FOURIER DECAY FOR HOMOGENEOUS SELF-AFFINE MEASURES

BORIS SOLOMYAK

Abstract. We show that for Lebesgue almost all \( d \)-tuples \((\theta_1, \ldots, \theta_d)\), with \(|\theta_j| > 1\), any self-affine measure for a homogeneous non-degenerate iterated function system \(\{Ax + a_j\}_{j=1}^m\) in \(\mathbb{R}^d\), where \(A^{-1}\) is a diagonal matrix with the entries \((\theta_1, \ldots, \theta_d)\), has power Fourier decay at infinity.

1. Introduction

For a finite positive Borel measure \(\mu\) on \(\mathbb{R}^d\), consider the Fourier transform
\[
\hat{\mu}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i \langle \xi, x \rangle} d\mu(x).
\]

We are interested in the decay properties of \(\hat{\mu}\) at infinity. The measure \(\mu\) is called Rajchman if
\[
\lim_{|\xi| \to \infty} \hat{\mu}(\xi) = 0,
\]
where \(|\xi|\) is a norm (say, the Euclidean norm) of \(\xi \in \mathbb{R}^d\). Whereas absolutely continuous measures are Rajchman by the Riemann-Lebesgue Lemma, it is a subtle question to decide which singular measures are such, see, e.g., the survey of Lyons [14]. A much stronger property, useful for many applications is the following.

Definition 1.1. For \(\alpha > 0\) let
\[
\mathcal{D}_d(\alpha) = \{\nu \text{ finite positive measure on } \mathbb{R}^d : |\hat{\nu}(t)| = O(|t|^{-\alpha}), \ |t| \to \infty\},
\]
and denote \(\mathcal{D}_d = \bigcup_{\alpha > 0} \mathcal{D}_d(\alpha)\). A measure \(\nu\) is said to have power Fourier decay if \(\nu \in \mathcal{D}_d\).

Many recent papers have been devoted to the question of Fourier decay for classes of “fractal” measures, see e.g., [2, 9, 11, 12, 13, 18, 23, 3, 1, 25, 17]. Here we continue this line of research, focusing on the class of homogeneous self-affine measures in \(\mathbb{R}^d\). A measure \(\mu\) is called self-affine if it is the invariant measure for a self-affine iterated function system (IFS) \(\{f_j\}_{j=1}^m\), with \(m \geq 2\), where \(f_j(x) = A_j x + a_j\), the matrices \(A_j : \mathbb{R}^d \to \mathbb{R}^d\) are invertible linear contractions (in some norm) and \(a_j \in \mathbb{R}^d\) are “digit” vectors. This means that for some probability vector \(p = (p_j)_{j=1}^m\) holds
\[
(1.1) \quad \mu = \sum_{j=1}^m p_j (\mu \circ f_j^{-1}).
\]

Date: June 10, 2021.

Supported in part by the Israel Science Foundation grant 911/19.
It is well-known that this equation defines a unique probability Borel measure. The self-affine IFS is \textit{homogeneous} if all \( A_j \) are equal to each other: \( A = A_j \) for \( j \leq m \). Denote the digit set by \( \mathcal{D} := \{a_1, \ldots, a_n\} \) and the corresponding self-affine measure by \( \mu(A, \mathcal{D}, p) \). We will write \( p > 0 \) if all \( p_j > 0 \). Following [8], we say that the IFS is \textit{affinely irreducible} if the attractor is not contained in a proper affine subspace of \( \mathbb{R}^d \). It is easy to see that this is a necessary condition for the self-affine measure to be Rajchman, so this will always be our assumption. By a conjugation with a translation, we can always assume that \( 0 \in \mathcal{D} \). In this case affine irreducibility is equivalent to the digit set \( \mathcal{D} \) being a \textit{cyclic family} for \( A \), that is, \( \mathbb{R}^d \) being the smallest \( A \)-invariant subspace containing \( \mathcal{D} \).

The IFS is \textit{self-similar} if all \( A_j \) are contracting similitudes, that is, \( A_j = \lambda_j O_j \) for some \( \lambda_j \in (0, 1) \) and orthogonal matrices \( O_j \). In many aspects, “genuine” (i.e., non-self-similar) self-affine and self-similar IFS are very different; of course, the distinction exists only for \( d \geq 2 \).

Every homogeneous self-affine measure can be expressed as an infinite convolution product

\[
\mu(A, \mathcal{D}, p) = \left( \prod_{n=0}^{\infty} \right) \sum_{j=1}^{m} p_j \delta_{A^n a_j},
\]

and for every \( p > 0 \) it is supported on the attractor (self-affine set)

\[
K_{A, \mathcal{D}} := \left\{ x \in \mathbb{R}^d : x = \sum_{n=0}^{\infty} A^n b_n, \ b_n \in \mathcal{D} \right\}.
\]

By the definition of the self-affine measure,

\[
\hat{\mu}(\xi) = \sum_{j=1}^{m} p_j \int e^{-2\pi i \langle \xi, Ax + a_j \rangle} \, d\mu = \left( \sum_{j=1}^{m} p_j e^{-2\pi i \langle \xi, a_j \rangle} \right) \hat{\mu}(A^t \xi),
\]

where \( A^t \) is the matrix transpose of \( A \). Iterating we obtain

\[
\hat{\mu}(\xi) = \prod_{n=0}^{\infty} \left( \sum_{j=1}^{m} p_j e^{-2\pi i \langle (A^t)^n \xi, a_j \rangle} \right) = \prod_{n=0}^{\infty} \left( \sum_{j=1}^{m} p_j e^{-2\pi i \langle \xi, A^n a_j \rangle} \right),
\]

where the infinite product converges, since \( \|A^n\| \to 0 \) exponentially fast.

