n_2^2,\ell}^{(\lambda)} \right\|_{L^2}^2 \lambda \right)^{\frac{1}{2}}.
$$

Both integrals on the right-hand side of the above equation are estimated by means of Theorem 3.3. This allows one to conclude that

$$
\int_{\Lambda_1} \sum_{n \geq 1} n^{-2} \left\| f_{n_\ell}^{(\lambda)} \right\|_{L^2}^2 \lambda \leq \infty.
$$

Proceeding similarly for the third region (with some choice of small $\varepsilon_1 > 0$)

$$(p, \lambda) \in [\frac{1}{3}, \frac{2}{3}] \times \left[ \left( \frac{245}{25} \right)^{\frac{3}{2}} + \varepsilon_1, 0.427 \frac{3}{2} \right]
$$

by means of the triple convolution method described above finishes the proof. \hfill \Box

5. CONDITIONING $\sum_{j=0}^{\infty} \omega_j \lambda^j$ ON BLOCKS OF DIGITS $\omega_n^{\ell+\ell}$

In Sections 2 and 4 it was shown that for a.e. $\lambda \in (\frac{1}{2}, 1)$ and every $p \in [\frac{1}{3}, \frac{2}{3}]$ such that $\lambda \in (e^{-h(\mu_p)}, 1)$ the conditional measures

$$
A \mapsto \mu_p(\Pi_\lambda(\omega)) \in A \mid \omega_n^{\ell+\ell} = \emptyset,
$$

are absolutely continuous with respect to Lebesgue measure. Moreover, this fails below $e^{-h(\mu_p)}$, see [16]. Furthermore, if $f_{n_\ell}^{(\lambda)}$ denotes the densities of these measures, then one has the $L^1$-Hölder bound

$$
\| f_{n_\ell}^{(\lambda)}(x) - f_{n_\ell}^{(\lambda)}(x + h) \|_1 \leq C \lambda 2^{\ell} n \| h \|_p
$$

for all $h$ with positive constants $C, c, \sigma$ depending on $\lambda$, see Proposition 4.2. Observe that for simplicity we no longer indicate the dependence of $f_{n_\ell}^{(\lambda)}$ on $p$. The purpose of this section is to apply (5.1) to show that for a.e. $\lambda \in (e^{-h(\mu_p)}, 1)$

$$
\lim_{n \to \infty} \frac{f_{n_\ell}^{(\lambda)}}{f_\lambda}(x) = 1 \quad \text{for a.e. } x \in \text{supp}(\nu_\lambda),
$$

where $f_\lambda = \frac{d\nu_\lambda}{dx}$. This means that conditioning on blocks of digits $\omega_n^{\ell+\ell}$ does not change the distribution of the sums $\sum_{j=0}^{\infty} \omega_j \lambda^j$ as $n \to \infty$. In fact, the same statement holds with $\ell, n \to \infty$ provided $\ell \leq n \rho$, where $\rho > 0$ is sufficiently small. Note that the denominator in (5.2) is a
weighted average of the numerators over \( \mathcal{Q} \). We now state, and prove, an equivalent form of Theorem 1.2. For this purpose we would like to point out that
\[
\frac{\mu_p[\omega_n^{n+\ell}] = \mathcal{Q} | \Pi_\lambda(\omega) = x]}{\mu_p[\omega_n^{n+\ell}] = \mathcal{Q}} = \frac{f_{n,\ell,\mathcal{Q}}^{(\lambda)}}{f_\lambda}(x).
\]

**Theorem 5.1.** For a.e. \( \lambda \in (1/2, 1) \) and all \( p \in \left[ \frac{1}{3}, \frac{2}{3} \right] \) such that \( H(p) > \log \lambda \), there exists \( \rho > 0 \) such that the following holds:

\[
\lim_{n \to \infty} \sup_{1 \leq \ell \leq np} \sup_{\theta \in \{\pm 1\}^{\ell+1}} \left| \frac{f_{n,\ell,\theta}^{(\lambda)}}{f_\lambda}(x) - 1 \right| = 0 \quad \text{for a.e. } x \in \text{supp}(\nu^p_\lambda)
\]

