An introduction to Fourier decay

Vilma Orgoványi and Alex Rutar

November 5, 2024

Contents

1 Introduction to Fourier analysis of fractal measures **1**

1.1 Self-similar sets and measures . . . 1

1.2 Bernoulli convolutions . . . 2

1.3 The L^1 Fourier transform . . . 3

1.4 Schwartz space and the L^2 Fourier transform . . . 5

1.5 Convolution and approximate identities . . . 7

1.6 A few words about tempered distributions . . . 9

2 Fourier decay of fractal measures **10**

2.1 An introduction to Fourier decay . . . 10

2.2 Fourier transform and dimension . . . 13

2.3 Power Fourier decay and absolute continuity . . . 16

1 Introduction to Fourier analysis of fractal measures

We begin with an introduction to Fourier analysis of fractal measures. These notes are mainly based on the books of Mattila [3] and Grafakos [2], along with Wolff's lecture notes [5], and occasionally Folland [1].

1.1 Self-similar sets and measures

Let $\Phi = \{\phi_i\}_{i\in\mathcal{I}}$ be a finite set of contractive similitudes on $\mathbb{R}^d$: that is, for each $i \in \mathcal{I}$,

$$\phi_i(x) = r_i O_i x + t_i$$

where $r_i \in (0,1)$, O_i is an orthogonal matrix, and $t_i \in \mathbb{R}^d$. It is well-known that there is a unique compact set Λ satisfying

$$\Lambda = \bigcup_{i \in \mathcal{I}} \phi_i(\Lambda).$$

We refer to this set as the *attractor* of the system, or more generally as a *self-similar set*.

In this document, however, we are more interested in fractal measures. Given a probability vector $\mathbf{p} = (p_i)_{i \in \mathcal{I}}$, we denote the corresponding *self-similar measure* ν, which is the unique Borel probability measure satisfying the invariance relationship

$$\nu = \sum_{i \in \mathcal{I}} p_i \, (S_i)_* \, \nu.$$

Equivalently, ν is the weak limit of the sequence $\nu_k = \sum_{\mathbf{i} \in \mathcal{I}^k} \delta_{\phi_{\underline{i}}(0)}$.

1.2 Bernoulli convolutions

The most important example for us will be a particularly well-studied class of self-similar iterated function systems: the *Bernoulli convolutions* with parameter $\lambda \in (0,1)$. Consider the IFS $\{S_{-1}^\lambda, S_1^\lambda\}$ where $S_t^\lambda(x) = \lambda x + t$ for $t \in \{-1, 1\}$. When $\lambda < 1/2$, the associated self-similar set is the usual Cantor set with subdivision ratio λ, but for $\lambda \geq 1/2$ the attractor is the entire interval.

Definition 1.1 (Bernoulli Convolution 1)**.** We call the associated self-similar measure μ_λ associated with the IFS $\{S_{-1}^\lambda, S_1^\lambda\}$ and the probability vector $(1/2, 1/2)$ the *Bernoulli convolution with parameter* λ.

We can also define the measure μ_λ as the limiting distribution of a random walk with decaying weights λ^k.

Definition 1.2 (Bernoulli Convolution 2)**.** Let X be a Rademacher random variable, i.e. $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/2$. Consider the i.i.d. sequence of random variables $(X_n)_{n=0}^\infty$ with the same distribution as X. Then μ_λ is the law of the random variable

$$\sum_{k=0}^\infty \lambda^k X_k.$$

We also write $Y_n = \sum_{k=0}^{n-1} \lambda^k X_k$.

Perhaps unsurprisingly from the name, the Bernoulli convolution also has a particular convolution structure.

Lemma 1.3. *Let $\lambda \in (0,1)$. Set*

$$\nu_\lambda^n := \frac{1}{2^n} \sum_{\underline{i} \in \{0,1\}^n} \delta_{S_{\underline{i}}(0)}.$$

Then ν_λ^n is the law of the discrete random variable Y_n and

$$\nu_\lambda^n = \mathop{\Large *}_{k=0}^{n-1} \frac{1}{2} \left(\delta_{\lambda^k} + \delta_{-\lambda^k} \right).$$

Moreover, ν_λ is the weak limit of ν_λ^n.