1.1. \textbf{Background.} We start with the known results on Fourier decay for classical Bernoulli convolutions \( \nu_\lambda \), namely, self-similar measures on the line, corresponding to the IFS \( \{\lambda x, \lambda x + 1\} \), with \( \lambda \in (0, 1) \) and probabilities \( \left( \frac{1}{2}, \frac{1}{2} \right) \) (often the digits \( \pm 1 \) are used instead; it is easy to see that taking any two distinct digits results in the same measure, up to an affine change of variable). Erdős [5] proved that \( \hat{\nu}_\lambda(t) \to 0 \) as \( t \to \infty \) when \( \theta = 1/\lambda \) is a \textit{Pisot number}. Recall that a Pisot number is an algebraic integer greater than one, whose algebraic (Galois) conjugates are all less than one in modulus. Salem [19] showed that if \( 1/\lambda \) is not a Pisot number, then \( \hat{\nu}_\lambda \) is a Rajchman measure.

In the other direction, Erdős [6] proved that for any \( [a, b] \subset (0, 1) \) there exists \( \alpha > 0 \) such that \( \nu_\lambda \in \mathcal{D}_\alpha(\alpha) \) for a.e. \( \lambda \in [a, b] \). Later, Kahane [10] indicated that Erdős’ argument actually gives
that $\nu_\lambda \in \mathcal{D}_1$ for all $\lambda \in (0, 1)$ outside a set of zero Hausdorff dimension. (We should mention that very few specific $\lambda$ are known, for which $\nu_\lambda$ has power Fourier decay, see Dai, Feng, and Wang [4].) In the original papers of Erdős and Kahane there were no explicit quantitative bounds; this was done in the survey [15], where the expression “Erdős-Kahane argument” was used first. The general case of a homogeneous self-similar measure on the line is treated analogously to Bernoulli convolutions: the self-similar measure is still an infinite convolution and the Erdős-Kahane argument on power Fourier decay goes through with minor modifications, see [4, 22]. Although one of the main motivations for the study of the Fourier transform has been the question of absolute continuity/singularity of $\nu_\lambda$, here we do not discuss it but refer the reader to the recent survey [24].

Next we turn to the non-homogeneous case on the line. Li and Sahlsten [12] proved that if $\mu$ is a self-similar measure on the line with contraction ratios $\{r_i\}_{i=1}^m$ and there exist $i \neq j$ such that $\log r_i/\log r_j$ is irrational, then $\mu$ is Rajchman. Moreover, they showed logarithmic decay of the Fourier transform under a Diophantine condition. A related result for self-conformal measures was recently obtained by Algom, Rodriguez Hertz, and Wang [1]. Brémont [3] obtained an (almost) complete characterization of (non)-Rajchman self-similar measures in the case when $r_j = \lambda^{\nu_j}$ for $j \leq m$. To be non-Rajchman, it is necessary for $1/\lambda$ to be Pisot. For “generic” choices of the probability vector $p$, assuming that $D \subset \mathbb{Q}(\lambda)$ after an affine conjugation, this is also sufficient, but there are some exceptional cases of positive co-dimension. Várjú and Yu [25] proved logarithmic decay of the Fourier transform in the case when $r_j = \lambda^{\nu_j}$ for $j \leq m$ and $1/\lambda$ is algebraic, but not a Pisot or Salem number. In [23] we showed that outside a zero Hausdorff dimension exceptional set of parameters, all self-similar measures on $\mathbb{R}$ belong to $\mathcal{D}_1$; however, the exceptional set is not explicit.

Turning to higher dimensions, we mention the recent paper by Rapaport [17], where he gives an algebraic characterization of self-similar IFS for which there exists a probability vector yielding a non-Rajchman self-similar measure. Li and Sahlsten [13] investigated self-affine measures in $\mathbb{R}^d$ and obtained power Fourier decay under some algebraic conditions, which never hold for a homogeneous self-affine IFS. Their main assumptions are total irreducibility of the closed group generated by the contraction linear maps $A_j$ and non-compactness of the projection of this group to $PGL(d, \mathbb{R})$. For $d = 2, 3$ they showed that this is sufficient.

1.2. Statement of results. We assume that $A$ is a matrix diagonalizable over $\mathbb{R}$. Then we can reduce the IFS, via a linear change of variable, to one where $A$ is a diagonal matrix with real entries. Given $A = \text{Diag}[\theta_1^{-1}, \ldots, \theta_d^{-1}]$, with $|\theta_j| > 1$, a set of digits $D = \{a_1, \ldots, a_m\} \subset \mathbb{R}^d$, and a probability vector $p$, we write $\theta = (\theta_1, \ldots, \theta_d)$ and denote by $\mu(\theta, D, p)$ the self-affine measure defined by (1.1). Our main motivation is the class of measures which can be viewed as “self-affine Bernoulli convolutions”, with $A = \text{Diag}[\theta_1^{-1}, \ldots, \theta_d^{-1}]$ a diagonal matrix with distinct real entries and $D = \{0, (1, \ldots, 1)\}$. In this special case we denote the self-affine measure by $\mu(\theta, p)$. 
Theorem 1.2. There exists an exceptional set $E \subset \mathbb{R}^d$, with $\mathcal{L}^d(E) = 0$, such that for all $\theta \in \mathbb{R}^d \setminus E$, with $\min_j |\theta_j| > 1$, for all sets of digits $D$, such that the IFS is affinely irreducible, and all $p > 0$, holds $\mu(\theta, D, p) \in \mathcal{D}_d$.