In particular, for fixed \( \ell \geq 1 \) and \( \theta \in \{\pm 1\}^{\ell+1} \),

\[
\lim_{n \to \infty} \frac{f_{n,\ell,\theta}^{(\lambda)}}{f_\lambda}(x) = 1 \quad \text{for a.e. } x \in \text{supp}(\nu^p_\lambda).
\]

**Proof.** Let \( \lambda \in (e^{-h(p)}, 1) \) be such that (5.1) holds. In particular, \( f_\lambda = \frac{d \mu_p}{dx} \) and each of the \( f_{n,\ell,\theta}^{(\lambda)} \) exist. Now fix some integers \( n, \ell \) and a vector \( \theta \in \{\pm 1\}^{\ell+1} \). Recall that \( \mu_p, \theta \) denotes the measure \( \mu_p \) conditioned on \( \omega_n^{n+\ell} = \theta \) (for simplicity we again do not indicate the dependence on \( p \)). If \( A \subset \mathbb{R} \) is a Borel set, then clearly

\[
\sum_{\mathcal{Z}_0 \in \{\pm 1\}^{\ell+1}} \mu_p(\mathcal{Z}) \mu_{n,\ell,\theta}^{(\lambda)} \left( \Pi_{\lambda}(\omega) \in A + \sum_{i=0}^{\ell} (\theta_i - \tau_i) \lambda^{i+1} \right) = \mu_p(\Pi_{\lambda}(\omega) \in A).
\]

Passing to densities yields

\[
f_\lambda(x) = \sum_{\mathcal{Z}_0 \in \{\pm 1\}^{\ell+1}} \mu_p(\mathcal{Z}) f_{n,\ell,\theta}^{(\lambda)} \left( x + \sum_{i=0}^{\ell} (\theta_i - \tau_i) \lambda^{i+1} \right).
\]

In view of (5.1) one obtains from (5.5) that

\[
\|f_\lambda - f_{n,\ell,\theta}^{(\lambda)}\|_1 \leq C n 2^\ell \lambda^{n\sigma}.
\]

For any \( \varepsilon > 0 \) let

\[
\mathcal{B}_n(\varepsilon) := \{ \omega \left| \sup_{1 \leq \ell \leq np} \sup_{\theta \in \{\pm 1\}^{\ell+1}} \left| \frac{f_{n,\ell,\theta}^{(\lambda)}}{f_\lambda}(x) - 1 \right| > \varepsilon \}.
\]

Then, in view of (5.6)

\[
\mu_p(\mathcal{B}_n(\varepsilon)) \leq \left| \frac{np}{\ell} \right| \sum_{\mathcal{Z}_0 \in \{\pm 1\}^{\ell}} \left| \varepsilon \int \left| \frac{f_{n,\ell,\theta}^{(\lambda)}}{f_\lambda}(\Pi_{\lambda}(\omega)) - 1 \right| d\mu_p(\omega)
\]

\[
\leq \left| \frac{np}{\ell} \right| \sum_{\mathcal{Z}_0 \in \{\pm 1\}^{\ell+1}} \left| \varepsilon \int \left| f_{n,\ell,\theta}^{(\lambda)} - f_\lambda \right|_1
\]

\[
\leq C n^2 2^{np} 2^{np} \lambda^{n\sigma} \varepsilon^{-1}.
\]
Thus, for small $\rho > 0$,
\[
\sum_{n=1}^{\infty} \mu_p(B_n(\varepsilon)) < \infty \implies \mu_p(\limsup_{n \to \infty} B_n(\varepsilon)) = 0.
\]
Since $\varepsilon$ was arbitrary, the proposition follows. \hfill \square

6. TAIL TRIVIALITY AND ENTROPY OF $\xi_{\lambda}$

Let $X = (X, \mathcal{A}, \mu)$ be a probability space with an automorphism $T : X \to X$.