The convolution structure appears since Y_n is a sum of the random variables with law $\frac{1}{2}(\delta_{\lambda^k} + \delta_{-\lambda^k})$. The convolution structure also implies that the Fourier transform of a Bernoulli convolution has a particularly convenient form. First, we can compute directly that

$$\widehat{\nu_\lambda^n}(\xi) = \prod_{k=0}^{n-1} \frac{\widehat{\delta_{\lambda^k}}(\xi) + \widehat{\delta_{-\lambda^k}}(\xi)}{2} = \prod_{k=0}^{n-1} \frac{e^{-2\pi i \xi \lambda^k} + e^{2\pi i \xi \lambda^k}}{2} = \prod_{k=0}^{n-1} \cos(2\pi i \xi \lambda^k). \tag{1.1}$$

Then since ν_λ is the weak limit of ν_λ^n, by Portmanteau's lemma:

$$\widehat{\nu_\lambda}(\xi) = \prod_{k=0}^{\infty} \cos(2\pi i \xi \lambda^k). \tag{1.2}$$

But before we say much more about the Fourier transform of Bernoulli convolutions, we first recall some basic facts from Fourier analysis that will be useful to us.

1.3 The L^1 Fourier transform

Let $\mathcal{M}(\mathbb{R}^d)$ denote the set of finite complex valued measures equipped with the norm

$$\|\mu\| = |\mu|(\mathbb{R}^d)$$

where $|\mu|$ denotes the total variation. We then define the *Fourier transform* of μ by

$$\widehat{\mu}(\xi) = \int_{\mathbb{R}^d} \mathrm{e}^{-2\pi i x \cdot \xi} \, d\mu(x). \tag{1.3}$$

Of course, if $\mu = f \in L^1(\mathbb{R}^d)$, then the Fourier transform is equivalently given by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^d} f(x)\, \mathrm{e}^{-2\pi i x\cdot\xi}\, dx. \tag{1.4}$$

Let us note a few general properties of the Fourier transform.

Lemma 1.4. *Let $\mu \in \mathcal{M}(\mathbb{R}^d)$. Then:*

1. *$\widehat{\mu}(0) = \mu(\mathbb{R}^d)$.*
2. *$\widehat{\mu}$ is bounded and uniformly continuous.*
3. *Suppose μ is has support contained in the ball $B(x, R)$ for some $R \geq 0$. Then μ is Lipschitz with Lipschitz constant $R|\mu|(\mathbb{R}^d)$.*

Proof. A direct computation gives that $\widehat{\mu}(0) = \mu(\mathbb{R}^d)$. To see that the Fourier transform is bounded, for $\xi \in \mathbb{R}^d$,

$$\widehat{\mu}(\xi) \leq \int |e^{-2\pi i x\cdot\xi}|\, \mathrm{d}\mu(x) \leq |\mu|(\mathbb{R}^d).$$

To see uniform continuity,

$$\begin{aligned}\widehat{\mu}(\xi + h) - \widehat{\mu}(\xi) &= \int \left(e^{-2\pi i x\cdot(\xi+h)} - e^{-2\pi i x\cdot\xi}\right) \mathrm{d}\mu(x)\\ &\leq \int \left|e^{-2\pi i x\cdot h} - 1\right| \mathrm{d}|\mu|(x).\end{aligned}$$

The final integral converges to 0 uniformly as h converges to 0, as claimed.

To prove the final fact, we recall the following inequality, which is an immediate consequence of the complex mean value theorem: for all $x, y \in \mathbb{R}$,

$$|e^{ix} - e^{iy}| \leq |x - y|.$$

Therefore, since $|x| \leq R$ for all $x \in \operatorname{supp}\mu$,

$$\begin{aligned}|\widehat{\mu}(\xi) - \widehat{\mu}(\xi')| &\leq \int \left|e^{-2\pi i x\cdot\xi} - e^{-2\pi i x\cdot\xi'}\right| \mathrm{d}\mu(x)\\ &\leq \int 2\pi|x\cdot\xi - x\cdot\xi'|\, \mathrm{d}|\mu|(x)\\ &\leq 2\pi R|\mu|(\mathbb{R}^d)|\xi - \xi'|\end{aligned}$$

as claimed. □

To conclude this section, we note the following type of idempotence of the Fourier transform, which is known as *Fourier inversion*. The proof of this result is somewhat more technical; for a proof, see [5, Theorem 3.4].