The theorem is a consequence of a more quantitative statement.

Theorem 1.3. Fix $1 < b_1 < b_2 < \infty$ and $c_1, \varepsilon > 0$. Then there exist $\alpha > 0$ and $\varepsilon = c_1 = M^{-1}$. Then the set
\[ E = \bigcup_{M=2}^\infty \mathcal{E}(M) \cup \{ \theta : \exists i \neq j, \theta_i = \theta_j \}. \]

Reduction of Theorem 1.2 to Theorem 1.3. For $M \in \mathbb{N}$ let $\mathcal{E}(M)$ be the exceptional set obtained from Theorem 1.3 with $b_1 = 1 + M^{-1}, b_2 = M$, and $\varepsilon = c_1 = M^{-1}$. Then the set
\[ E = \bigcup_{M=2}^\infty \mathcal{E}(M) \cup \{ \theta : \exists i \neq j, \theta_i = \theta_j \}. \]

has the desired properties.

The proof of Theorem 1.3 uses a version of the Erdős-Kahane technique. We follow the general scheme of [15, 22], but this is not a trivial extension.

In view of the convolution structure, Theorem 1.3 yields some information on absolute continuity of self-affine measures, by a standard argument.

Corollary 1.4. Fix $1 < b_1 < b_2 < \infty$ and $c_1, \varepsilon > 0$. Then there exist a sequence $n_k \to \infty$ and $\tilde{E}_k \subset \mathbb{R}^d$, depending on these parameters, such that $\mathcal{L}^d(\tilde{E}_k) = 0$ and for all $\theta \notin \tilde{E}_k$ satisfying
\[ b_1 \leq \min_j |\theta_j| < \max_j |\theta_j| \leq b_2 \quad \text{and} \quad |\theta_i - \theta_j| \geq c_1, i \neq j, \]

for all digit sets $D$ such that the IFS is affinely irreducible, and all $p$ such that $\min_j p_j \geq \varepsilon$, the measure $\mu(\theta, D, p)$ is absolutely continuous with respect to $\mathcal{L}^d$, with a Radon-Nikodym derivative in $C^k(\mathbb{R}^d)$, $k \geq 0$.

Proof (derivation). Let $n \geq 2$. It follows from (1.2) that
\[ \mu(A, D, p) = \mu(A^n, D, p) \ast \mu(A^n, AD, p) \ldots \ast \mu(A^n, A^{n-1}D, p). \]

It is easy to see that if the original IFS is affinely irreducible, then so are the IFS associated with $(A^n, A^iD)$, and moreover, these IFS are all affine conjugate to each other. Therefore, if $\mu(A^n, D, p) \in \mathcal{D}_d(\alpha)$, then $\mu(A, D, p) \in \mathcal{D}_d(n\alpha)$. As is well-known,
\[ \mu \in \mathcal{D}_d(\beta), \beta > d + k \implies \frac{d\mu}{d\mathcal{L}^d} \in C^k(\mathbb{R}^d), \]
so we can take \( n_k \) such that \( n_k \alpha > d + k \), and \( \mathcal{E}_k = \{ \theta : \theta^{n_k} \in \mathcal{E} \} \), where \( \alpha \) and \( \mathcal{E} \) are from Theorem 1.3.

Remark 1.5. (a) In general, the power decay cannot hold for all \( \theta \); for instance, it is easy to see that the measure \( \mu(\theta, p) \) is not Rajchman if at least one of \( \theta_k \) is a Pisot number. Thus in the most basic case with two digits, the exceptional set has Hausdorff dimension at least \( d - 1 \).

(b) It is natural to ask what happens if \( A \) is not diagonalizable over \( \mathbb{R} \). A complex eigenvalue of \( A \) corresponds to a 2-dimensional homogeneous self-similar IFS with rotation, or an IFS of the form \( \{ \lambda z + a_j \}_{j=1}^m \), with \( \lambda \in \mathbb{C} \), \( |\lambda| < 1 \), and \( a_j \in \mathbb{C} \). In [21] it was shown that for all \( \lambda \) outside a set of Hausdorff dimension zero, the corresponding self-similar measure belongs to \( \mathcal{D}_2 \). It may be possible to combine the methods of [21] with those of the current paper to obtain power Fourier decay for a typical \( A \) diagonalizable over \( \mathbb{C} \). It would also be interesting to consider the case of non-diagonalizable \( A \), starting with a single Jordan block.

(c) In the special case of \( d = 2 \) and \( m = 2 \), our system reduces to a planar self-affine IFS, conjugate to \( \{ (\lambda x, \gamma y) \pm (-1,1) \} \) for \( 0 < \gamma < \lambda < 1 \). This system has been studied by many authors, especially the dimension and topological properties of its attractor, see [7] and the references therein. For our work, the most relevant is the paper by Shmerkin [20]. Among other results, he proved absolute continuity with a density in \( L^2 \) of the self-affine measure (with some fixed probabilities) almost everywhere in some region, in particular, in some explicit neighborhood of \((1,1)\). He also showed that if \((\lambda^{-1}, \gamma^{-1})\) for a Pisot pair, then the measure is not Rajchman and hence singular.

1.3. Rajchman self-affine measures. The question “when is \( \mu(A, \mathcal{D}, p) \) is Rajchman?” is not addressed here. Recently Rapaport [17] obtained an (almost) complete characterization of self-similar Rajchman measures in \( \mathbb{R}^d \). Of course, our situation is vastly simplified by the assumption that the IFS is homogeneous, but still it is not completely straightforward. The key notion here is the following.