**Definition 6.1.** $T$ is said to be a $K$-automorphism if there exists a $\sigma$-subalgebra $\xi \subset \mathcal{A}$ (which in Lebesgue spaces is the same as a measurable partition) satisfying the following conditions, see [3] chapter 10:

1. $T\xi \supseteq \xi$
2. $\bigvee_{n=-\infty}^{\infty} T^n\xi = \mathcal{A}$ mod $\mu$
3. $\bigwedge_{n=-\infty}^{\infty} T^n\xi = \emptyset, X \mod \mu$.

Such a partition is called a $K$-partition.

Clearly, a $K$-partition as defined in the introduction satisfies these conditions. In a Lebesgue space, the converse is also easily verified, since any $\sigma$-algebra is countably generated mod $\mu$.

The main purpose of this section is to show that property (ii) from definition of $K$-partitions is satisfied and to calculate the entropies of these partitions. This will follow from the following two — very general — lemmas, which for notational simplicity only we state for the class of Bernoulli measures.

**Lemma 6.2.** If for all nonnegative integers $\ell$, and $\theta \in \{\pm1\}^{\ell+1}$
\[
(6.1) \quad \mu_p[\omega^\ell = \theta | T^{-n}\xi_{\lambda}(\omega)] \to \mu_p[\omega^\ell = \theta] \text{ as } n \to \infty
\]
then
\[
T_{\lambda} := \bigwedge_{n=0}^{\infty} T^{-n}\xi_{\lambda} = \emptyset, X \mod \mu.
\]

**Proof.** Let $E \in T_{\lambda}$. Fix an $\varepsilon > 0$ and let $E_N$ be a cylinder set with unspecified coordinates in $\mathbb{Z} \setminus (-\infty, -N) \cup (N, \infty)$ so that
\[
\mu(E \Delta E_N) < \varepsilon.
\]
Since $E_N$ is a cylinder set, depending on a finite number of coordinates, (6.1) implies that
\[
\mathbb{E}(1_{E_N} | T_{\lambda}) = \lim_{n \to \infty} \mathbb{E}(1_{E_N} | T^{-n}\xi_{\lambda}) = \mu_p(E_N).
\]
Since $E \in T_{\lambda}$,
\[
\mathbb{E}(1_{E} | T_{\lambda}) = 0 \text{ or } 1 \quad \text{a.s.}
\]
The conditional expectation operator is a projection, and in particular does not increase the $L^1$-norm of $1_E - 1_{E_N}$, and this implies that
\[
\mu_p(E_N) \in [0, \varepsilon] \cup [1 - \varepsilon, 1].
\]
Since $\mu(E \Delta E_N) < \varepsilon$ and $\varepsilon$ is arbitrary, we see that $\mu(E) = 0$ or 1. \hfill \square

The following lemma shows that the entropy formula (1.5) holds whenever $f_{\lambda}$ is absolutely continuous.
Lemma 6.3. Let $\mu_p$ be Bernoulli with some fixed $p$ and assume that for some fixed $\lambda \in (1/2, 1)$ the distribution $\nu_\lambda$ is absolutely continuous with density $f_\lambda$. Then the entropy of $\xi_\lambda$ is given by

$$h(\xi_\lambda|T^{-1}\xi_\lambda) = h_{\mu_p}(T) + \log \lambda,$$

with $T$ denoting the right shift.

Proof. Recall that

$$(6.2) \quad h(\xi_\lambda|T^{-1}\xi_\lambda) = -\int_{\Omega} \log \mu_p(\omega_0 = \tau_0 | \Pi_\lambda(\omega) = \Pi_\lambda(\zeta)) d\mu_p(\zeta).$$

The integrand on the right-hand side is defined for every $\lambda$ for which $f_\lambda = \frac{df_\lambda}{d\lambda}$ exists. Indeed (for a.e. $\zeta$),

$$\mu_p(\omega_0 = \tau_0 | \Pi_\lambda(\omega) = \Pi_\lambda(\zeta)) = \frac{\lim_{\varepsilon \to 0^+} \frac{1}{2\varepsilon} \mu_p(\omega_0 = \tau_0, | \sum_{j=1}^{\infty} \omega_j \lambda^j - \sum_{j=0}^{\infty} \tau_j \lambda^j | < \varepsilon)}{\frac{1}{2\varepsilon} \mu_p(\sum_{j=0}^{\infty} \omega_j \lambda^j - \sum_{j=0}^{\infty} \tau_j \lambda^j | < \varepsilon)}$$

$$= \frac{\mu_p(\omega_0 = \tau_0) f_{\lambda}(\Pi_\lambda(T^{-1}\zeta))}{f_{\lambda}(\Pi_\lambda(\zeta))}.$$ (6.3)