Theorem 1.5 (Fourier Inversion)**.** *If $f \in L^1$ and $\widehat{f} \in L^1$ then*

$$f(x) = \widehat{\widehat{f}}(-x) \tag{1.5}$$

1.4 Schwartz space and the L^2 Fourier transform

In the previous section, we defined the Fourier transform essentially for L^1 functions. However, it would also be nice to define a notion of Fourier transform on spaces which are larger than L^1. Unfortunately, the integral Eq. (1.3) defining the Fourier transform does not immediately make sense in more general spaces, such as L^p for $p > 1$.

Instead, the approach in this section is to consider the Fourier transform on a special class of functions which are particularly well-behaved, and then extend the Fourier transform to the larger space by defining it on the class of functions and moreover requiring that the Fourier transform interact in a reasonable way with the topology on the respective spaces.

The nice class of functions that we will consider is the following. Essentially, these are smooth functions whose derivatives decay sufficiently rapidly at infinity and includes, for instance, the class of compactly supported smooth functions.

Definition 1.6. We denote the *Schwartz space* of functions $\mathcal{S}$ as the set of functions $f \in C^\infty$ such that $x^\alpha D^\beta f$ is a bounded function for each α, β.

The Schwartz space is equipped with the family of semi-norms

$$\|f\|_{\alpha\beta} = \|x^\alpha D^\beta f\|_\infty.$$

This family of norms induces the structure of a *Frechet space* on $\mathcal{S}$. However, we will not go into detail about this; let us just note that a sequence (f_k) converges to f in $\mathcal{S}$ if

$$\lim_{k\to\infty} \|f_k - f\|_{\alpha\beta} = 0 \quad \text{for all} \quad \alpha, \beta.$$

There are a number of equivalent characterizations of Schwartz space.

Proposition 1.7. *Let $f \in C^\infty$. Then the following are equivalent.*

- *$f \in \mathcal{S}$.*

- $(1+|x|)^N D^\beta f$ *is bounded for each* N *and* β.
- $\lim_{n\to\infty} x^\alpha D^\beta f = 0$ *for each* α,β.
- $\|x^\alpha D^\beta f\|_1 < \infty$ *for all* α, β.

More generally, Schwartz space has a many very nice properties that makes it convenient to work with.

Proposition 1.8. *The Schwartz space* $\mathcal{S}$ *has the following key properties.*

- *The Fourier transform* $\mathcal{F}$ *restricts to a continuous is continuous and surjective function on* $\mathcal{F}: \mathcal{S} \to \mathcal{S}$.
- *Plancherel theorem* $\mathcal{F}$ *is an* L^2*-isometry on* $\mathcal{S}$*:*

$$\int_{\mathbb{R}^d} f(x)\overline{g(x)}dx = \int_{\mathbb{R}^d} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}d\xi \tag{1.6}$$

Of course, every compactly supported smooth function is a Schwartz function. The Schwartz space also includes a number of other important examples; perhaps the most significant is the *Gaussian*

$$g(x) = e^{-\pi x^2}.$$

The Gaussian is a fixed point of the Fourier transform: $\widehat{g} = g$.

The fact that the class of Schwartz functions is preserved by the Fourier transform is a key motivation for this class, and also why in general it is somewhat more useful than the more restrictive class of compactly supported smooth functions. In fact, if f is compactly supported, then $\widehat{f}$ will *never* be compactly supported.

The Schwartz space having nice properties is quite insufficient; it is also important that Schwartz space contains an abundance of functions.

Fact 1.9. *The space* $\mathcal{S}$ *is dense in* L^p *for all* $1 \le p < \infty$.

Actually, we have the following somewhat better simultaneous approximation property. The proof can be found in [5, Lemma 3.3], which says that we can approximate L^p functions by functions in Schwartz space simultaneously in all p.

Lemma 1.10. *Suppose* $f \in L^1_{loc}$. *Then there is a fixed sequence* $\{g_k\} \subset C_0^\infty$ *such that if* $p \in [1,\infty)$ *and* $f \in L^p$, *then* $g_k \to f$ *in* L^p. *If moreover* f *is continuous and decays to zero at* ∞, *then* $g_k \to f$ *uniformly.*

We can now define the Fourier transform in L^2. In particular, we have the following key corollary (and generalization) of the Plancherel theorem on Schwartz space.