Definition 1.6. A collection of numbers \((\theta_1, \ldots, \theta_m)\) (real or complex) is called a Pisot family or a P.V. \( m \)-tuple if

(i) \(|\theta_j| > 1\) for all \( j \leq m \) and

(ii) there is a monic integer polynomial \( P(t) \), such that \( P(\theta_j) = 0 \) for all \( j \leq m \), whereas every other root \( \theta' \) of \( P(t) \) satisfies \(|\theta'| < 1\).

It is not difficult to show, using the classical techniques of Pisot [16] and Salem [19], as well as some ideas from [17] Section 5] that

- If \( \mu(A, \mathcal{D}, p) \) is not a Rajchman measure and the IFS is affinely irreducible, then the spectrum \( \text{Spec}(A^{-1}) \) contains a Pisot family;
- if \( \text{Spec}(A^{-1}) \) contains a Pisot family, then for a “generic” choice of \( \mathcal{D} \), with \( m \geq 3 \), the measure \( \mu(A, \mathcal{D}, p) \) is Rajchman; however,
• if Spec($A^{-1}$) contains a Pisot family, then under appropriate conditions the measure $\mu(A, D, p)$ is not Rajchman. For instance, this holds if there is at least one conjugate of the elements of the Pisot family less than 1 in absolute value, $m = 2$, and $A$ is diagonalizable over $\mathbb{R}$.

We omit the details.

2. Proofs

The following is an elementary inequality.

**Lemma 2.1.** Let $p = (p_1, \ldots, p_m) > 0$ be a probability vector and $\alpha_1 = 0$, $\alpha_j \in \mathbb{R}$, $j = 2, \ldots, m$. Denote $\varepsilon = \min_j p_j$ and write $\|x\| = \text{dist}(x, \mathbb{Z})$. Then for any $k \leq m$,

$$\left| \sum_{j=1}^{m} p_j e^{-2\pi i \alpha_j} \right| \leq 1 - 2\pi \varepsilon \|\alpha_k\|^2.$$  

**Proof.** Fix $k \in \{2, \ldots, m\}$. We can estimate

$$\left| \sum_{j=1}^{m} p_j e^{-2\pi i \alpha_j} \right| = \left| p_1 + \sum_{j=2}^{m} p_j e^{-2\pi i \alpha_j} \right| \leq |p_1 + p_k e^{-2\pi i \alpha_k}| + (1 - p_1 - p_k).$$

Assume that $p_1 \geq p_k$, otherwise, write $|p_1 + p_k e^{-2\pi i \alpha_k}| = |p_1 e^{2\pi i \alpha_k} + p_k|$ and repeat the argument. Then observe that $|p_1 + p_k e^{-2\pi i \alpha_k}| \leq (p_1 - p_k) + p_k|1 + e^{-2\pi i \alpha_k}|$ and $|1 + e^{-2\pi i \alpha_k}| = 2|\cos(\pi \alpha_k)| \leq 2(1 - \pi \|\alpha_k\|^2)$. This implies the desired inequality.

Recall (1.3):

$$\hat{\mu}(\xi) = \prod_{n=0}^{\infty} \left( \sum_{j=1}^{m} p_j e^{-2\pi i \langle \xi, A^n a_j \rangle} \right).$$

For $\xi \in \mathbb{R}^d$, with $\|\xi\|_\infty \geq 1$, let $\eta(\xi) = (A^t)^{N(\xi)} \xi$, where $N(\xi) \geq 0$ is maximal, such that $\|\eta(\xi)\|_\infty \geq 1$. Then $\|\eta(\xi)\|_\infty \in [1, \|A^t\|_\infty]$ and (1.3) implies

$$\hat{\mu}(\xi) = \hat{\mu}(\eta(\xi)) \cdot \prod_{n=1}^{N(\xi)} \left( \sum_{j=1}^{m} p_j e^{-2\pi i \langle \eta(\xi), A^{-n} a_j \rangle} \right).$$

2.1. **Proof of Theorem 1.3.** First we show that the case of a general digit set may be reduced to $D = \{0, 1, \ldots, 1\}$. We start with the formula (2.2), which under the current assumptions becomes

$$\hat{\mu}(\xi) = \hat{\mu}(\eta(\xi)) \cdot \prod_{n=1}^{N(\xi)} \left( \sum_{j=1}^{m} p_j \exp \left[ -2\pi i \sum_{k=1}^{d} \eta_k a_j^{(k)} \theta_k^n \right] \right),$$

where $\theta_k = e^{2\pi i \alpha_k}.$
where \( a_j = (a_j^{(k)})_{k=1}^d \) and \( \eta(\xi) = (\eta_k)^d_{k=1} \). Note that \( \|\eta(\xi)\|_\infty \in [1, \max_j |\theta_j|] \). Assume without loss of generality that \( a_1 = 0 \), then we have by (2.1), for any fixed \( j \in \{2, \ldots, m\} \):

\[
|\hat{\mu}(\xi)| \leq \prod_{n=1}^{N(\xi)} \left( 1 - 2\pi \| \sum_{k=1}^d \eta_k a_j^{(k)} \theta^n_k \| \right),
\]

where \( \| \cdot \| \) denotes the distance to the nearest integer. Further, we can assume that all the coordinates of \( a_j \) are non-zero; otherwise, we can work in the subspace

\[ H := \{ x \in \mathbb{R}^d : x_k = 0 \iff a_j^{(k)} = 0 \} \]

and with the corresponding variables \( \theta_k \), and then get the exceptional set of zero \( \mathcal{L}^d \) measure as a product of a set of zero measure in \( H \) and the entire \( H^\perp \). Finally, apply a linear change of variables, so that \( a_j^{(k)} = 1 \) for all \( k \), to obtain:

\[
(2.3) \quad |\hat{\mu}(\xi)| \leq \prod_{n=1}^{N(\xi)} \left( 1 - 2\pi \| \sum_{k=1}^d \eta_k \theta^n_k \| \right).
\]

This is exactly the situation corresponding to the measure \( \mu(\theta, p) \), and we will be showing (typical) power decay for the right-hand side of (2.3). This completes the reduction.