Let $\psi_\lambda(\zeta)$ denote the logarithm of (6.3). Observe that $\psi_\lambda$ is finite for a.e. $\zeta$. According to (6.2),

$$h(\xi_\lambda|T^{-1}\xi_\lambda) = -\int \psi_\lambda(\zeta) d\mu_p(\zeta).$$

By the ergodic theorem,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} -\psi_\lambda(T^{-k}\zeta) = h(\xi_\lambda|T^{-1}\xi_\lambda)$$

for a.e. $\zeta$. On the other hand,

$$\frac{1}{n} \sum_{k=0}^{n-1} -\psi_\lambda(T^{-k}\zeta) = h_{\mu_p}(T) + \log \lambda + \frac{1}{n} \left[ \log(f_{\lambda}(\Pi_\lambda(\zeta))) - \log(f_{\lambda}(\Pi_\lambda(T^{-n}\zeta))) \right].$$

Since for a.e. $\zeta$

$$\lim \inf_{n \to \infty} \frac{1}{n} \left| \log(f_{\lambda}(\Pi_\lambda(\zeta))) - \log(f_{\lambda}(\Pi_\lambda(T^{-n}\zeta))) \right| = 0,$$

the lemma follows. \qed

This establishes Theorem 1.1 from Section 1.

7. The case where $\lambda$ is the inverse of a Pisot or Salem number

It was known to Erdős [5] that $\nu_\lambda$ is singular if $\lambda^{-1}$ is a Pisot number. It is not known whether $\nu_\lambda$ is singular for inverses of Salem numbers, but it is not difficult to see that for such $\lambda$ the measure $\nu_\lambda$ is not in $L^{2\gamma}$ for any $\gamma > 0$, see [14] Section 5 for more details and references to earlier work. Recall that a Pisot number is an algebraic integer $\beta$ so that all the conjugates of $\beta$ lie strictly inside the unit circle, whereas a Salem number has all its conjugates inside the unit
disk, with at least one on the unit circle. See [14] for further details. An example of a Pisot number is the golden mean
\[ \beta = \frac{1 + \sqrt{5}}{2}. \]

In this section we show that contrary to the generic behavior discussed in the previous section, if $\lambda^{-1}$ is a Pisot number, then knowing the sum $\sum_{j=0}^\infty \omega_j \lambda^j$ has a drastic effect on the distribution of the block of digits $\omega_n^{n+\ell}$ even as $n \to \infty$.

Consider first the Pisot case. Fix some Pisot number $\beta$ and let $\lambda = \beta^{-1}$. It was shown by Garsia [6] that for any $\omega_0^N \neq z_0^N$

\[ \left| \sum_{j=0}^N \omega_j \lambda^j - \sum_{j=0}^N \tau_j \lambda^j \right| \geq c_0 \lambda^N \tag{7.1} \]

with some constant $c_0$ depending only on $\beta$. For the convenience of the reader we recall the simple argument. First, suppose that

\[ \sum_{k=0}^N a_k \beta^k \neq 0 \]

where $a_k \in \mathbb{Z}$, $|a_k| \leq M$. Since the minimal polynomial of $\beta_1 = \beta$ equals the minimal polynomial of the conjugates $\beta_2, \ldots, \beta_d$ of $\beta$, one obtains

\[ \prod_{j=1}^d \left| \sum_{k=0}^N a_k \beta_j^k \right| \neq 0. \tag{7.3} \]

Expanding the left-hand side of (7.3) and observing that the symmetric polynomials of $\beta_1, \ldots, \beta_d$ are integer-valued, implies that in fact

\[ \prod_{j=1}^d \left| \sum_{k=0}^N a_k \beta_j^k \right| \geq 1. \tag{7.4} \]

Therefore,

\[ \left| \sum_{k=0}^N a_k \beta^k \right| \geq M^{-d+1} \prod_{j=2}^d (1 - |\beta_j|) = c_0(\beta). \tag{7.5} \]

Clearly, (7.5) with $M = 2$ implies (7.1).