Theorem 1.11. *There is a unique bounded operator $\mathcal{F} : L^2 \to L^2$ such that for $f \in \mathcal{S}$, $\mathcal{F}(f) = \widehat{f}$. Moreover, $\mathcal{F}$ is a unitary operator and $\mathcal{F} f = \widehat{f}$ for all $f \in L^1 \cap L^2$.*

In particular, if $\nu = \mu + f\,\mathrm{d}x$ where $\mu \in \mathcal{M}(\mathbb{R}^d)$ and $f \in L^2$,

$$\int \widehat{\nu}\phi = \int \widehat{\phi}\,\mathrm{d}\nu \text{ for all } \phi \in \mathcal{S}.$$

The proof is given in [5, Theorem 3.10].

1.5 Convolution and approximate identities

In this section, we introduce a useful general technical tool which provides well-behaved approximations of general measures.

Let us begin by recalling the convolution operator. Given functions f, g, we would like to define the convolution

$$f * g(x) = \int_{\mathbb{R}^d} f(x-y)g(y)dy. \tag{1.7}$$

However, we need to make some assumptions concerning the functions to ensure that this integral actually exists.

First, let us observe that this definition is entirely sensible if f and g are L^1 functions.

Definition 1.12. If $\mu, \nu \in M(\mathbb{R}^n)$, we define their convolution $\mu * \nu \in M(\mathbb{R}^n)$ by $\mu * \nu(E) = \mu \times \nu\left(\alpha^{-1}(E)\right)$ where $\alpha : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is addition, $\alpha(x, y) = x + y$. Equivalently,

$$\mu \times \nu(E) = \iint \chi_E(x+y)d\mu(x)d\nu(y).$$

The convolution operation for measures has the following properties.

Proposition 1.13.

1. *Convolution of measures is commutative and associative.*
2. *For any bounded Borel measurable function h,*

$$\int h d(\mu * \nu) = \iint h(x+y)d\mu(x)d\nu(y).$$

3. $\|\mu * \nu\| \leq \|\mu\|\|\nu\|$.

4. *If $d\mu = f\,dm$ and $d\nu = g\,dm$, then $d(\mu * \nu) = (f * g)dm$; that is, on L^1 the new and old definitions of convolution coincide.*

We can also define the convolution for more general functions in L^p for $1 \le p \le \infty$. This fact is known as *Young's inequality*.

Proposition 1.14 (Young's inequality)**.** *If $f \in L^1$ and $g \in L^p (1 \le p \le \infty)$, then $f * g(x)$ exists for almost every x, $f * g \in L^p$, and $\|f * g\|_p \le \|f\|_1 \|g\|_p$.*

The convolution operation has the following smoothing property.

Lemma 1.15. *If $f \in C_0^\infty$ and $g \in L^1_{loc}$ then $f * g$ is C^∞ and*

$$D^\alpha (f * g) = (D^\alpha f) * g \tag{1.8}$$

Remark 1.16. For the random variables X, Y with law μ, ν respectively, the law of $X + Y$ is $\mu * \nu$.

Now, we can introduce the notion of an approximate identity.

Definition 1.17. We say that a family of functions $\{\psi_\varepsilon : \varepsilon > 0\}$ is an *approximate identity* if $\operatorname{supp} \psi_\varepsilon \subset B(0, \varepsilon)$ and $\int \psi_\varepsilon = 1$ for all $\varepsilon > 0$.

A standard way to construct an approximate identity is to begin with a smooth compactly supported function $\phi \in C_c^\infty$ with $\int \phi = 1$ and define

$$\phi_\varepsilon(x) = \varepsilon^{-n} \phi(x/\varepsilon)\} \tag{1.9}$$

We will require the following key property of approximate identities.

Theorem 1.18. *Let $\{\psi_\varepsilon(x)\}$ be an approximate identity and μ a locally finite Borel measure on $\mathbb{R}^n$. Then $\psi_\varepsilon * \mu$ converges weakly to μ as $\varepsilon \to 0$: more precisely,*

$$\int \psi(x)(\psi_\varepsilon * \mu)\,\mathrm{d}x \to \int \psi(x)\,\mathrm{d}\mu(x),$$

for all $\psi \in C_c^0(\mathbb{R}^d)$.