Next we use a variant of the Erdős-Kahane argument, see e.g. [15, 22] for other versions of it. Intuitively, we will get power decay if \( |\sum_{k=1}^d \eta_k \theta^n_k| \) is uniformly bounded away from zero for a set of \( n \)'s of positive lower density, uniformly in \( \eta \).

Fix \( c_1 > 0 \) and \( 1 < b_1 < b_2 < \infty \), and consider the compact set

\[ H = \{ \theta = (\theta_1, \ldots, \theta_d) \in (-b_2, -b_1] \cup [b_1, b_2]^d : |\theta_i - \theta_j| \geq c_1, \; i \neq j \}. \]

We will use the notation \( [N] = \{1, \ldots, N\} \), \( [n, N] = \{n, \ldots, N\} \). For \( \rho, \delta > 0 \) we define the “bad set” at scale \( N \):

\[
(2.4) \quad E_{H,N}(\delta, \rho) = \left\{ \theta \in H : \max_{\eta : 1 \leq |\eta|_\infty \leq b_2} \frac{1}{N} \left\{ n \in [N] : \left\| \sum_{k=1}^d \eta_k \theta^n_k \right\| < \rho \right\} > 1 - \delta \right\}.
\]

Now we can define the exceptional set:

\[ \mathcal{E}_H(\delta, \rho) := \bigcup_{N_0 = 1}^\infty \bigcup_{N = N_0}^\infty E_{H,N}(\delta, \rho). \]

Theorem [1.3] will immediately follow from the next two propositions.

**Proposition 2.2.** For any positive \( \rho \) and \( \delta \), we have \( \mu(\theta, p) \in \mathcal{D}(\alpha) \) whenever \( \theta \in H \setminus \mathcal{E}_H(\delta, \rho) \), where \( \alpha \) depends only on \( \delta, \rho, H \), and \( \varepsilon = \min\{p, 1 - p\} \).

**Proposition 2.3.** There exist \( \rho = \rho_H > 0 \) and \( \delta = \delta_H > 0 \) such that \( \mathcal{L}^d(\mathcal{E}_H(\delta, \rho)) = 0 \).
Proof of Proposition 2.2. Suppose that \( \theta \in \mathcal{H} \setminus \mathcal{E}_H(\delta, \rho) \). This implies that there is \( N_0 \in \mathbb{N} \) such that \( \theta \notin \mathcal{E}_{H,N}(\delta, \rho) \) for all \( N \geq N_0 \). Let \( \xi \in \mathbb{R}^d \) be such that \( \|\xi\|_{\infty} > b_2^{N_0} \). Then \( N = N(\xi) \geq N_0 \), where \( \eta = \eta(\xi) = A^N(\xi) \) and \( N(\xi) \) is maximal with \( \|\eta\|_{\infty} \geq 1 \). From the fact that \( \theta \notin \mathcal{E}_{H,N}(\delta, \rho) \) it follows that

\[
\frac{1}{N} \left| \left\{ n \in [N] : \sum_{k=1}^d \eta_k \theta_k^n < \rho \right\} \right| \leq 1 - \delta.
\]

Then by (2.3),

\[
\left| \hat{\mu}(\theta, p)(\xi) \right| \leq (1 - 2\pi \varepsilon \rho^2)^{|\delta N|}.
\]

By the definition of \( N = N(\xi) \) we have

\[
\|\xi\|_{\infty} \leq b_2^{N+1}.
\]

It follows that

\[
\left| \hat{\mu}(\theta, p)(\xi) \right| = O_{H, \varepsilon}(1) \cdot \|\xi\|_{\infty}^{-\alpha},
\]

for \( \alpha = -\delta \log(1 - 2\pi \varepsilon \rho^2) / \log b_2 \), and the proof is complete. \( \square \)

Proof of Proposition 2.3. It is convenient to express the exceptional set as a union, according to a dominant coordinate of \( \eta \) (which may be non-unique, of course): \( \mathcal{E}_{H,N}(\delta, \rho) = \bigcup_{j=1}^d \mathcal{E}_{H,N,j}(\delta, \rho) \), where

\[(2.5) E_{H,N,j}(\delta, \rho) := \left\{ \theta \in H : \exists \eta, \quad 1 \leq |\eta_j| = \|\eta\|_{\infty} \leq b_2, \quad \frac{1}{N} \left| \left\{ n \in [N] : \sum_{k=1}^d \eta_k \theta_k^n < \rho \right\} \right| > 1 - \delta \right\}.
\]

It is easy to see that \( E_{H,N,j}(\delta, \rho) \) is measurable. Observe that

\[
\mathcal{E}_H(\delta, \rho) := \bigcup_{j=1}^d \mathcal{E}_{H,j}(\delta, \rho), \quad \text{where} \quad \mathcal{E}_{H,j}(\delta, \rho) := \bigcap_{N=1}^\infty \bigcup_{N_0=1}^{N_0} E_{H,N,j}(\delta, \rho).
\]