Our goal is to show that for large $\ell$, conditioning on the value of $\sum_{j=0}^\infty \omega_j \lambda^j$ places a strong restriction on any block of digits $\omega_n^{n+\ell}$, regardless of $n$. Indeed, suppose

\[ \sum_{j=0}^\infty \omega_j \lambda^j = x \]

and fix some $n$ and $\ell$. Then

\[ \left| \sum_{j=0}^{n-1} \omega_j \lambda^j - x \right| \leq \frac{\lambda^n}{1 - \lambda}, \quad \left| \sum_{j=0}^{n+\ell} \omega_j \lambda^j - x \right| \leq \frac{\lambda^{n+\ell+1}}{1 - \lambda}. \tag{7.6} \]
In view of (7.1) this implies that the sums
\[ \sum_{j=0}^{n-1} \omega_j \lambda^j \quad \text{and} \quad \sum_{j=0}^{n+\ell} \omega_j \lambda^j \]
can attain only
\[ B := \frac{\lambda}{c_0(1-\lambda)} \]
many different values. Thus conditioning on the infinite sum implies that the finite sum
\[ (7.7) \quad \sum_{j=0}^{\ell} \omega_{n+j} \lambda^j \]
can attain only \( B^2 \) different values, regardless of the choice of \( n \) and \( \ell \). For large \( \ell \), this presents a strong restriction on the allowed sequences \( \{\omega_j\}_{n+\ell}^n \), since without conditioning, the sums (7.7) are \( 2\lambda^\ell \) dense in \( [-\frac{1}{1-\lambda}, \frac{1}{1-\lambda}] \). A proof of this easy fact can be found e.g. in Corollary 2.2 of [17]. This immediately implies that \( \xi_\lambda \) is not a \( K \)-partition.

Now let \( \lambda = \beta^{-1} \) where \( \beta \) is a Salem number. Under the assumption
\[ \sum_{0}^{N} a_k \beta^k \neq 0 \]
with \( a_k \in \mathbb{Z} \) satisfying \( |a_k| \leq M \) one concludes from (7.4) that
\[ (7.9) \quad \left| \sum_{k=0}^{N} a_k \beta^k \right| \geq \prod_{j=2}^{d} \left[ \sum_{k=0}^{N} |a_k| \beta_j^k \right]^{-1} \geq M^{-d+1}(N+1)^{-d+1}. \]

Setting \( \sum_{j=0}^{\infty} \omega_j \lambda^j = x \) the calculation following (7.6) in conjunction with (7.9) shows that there are at most
\[ C(n+\ell)^{2d} \]
many possibilities for the sum
\[ \sum_{j=0}^{\ell} \omega_{n+j} \lambda^j. \]

As above, unconstrained, this sum attains at least \( c\lambda^{-\ell} \) many values since it is \( \lambda^\ell \) dense in an interval of size one. This shows that
\[ \mu_p[\omega_{n+\ell}^n] = \theta \left| \Pi_\lambda(\omega) \right| = 0 \]
for many choices of \( \theta \) as soon as \( \ell > C \log n \). In particular, (1.7) fails for inverses of Salem numbers.
8. General Bernoulli Measures

In this section we show that any abstract Bernoulli system has K-partitions of any entropy between 0 and that of the system.

Theorem 8.1. Let \((X,T,\mu)\) be a Bernoulli system. Then \(X\) has a measurable partition of entropy \(h\) for any \(h \in (0, h(X,T,\mu))\).