1.6 A few words about tempered distributions

In the previous section, we defined the Fourier transform on L^2. One can also check for a function $f_1 + f_2 \in L^1 + L^2$ that it is well-defined to set

$$\mathcal{F}(f_1 + f_2) = \widehat{f_1} + \widehat{f_2}$$

where these are the L^1 and L^2 Fourier transforms respectively; this definition does not depend on the particular decomposition of the functions. In particular, since every function in L^p for $1 \leq p \leq 2$ can be decomposed as a sum $f_1 + f_2 \in L^1 + L^2$, this gives an unambiguous definition of the Fourier transform on L^p with $1 \leq p \leq 2$.

Unfortunately, things break down for $p > 2$; and then one might also wish to speak sensibly about the Fourier transform of more general functions, such as of infinite (but locally finite) measures, or unbounded functions.

The most standard way to extend the Fourier transform is through the machinery of *tempered distributions*. Recall the definition of the Schwartz space $\mathcal{S}$ from the previous section, equipped with the structure of a Frechet space induced by the family of seminorms $\cdot_{\alpha\beta}$.

Definition 1.19. The space $\mathcal{S}'$ of *tempered distributions* is the linear dual of $\mathcal{S}$; that is, the space of continuous linear functionals from $\mathcal{S}$ to $\mathbb{C}$.

The space of tempered distributions includes a number of objects which we have considered in previous sections. Of course, the forward direction of the Riesz representation theorem implies that measures define tempered distributions by integration:

$$\phi \mapsto \int \phi \, \mathrm{d}\mu.$$

The space of tempered distributions also includes the L^p spaces.

Example 1.20. For any $1 \leq p \leq \infty$ if $f \in L^p$, then f defines a tempered distribution via the action by integration

$$\phi \mapsto \int \phi(x) f(x) \, \mathrm{d}x.$$

Even more generally, since $\mathcal{S}$ consists of functions which are decaying rapidly at infinity, if f is an unbounded function but increasing sufficiently slowly, then it also defines a functional by integration. For instance, any function $f \in L^1_{loc}$ which has

$$\int (1 + |x|)^N |f(x)| \, \mathrm{d}x < \infty$$

for some $N \in \mathbb{N}$ also defines a tempered distribution by integration, in the same way as above.

In particular, the space of tempered distributions even includes the functions which occur as Fourier transforms. For instance, if $\mu \in \mathcal{M}(\mathbb{R}^d)$, recall that $\widehat{\mu}$ is a uniformly continuous and bounded function; in particular, it satisfies the above conditions.

The key usefulness of the space of tempered distributions, again exploiting invariance of the Fourier transform on Schwartz space, is that we can extend the Fourier transform by duality. Given $\Gamma \in \mathcal{S}'$, we define its Fourier transform $\widehat{\Gamma} \in \mathcal{S}'$ by the rule

$$\widehat{\Gamma}(\phi) = \Gamma(\widehat{\phi}).$$

In "integral notation", $f = \widehat{g}$ if for all $\phi \in \mathcal{S}$

$$\int \phi(x) g(x)\,\mathrm{d}x = \int \widehat{\phi}(x) f(x)\,\mathrm{d}x.$$

It takes a decent amount of work to rigorously build up the theory of the Fourier transform on $\mathcal{S}'$, and to check that the Fourier transform behaves as expected on the familiar objects. We do not undertake this development here.

Instead, let us content ourselves with a computation. Let f be the constant 1 function on $\mathbb{R}^d$. To understand $\widehat{f}$, let us see how $\widehat{f}$ acts on Schwartz functions. Indeed, for $\phi \in \mathcal{S}$,

$$\widehat{f}(\phi) = \int \widehat{\phi}(x)\,\mathrm{d}x = \int \widehat{\phi}(x) e^{2\pi i x\cdot 0}\,\mathrm{d}x = \phi(0) = \int \phi(x)\delta_0(x)\,\mathrm{d}x.$$

In other words, $\widehat{f}$ acts as integration against δ_0, so in fact $\widehat{f} = \delta_0$.

2 Fourier decay of fractal measures

In this section, we now are now in the heart of the matter: we begin the study the Fourier decay of fractal measures.

Throughout this section, we specialize to the space $M(\mathbb{R}^d)$ of compactly supported finite Borel measures on $\mathbb{R}^d$.

2.1 An introduction to Fourier decay

The weakest possible decay which one can hope for is eventual convergence to 0 at infinity.

Definition 2.1. We say that a measure μ is *Rajchman* if $|\widehat{\mu}(\xi)|$ converges to 0 as $|\xi|$ diverges to infinity.