It is, of course, sufficient to show that \( \mathcal{L}^d(\mathcal{E}_{H,j}(\delta, \rho)) = 0 \) for every \( j \in [d] \), for some \( \delta, \rho > 0 \). Without loss of generality, assume that \( j = d \). Since \( \mathcal{E}_{H,d}(\delta, \rho) \) is measurable, the desired claim will follow if we prove that every slice of \( \mathcal{E}_{H,d}(\delta, \rho) \) in the direction of the \( x_d \)-axis has zero \( \mathcal{L}^1 \) measure. Namely, for fixed \( \theta' = (\theta_1, \ldots, \theta_{d-1}) \) let

\[
\mathcal{E}_{H,d}(\delta, \rho, \theta') := \{ \theta_d : (\theta', \theta_d) \in \mathcal{E}_{H,d}(\delta, \rho) \}.
\]

We want to show that \( \mathcal{L}^1(\mathcal{E}_{H,d}(\delta, \rho, \theta')) = 0 \) for all \( \theta' \). Clearly,

\[
\mathcal{E}_{H,d}(\delta, \rho, \theta') := \bigcap_{N_0=1}^{N_0} \bigcup_{N=1}^{N_0} E_{H,N,d}(\delta, \rho, \theta'),
\]
where
\begin{equation}
E_{H,N,d}(\delta, \rho, \theta') = \left\{ \theta_d : (\theta', \theta_d) \in H : \max_{\eta: 1 \leq |\eta_d| \leq b_2} \frac{1}{N} \left| \left\{ n \in [N] : \left| \sum_{k=1}^{d} \eta_k \theta_k^n \right| < \rho \right\} \right| > 1 - \delta \right\}
\end{equation}

**Lemma 2.4.** There exists a constant $\rho > 0$ such that, for any $N \in \mathbb{N}$ and $\delta \in (0, \frac{1}{2})$, the set $E_{H,N,d}(\delta, \rho, \theta')$ can be covered by $\exp(O_{H}(\delta \log(1/\delta)N))$ intervals of length $b_1^{-N}$.

We first complete the proof of the proposition, assuming the lemma. By Lemma 2.4,
\begin{equation}
\mathcal{L}^1 \left( \bigcup_{N=N_0}^{\infty} E_{H,N,d}(\delta, \rho, \theta') \right) \leq \sum_{N=N_0}^{\infty} \exp(O_{H}(\delta \log(1/\delta)N)) \cdot b_1^{-N} \to 0, \quad N_0 \to \infty,
\end{equation}

provided $\delta > 0$ is so small that $\log b_1 > O_{H}(\delta \log(1/\delta))$. Thus $\mathcal{L}^1(E_{H,d}(\delta, \rho, \theta')) = 0$. \hfill \Box

**Proof of Lemma 2.4.** Fix $\theta'$ in the projection of $H$ to the first $(d-1)$ coordinates and $\eta \in \mathbb{R}^d$, with $1 \leq |\eta_d| = \|\eta\|_{\infty} \leq b_2$. Below all the constants implicit in the $O(\cdot)$ notation are allowed to depend on $H$ and $d$. Let $\theta_d$ be such that $(\theta', \theta_d) \in H$ and write
\begin{equation}
\sum_{k=1}^{d} \eta_k \theta_k^n = K_n + \varepsilon_n, \quad n \geq 0,
\end{equation}

where $K_n \in \mathbb{Z}$ is the nearest integer to the expression in the left-hand side, so that $|\varepsilon_n| \leq \frac{1}{2}$. We emphasize that $K_n$ depends on $\eta$ and on $\theta_d$. Define $A_n^{(0)} = K_n$, $\tilde{A}_n^{(0)} = K_n + \varepsilon_n$, and then for all $n$ inductively:
\begin{equation}
A_n^{(j)} = A_{n+1}^{(j-1)} - \theta_j A_n^{(j-1)}; \quad \tilde{A}_n^{(j)} = \tilde{A}_{n+1}^{(j-1)} - \theta_j \tilde{A}_n^{(j-1)}, \quad j = 1, \ldots, d - 1.
\end{equation}

It is easy to check by induction that
\begin{equation}
\tilde{A}_n^{(j)} = \sum_{i=j+1}^{d} \eta_i \prod_{k=1}^{j} (\theta_i - \theta_k) \theta_i^n, \quad j = 1, \ldots, d - 1,
\end{equation}

hence
\begin{equation}
\tilde{A}_n^{(d-1)} = \eta_d \prod_{k=1}^{d-1} (\theta_d - \theta_k) \theta_d^n; \quad \theta_d = \frac{\tilde{A}_{n+1}^{(d-1)}}{\tilde{A}_n^{(d-1)}}, \quad n \in \mathbb{N}.
\end{equation}