Proof. Let \(I = [H(1/3), \log 2]\). There exists \(B_I\) such that any \(B > B_I\) can be written as a finite sum of at least two elements of \(I\). First assume \(h(X,T,\mu) > B_I\). Let \(X = \{\pm1\}^\mathbb{Z}\), and for \(n \geq 1\), denote by \(\sigma^n : \Omega^n \rightarrow \Omega^n\) the shift map

\[
(\omega_1, \ldots, \omega_n) \mapsto (\sigma \omega_1, \ldots, \sigma \omega_n).
\]

By the Ornstein isomorphism theorem (see [20]), the system \((X,T,\mu)\) is isomorphic to

\[
(\Omega^n, \sigma^n, \mu_{p_1} \times \cdots \times \mu_{p_n})
\]

as long as \(\sum_{i=1}^{n} H(p_i) = h(X,T,\mu)\). If \(\xi_i\) is a K-partition for \((\Omega, \sigma, \mu)\) with entropy \(h_i\), then \(\bigvee_{i=1}^{n} \xi_i\) is a K-partition for \(\Omega^n \cong X\) of entropy \(\sum_{i=1}^{n} h_i\). From Theorem 1.1, we know in particular how to construct a K-partition for \((\Omega, \sigma, \mu)\) of a.e. entropy in \((0, H(p_i))\) as long as \(p_i \in (\frac{1}{3}, \frac{2}{3})\).

Thus, if \(X\) can be presented as the product of at least two \((\Omega, \sigma, \mu)\) with \(p_i \in (\frac{1}{3}, \frac{3}{2})\) then the set of entropies attainable by a K-partition in \(X\) contains

\[
H_1 + H_2 + \cdots + H_n
\]

with \(H_i \in (0, H(p_i))\) and \(\mathcal{L}((0,H(p_i)) \setminus H_i) = 0\). Thus, \(H_1 + H_2 + \cdots + H_n = (0, h(X,T,\mu))\).

Furthermore, the usual construction of K-partitions from finite separating partitions easily shows one has K-partitions of entropy \(h(X,T,\mu)\) in \(X\), so the theorem is proved for all \(X\) with \(h(X,T,\mu) > B_I\).

We now pass to the general case. Consider the mixing Markov chain \(\{\tau_n\}\) (with memory \(N + 1\)) on the symbols \(\{-1,0,1\}\), characterized as follows: after each nonzero symbol, there is a run of \(N\) or \(N + 1\) zeros (with probabilities \(1 - \alpha\) and \(\alpha\), respectively). This run is then followed by a 1 with probability \(p \in [1/3, 2/3]\) and a -1 with probability \(1 - p\).

Define \(\omega(\tau)\) to be the bi-infinite word with symbols in \(\{\pm1\}\), obtained from \(\tau\) by deleting all zeros, where \(\omega(\tau)\) is the first nonzero element in \(\tau_0^{\infty}\).

We construct a realization of the Markov chain \(\{\tau_n\}\) by means of the process \(\{V_n(\tau)\}\),

\[
V_n(\tau) := \gamma_0^{(\lambda)}[\omega(T^{-n}\tau)],
\]

where \(T\) is the right shift. Note that almost surely

\[
V_n(\tau) = V_{n+1}(\tau)
\]

only for \(n\) such that \(\tau_n = 0\). Hence the \(\{V_n(\tau)\}_{n \in \mathbb{Z}}\) generate modulo null sets the \(\sigma\)-algebra of measurable sets.
We claim that for almost every $\lambda \in (e^{-H(p)}, 1)$, specifically, those $\lambda$ for which Theorem 1.1 holds, the left tail of the $\{V_n\}$ is trivial. Indeed if $n < 0$ and $k = \lfloor n/(N + 1) \rfloor$, then

\[ \sigma(V_j; j \leq n) \subset \sigma(\tau_j; j < 0) \lor \sigma\left(Y_j^{(\lambda)}[\omega(\cdot)]; j \leq k\right) \tag{8.2} \]

and so since the two $\sigma$-algebras on the right hand side of (8.2) are independent, and using the left tail-triviality of $\{Y_j^{(\lambda)}[\cdot]\}$,

\[ \bigwedge_{n=-\infty}^0 \sigma(V_j; j \leq n) \subset \sigma(\tau_j; j < 0) \lor \bigwedge_{k=-\infty}^0 \sigma\left(Y_j^{(\lambda)}[\omega(\cdot)]; j \leq k\right) = \sigma(\tau_j; j < 0) \tag{8.3} \]

since a tail $\sigma$-algebra is $\mathcal{T}$-invariant, equation (8.3) implies that

\[ \bigwedge_{n=-\infty}^0 \sigma(V_j; j \leq n) \subset \bigwedge_{k=-\infty}^0 \sigma(\tau_j; j < k) \]

This establishes the triviality of the left tail of $\{V_n\}$.