A fundamental fact is that measures which are absolutely continuous with respect to Lebesgue measure are Rajchman.

Lemma 2.2 (Riemann–Lebesgue). *Suppose $\mu = f \in L^1$. Then μ is Rajchman.*

The proof can be found, for instance, in [2, Proposition 2.2.17].

However, there are already many measures which are far from being Rajchman.

Example 2.3. Let W be an affine subspace of $\mathbb{R}^d$, and suppose μ is supported on W. Then for all $\xi \in W^\perp$, since $x \cdot \xi = 0$ for all $x \in \operatorname{supp} \mu$,

$$\widehat{\mu}(\xi) = \int e^{-2\pi i x \cdot \xi} \, \mathrm{d}\mu(x) = \mu(\mathbb{R}^d).$$

In other words, $\widehat{\mu}$ has *no Fourier decay at all* on $W^\perp$; the situation is particularly bad when in fact W is a proper affine subspace.

Despite this, being trapped in a subspace does not immediately preclude Fourier decay. A key feature of the following example is *curvature*.

Example 2.4. Let σ^{d-1} denote the surface measure on the unit sphere S^{d-1}. Then there is a constant $C(d) > 0$ so that

$$|\widehat{\sigma^{d-1}}(\xi)| \leq C(d)|\xi|^{(1-d)/2}.$$

See, for instance, [3, Section 3.3].

The previous example provides some additional motivation: we are not only interested in establishing Fourier decay, but also interested in obtaining information on the rate of decay of the Fourier transform.

Definition 2.5. We say that a measure μ has *power Fourier decay* if there is an exponent $s > 0$ so that

$$|\widehat{\mu}(\xi)| \lesssim |\xi|^{-s}.$$

To motivate why one might be interested in power Fourier decay, let us note some consequences.

Theorem 2.6. *Let $\mu \in M(\mathbb{R}^d)$. Then:*

1. *If $\widehat{\mu} \in L^2$, then μ is absolutely continuous with L^2-density.*

2. *If $\widehat{\mu} \in L^1$, then μ is absolutely continuous with continuous density.*

Proof. First suppose $\widehat{\mu} \in L^2$. Since the L^2 Fourier transform is surjective, get $f \in L^2$ so that $\widehat{f} = \widehat{\mu}$. We wish to show that $f = \mu$: to do this, we will use the Riesz representation theorem. More precisely, it suffices to show that for all $\psi \in C_c^0$ that

$$\int \psi(x) f(x) \, \mathrm{d}x = \int \psi(x) \, \mathrm{d}\mu(x).$$

First, if $\phi \in \mathcal{S}$, then since $f \in L^2$, by Theorem 1.11,

$$\int \phi f = \int \widehat{\phi} \widehat{f} = \int \widehat{\phi} \widehat{\mu} = \int \phi \, \mathrm{d}\mu.$$

Therefore the inequality holds on $\mathcal{S}$.

To complete the proof, by Lemma 1.10, let ϕ_n be a sequence in $\mathcal{S}$ converging uniformly and in L^2 to ϕ. Then by the forward direction of the Riesz representation theorem and the Cauchy–Schwarz inequality, the desired inequality holds for all $\phi \in \mathcal{S}$.

Next, suppose $\widehat{\mu} \in L^1$. Let $\{\phi_\varepsilon : \varepsilon > 0\}$ be a smooth approximate identity, so for all $\varepsilon > 0$, μ_ε is smooth and compactly supported. In particular, $\widehat{\mu_\varepsilon} \in \mathcal{S}$. Therefore by Fourier inversion,

$$\mu_\varepsilon(x) = \int \widehat{\mu_\varepsilon} e^{2\pi i x \cdot \xi} \, \mathrm{d}\xi = \int \widehat{\psi_\varepsilon}(\xi) \widehat{\mu}(\xi) e^{2\pi i x \cdot \xi} \, \mathrm{d}\xi.$$

Since $\widehat{\mu} \in L^1$ and $\widehat{\psi_\varepsilon}(\xi) e^{2\pi i x \cdot \xi}$ is uniformly bounded in ε, by the dominated convergence theorem,

$$g(x) := \lim_{\varepsilon \to 0} \mu_\varepsilon = \int \widehat{\mu} e^{2\pi i x \cdot \xi} \, \mathrm{d}\xi.$$

In fact, g is continuous and bounded as argued in Lemma 1.4. On the other hand, we recall that μ_ε converges weakly to μ. It follows that $\mu \in L^1$ with continuous and bounded density g. □

In particular, sufficiently good power decay is sufficient to establish absolute continuity. This is a sort of converse to the Riemann–Lebesgue lemma. Unfortunately, having sufficiently fast decay so that $\widehat{\mu} \in L^1$ is in many situations too much to ask for to be useful.