We have $\|\eta\|_{\infty} \leq b_2$ and $|\tilde{A}_n^{(0)} - A_n^{(0)}| \leq |\varepsilon_n|$, and then by induction, by (2.7),
\begin{equation}
|\tilde{A}_n^{(j)} - A_n^{(j)}| \leq (1 + b_2)^j \max\{|\varepsilon_n|, \ldots, |\varepsilon_{n+j}|\}, \quad j = 1, \ldots, d - 1.
\end{equation}
Another easy calculation gives
\[ K_{n+d+1} = \theta_1 K_{n+d} + A_{n+d}^{(1)} + \cdots \]
(2.10)
\[ = [\theta_1 K_{n+d} + \theta_2 A_{n+d-1}^{(1)} + \cdots + \theta_{d-1} A_{n+2}^{(d-2)}] + A_{n+2}^{(d-1)} \]
Since \( \frac{A_{n+2}^{(d-1)}}{A_{n+1}^{(d-1)}} \approx A_{n+1}^{(d-1)} = \theta_d \), we have
\[ K_{n+d+1} \approx \left[ \theta_1 K_{n+d} + \theta_2 A_{n+d-1}^{(1)} + \cdots + \theta_{d-1} A_{n+2}^{(d-2)} \right] + \frac{(A_{n+1}^{(d-1)})^2}{A_{n+1}^{(d-1)}} \]
(2.11)
where \( R_{\theta_1, \ldots, \theta_{d-1}}(K_n, \ldots, K_{n+d}) \) is a rational function, depending on the (fixed) parameters \( \theta_1, \ldots, \theta_{d-1} \). To make the approximate equality precise, note that by (2.8) and our assumptions,
\[ |\tilde{A}_n^{(d-1)}| \geq c_1^{-1} b_1^n, \]
where \( b_1 > 1 \), and \( |\tilde{A}_n^{(d-1)} - A_n^{(d-1)}| \leq (1 + b_2)d^{-1}/2 \) by (2.9). Hence
\[ |A_n^{(d-1)}| \geq c_1^{-1} b_1^n/2 \text{ for } n \geq n_0 = n_0(H), \]
and so
\[ |A_n^{(d-1)} / A_n^{(d-1)}| \leq O(1), \quad n \geq n_0. \]
In the next estimates we assume that \( n \geq n_0(H) \). In view of the above, especially (2.9) for \( j = d-1 \),
\[ \left| \frac{A_{n+1}^{(d-1)}}{A_n^{(d-1)}} - \theta_d \right| = \left| \frac{A_{n+1}^{(d-1)}}{A_n^{(d-1)}} - \tilde{A}_{n+1}^{(d-1)} / \tilde{A}_n^{(d-1)} \right| \]
\[ \leq \left| \frac{A_{n+1}^{(d-1)} - \tilde{A}_{n+1}^{(d-1)}}{A_n^{(d-1)}} \right| + \left| \tilde{A}_{n+1}^{(d-1)} \right| \cdot \left| \frac{1}{A_n^{(d-1)}} - \frac{1}{\tilde{A}_n^{(d-1)}} \right| \]
\[ \leq O(1) \cdot \max\{|\varepsilon_n|, \ldots, |\varepsilon_{n+d}|\} \cdot |A_n^{(d-1)}|^{-1}. \]
It follows that, on the one hand,
\[ \left| \frac{A_{n+1}^{(d-1)}}{A_n^{(d-1)}} - \theta_d \right| \leq O(1) \cdot b_1^{-n}, \]
(2.13)
and on the other hand,
\[ \left| \left( \frac{A_{n+1}^{(d-1)}}{A_n^{(d-1)}} \right)^2 - A_{n+2}^{(d-1)} \right| \leq O(1) \cdot \max\{|\varepsilon_n|, \ldots, |\varepsilon_{n+d+1}|\}. \]
(2.14)
Note that $A_n^{(j)}$, for $j \in [d-1]$, is a linear combination of $K_n, K_{n+1}, \ldots, K_{n+j}$ with coefficients that are polynomials in the (fixed) parameters $\theta_1, \ldots, \theta_{d-1}$, hence the inequality (2.13) shows that

\[(2.15) \quad \text{given } K_n, \ldots, K_{n+d}, \text{ we have an } O(1) \cdot b_1^{-n}\text{-approximation of } \theta_d.\]

The inequality (2.14) yields, using (2.11) and (2.10), that, for $n \geq n_0$,

\[|K_{n+d+1} - R_{\theta_1, \ldots, \theta_{d-1}}(K_n, \ldots, K_{n+d})| \leq O(1) \cdot \max\{|\varepsilon_n|, \ldots, |\varepsilon_{n+d+1}|\}.\]

Thus we have:

(i) Given $K_n, \ldots, K_{n+d}$, there are at most $O(1)$ possible values for $K_{n+d+1}$, uniformly in $\eta$ and $\theta_1, \ldots, \theta_{d-1}$. There are also $O(1)$ possible values for $K_1, \ldots, K_{n_0}$ since $\|\eta\|_\infty$ and $\|\theta\|$ are bounded above by $b_2$.

(ii) There is a constant $\rho = \rho(H) > 0$ such that if $\max\{|\varepsilon_n|, \ldots, |\varepsilon_{n+d+1}|\} < \rho$, then $K_n, \ldots, K_{n+d}$ uniquely determine $K_{n+d+1}$, as the nearest integer to $R_{\theta_1, \ldots, \theta_{d-1}}(K_n, \ldots, K_{n+d})$, again independently of $\eta$ and $\theta_1, \ldots, \theta_{d-1}$.

Fix an $N$ sufficiently large. We claim that for each fixed set $J \subset [N]$ with $|J| \geq (1 - \delta)N$, the set

\[\{(K_n)_{n \in [N]} : \varepsilon_n = \|\sum_{k=1}^d \eta_k \theta_k^n\| < \rho \text{ for some } \theta_d, \eta \text{ and all } n \in J\}\]

has cardinality $\exp(O(\delta N))$. Indeed, fix such a $J$ and let

\[\tilde{J} = \{i \in [n_0 + (d + 1), N] : i, i - 1, \ldots, i - (d + 1) \in J\}.

We have $|\tilde{J}| \geq (1 - (d + 2)\delta)N - n_0 - (d + 1)$. If we set

\[\Lambda_j = (K_i)_{i \in [j]},\]

then (i), (ii) above show that $|\Lambda_{j+1}| = |\Lambda_j|$ if $j \in \tilde{J}$ and $|\Lambda_{j+1}| = O(|\Lambda_j|)$ otherwise. Thus $|\Lambda_N| \leq O(1)^{(d+2)\delta N}$, as claimed.