The conditional entropies can be estimated as follows:

\[ h\left(V_n \big| V_{n-1}, V_{n-2}, \ldots \right) = h\left(\tau_n \big| V_{n-1}, V_{n-2}, \ldots \right) + \mathbb{P}[\tau_n \neq 0]h(Y_0^{(\lambda)}[\cdot] \big| Y_{-1}^{(\lambda)}[\cdot] \ldots) \]

\[ = H(\alpha) + \frac{H(p) + \log \lambda}{N + \alpha + 1}. \]

Using this as the basic building block, we can represent any Bernoulli system as a product of arbitrarily many such building blocks with each $p$ in $[1/3, 2/3]$, and conclude the proof as before. \qed

Finally, we present a sketch of an alternative proof of Theorem 1.4. Let $\Gamma$ be an irreducible lattice in $SL_2(\mathbb{R}) \times SL_2(\mathbb{R})$. We take $X = SL_2(\mathbb{R}) \times SL_2(\mathbb{R})/\Gamma$ equipped with Haar measure as our measure space, and consider the action on $X$ from the left of the element

\[ g_{\alpha, \beta} = \left( \begin{pmatrix} e^\alpha & 0 \\ 0 & e^{-\alpha} \end{pmatrix}, \begin{pmatrix} e^\beta & 0 \\ 0 & e^{-\beta} \end{pmatrix} \right) \]

The entropy of this system (for $\alpha, \beta > 0$) is easily seen to be $2\alpha + 2\beta$. This system is Bernoulli by Dani [4].

Consider now the foliation of $X$ into the orbits of the group

\[ U = \left\{ \left( \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \right\}. \]

This foliation is invariant under $g_{\alpha, \beta}$, with each leaf expanding by a factor of $e^{2\alpha}$. Take $\xi$ to be a measurable partition subordinate to this foliation, i.e. that $\xi$ satisfies the following: for almost every $x$ the atom of $\xi$ containing $x$ is a subset of the $U$-orbit of $x$, and contains a small neighborhood of $x$ in the leaf (i.e. $U$-orbit). The existence of such a partition is easily established directly; it also follows from the much more general construction in [10],[11].

Using the ergodicity of the action of $U$ on $X$, one can show that any $\xi$ of this form is a $K$-partition. Furthermore, it is a rather standard fact that the entropy of this partition is the logarithm of the expansion rate – i.e. $2\alpha$. We omit the proof of these two claims.
By fixing the entropy \( h = 2\alpha + 2\beta \), and varying \( \alpha \), one obtains \( K \)-partitions of any entropy \( (0, h) \), for a Bernoulli system of arbitrary finite entropy \( h \). By considering products of such systems one also obtains \( K \)-partitions of arbitrary entropy in an infinite entropy Bernoulli system.

**Acknowledgement.** We are grateful to Ya. G. Sinai for suggesting the problem that led to this work, and to Benji Weiss for several helpful discussions.

**References**

17. Peres, Y., Solomyak, B. *Approximation by polynomials with coefficients \( \pm 1 \).* J. Number Theory, to appear.

Institute for Advanced Study, Olden Lane, Princeton, N.J. 08540
E-mail address: elon@ias.edu

Department of Statistics, University of California, Berkeley, CA 94720, and Department of Mathematics, Hebrew University, Jerusalem, Israel
E-mail address: peres@stat.berkeley.edu

Department of Mathematics, Princeton University, Fine Hall, Princeton N.J. 08544
E-mail address: schlag@math.princeton.edu