The number of subsets $A$ of $[N]$ of size $\geq (1 - \delta)N$ is bounded by $\exp(O(\delta \log(1/\delta)N))$ (using e.g. Stirling's formula), so we conclude that there are

\[\exp(O(\delta \log(1/\delta)N)) \cdot \exp(O(\delta N)) = \exp(O(\delta \log(1/\delta)N))\]

sequences $K_1, \ldots, K_N$ such that $|\varepsilon_n| < \rho$ for at least $(1 - \delta)N$ values of $n \in [N]$. Hence by (2.15) the set (2.6) can be covered by $\exp(O_H(\delta \log(1/\delta)N))$ intervals of radius $b_1^{-N}$, as desired. \hfill \Box

The proof of Theorem 1.3 is now complete.

Acknowledgement. Thanks to Ariel Rapaport for corrections and helpful comments on a preliminary version.
References

[1] Amir Algom, Federico Rodríguez Hertz, and Zhiren Wang. Pointwise normality and Fourier decay for self-conformal measures. arXiv e-prints, December 2020. arXiv:2012.06529.

[2] Jean Bourgain and Semyon Dyatlov. Fourier dimension and spectral gaps for hyperbolic surfaces. Geom. Funct. Anal., 27(4):744–771, 2017.

[3] Julien Brémont. Self-similar measures and the Rajchman property. arXiv e-prints, October 2019. arXiv:1910.03463.

[4] Xin-Rong Dai, De-Jun Feng, and Yang Wang. Refinable functions with non-integer dilations. J. Funct. Anal., 250(1):1–20, 2007.

[5] Paul Erdős. On a family of symmetric Bernoulli convolutions. Amer. J. Math., 61:974–976, 1939.

[6] Paul Erdős. On the smoothness properties of a family of Bernoulli convolutions. Amer. J. Math., 62:180–186, 1940.

[7] Kevin G. Hare and Nikita Sidorov. On a family of self-affine sets: topology, uniqueness, simultaneous expansions. Ergodic Theory Dynam. Systems, 37(1):193–227, 2017.

[8] Michael Hochman. On self-similar sets with overlaps and inverse theorems for entropy in $\mathbb{R}^d$. arXiv e-prints, Memoirs of the AMS, to appear, page arXiv:1503.09043, March 2015.

[9] Thomas Jordan and Tuomas Sahlsten. Fourier transforms of Gibbs measures for the Gauss map. Math. Ann., 364(3-4):983–1023, 2016.

[10] J.-P. Kahane. Sur la distribution de certaines séries aléatoires. Bull. Soc. Math. France, Mém. No. 25, Soc. Math. France Paris, pages 119–122, 1971.

[11] Jialun Li. Decrease of Fourier coefficients of stationary measures. Math. Ann., 372(3-4):1189–1238, 2018.

[12] Jialun Li and Tuomas Sahlsten. Trigonometric series and self-similar sets. arXiv e-prints, Feb 2019. arXiv:1902.00426.

[13] Jialun Li and Tuomas Sahlsten. Fourier transform of self-affine measures. arXiv e-prints, Sep 2020.

[14] Russell Lyons. Seventy years of Rajchman measures. J. Fourier Anal. Appl., pages 363–377, 1995.

[15] Yuval Peres, Wilhelm Schlag, and Boris Solomyak. Sixty years of Bernoulli convolutions. In Fractal geometry and stochastics, II (Greifswald/Koserow, 1998), volume 46 of Progr. Probab., pages 39–65. Birkhäuser, Basel, 2000.

[16] Charles Pisot. La répartition modulo 1 et les nombres algébriques. Ann. Scuola Norm. Super. Pisa Cl. Sci. (2), 7(3-4):205–248, 1938.

[17] Ariel Rapaport. On the Rajchman property for self-similar measures on $\mathbb{R}^d$. arXiv e-prints, April 2021. arXiv:2104.03095.

[18] Tuomas Sahlsten and Connor Stevens. Fourier decay in nonlinear dynamics. arXiv e-prints, Oct 2018. arXiv:1810.01378.

[19] Raphael Salem. Sets of uniqueness and sets of multiplicity. Trans. Amer. Math. Soc., 54:218–228, 1943.

[20] Pablo Shmerkin. Overlapping self-affine sets. Indiana Univ. Math. J., 55(4):1291–1331, 2006.

[21] Pablo Shmerkin and Boris Solomyak. Absolute continuity of complex Bernoulli convolutions. Math. Proc. Cambridge Philos. Soc., 161(3):435–453, 2016.

[22] Pablo Shmerkin and Boris Solomyak. Absolute continuity of self-similar measures, their projections and convolutions. Trans. Amer. Math. Soc., 368(7):5125–5151, 2016.

[23] Boris Solomyak. Fourier decay for self-similar measures. arXiv e-prints, Proc. of the AMS, to appear, June 2019. arXiv:1906.12164.
[24] Péter P. Varjú. Recent progress on Bernoulli convolutions. In European Congress of Mathematics, pages 847–867. Eur. Math. Soc., Zürich, 2018.

[25] Péter P. Varjú and Han Yu. Fourier decay of self-similar measures and self-similar sets of uniqueness. arXiv e-prints, April 2020. arXiv:2004.09358.

Boris Solomyak, Department of Mathematics, Bar-Ilan University, Ramat Gan, 5290002 Israel

Email address: bsolom3@gmail.com