Corollary 2.7. *Let $\mu \in M(\mathbb{R}^d)$. If $|\widehat{\mu}(\xi)| = O(|\xi|^{-\alpha})$ for some $\alpha > d$, then $\widehat{\mu} \in L^1$.*

2.2 Fourier transform and dimension

In this section, we consider the relationship between Fourier decay and fractal dimension through an intermediate representation known as the *energy* of a measure. The main point is to provide some demonstration of the following phenomenon: *if μ has sufficiently large dimension, then μ must exhibit some form of Fourier decay, at least on average*.

Let us begin with an L^2 notion of dimension.

Definition 2.8. Let $\mu \in M(\mathbb{R}^d)$. Then the *(lower) correlation dimension* or *L^2 dimension* of μ is given by

$$\dim_2 \mu = \liminf_{r \searrow 0} \frac{\log \int \mu(B(x,r)) d\mu(x)}{\log r}.$$

In general, $\dim_2 \mu \leq \dim_{\mathrm{H}} \mu$ for any measure μ, and strict inequality is possible. Heuristically, the correlation dimension is much more sensitive to the local oscillations of the measure than the Hausdorff dimension.

We now define the *s-energy* of a measure μ by

$$I_s \mu = \iint \frac{d\mu(x) d\mu(y)}{|x-y|^s}.$$

The s-energy and the correlation dimension are closely related.

Lemma 2.9. *Let $\mu \in M(\mathbb{R}^d)$ and let $R_0 = \operatorname{diam} \operatorname{supp} \mu$. Then*

$$\dim_2 \mu = \sup \left\{ s \geq 0 : I_s \mu < +\infty \right\}.$$

Proof. Before we proceed with the proof, let us begin with a useful identity. Fix $t > 0$. By change of variables, for all x,

$$\begin{aligned} \int |x-y|^{-t} d\mu(y) &= \int_0^\infty \mu(\{y : |x-y|^{-t} \geq r\}) \, \mathrm{d}r \\ &= \int_0^\infty \mu(\{B(x, r^{-1/t})\}) \, \mathrm{d}r \\ &= \int_0^\infty t \mu(B(x,r)) r^{-(t+1)} \, \mathrm{d}r. \end{aligned}$$

Now, we first show that $\dim_2 \mu \leq \sup \left\{ s \geq 0 : I_s \mu < +\infty \right\}$. Let $t < s < \dim_2 \mu$. Let r_0 be sufficiently small so that for all $0 < r \leq r_0$, $r^s > \int \mu(B(x,r)) \, \mathrm{d}\mu(x)$. Using the identity proven above, we split the integral into small and large scales. For the small scales, by Fubini,

$$\begin{aligned}\int \int_0^{r_0} t\mu(B(x,r))r^{-(t+1)}\,\mathrm{d}r\,\mathrm{d}x &= \int_0^{r_0} tr^{-(t+1)} \int \mu(B(x,r))\,\mathrm{d}x\,\mathrm{d}r \\ &\leq \int_0^{r_0} tr^{-(t+1)}r^s < \infty\end{aligned}$$

since $t < s$. For the large scales, since μ is compactly supported, applying the trivial bound $\mu(B(x,r)) \leq \mu(\mathbb{R}^d)$,

$$\begin{aligned}\int \int_{r_0}^{\infty} t\mu(B(x,r))r^{-(t+1)}\,\mathrm{d}r\,\mathrm{d}x &= \int \int_{r_0}^{R_0} t\mu(B(x,r))r^{-(t+1)}\,\mathrm{d}r\,\mathrm{d}x \\ &\leq \mu(\mathbb{R}^d) \int \int_{r_0}^{R_0} tr^{-(t+1)}\,\mathrm{d}r\,\mathrm{d}x < \infty.\end{aligned}$$

Combining these two bounds and applying the initial identity yields the claim.

For the converse direction, we prove something somewhat stringer: if $I_s(\mu) < \infty$ then the restriction of the measure μ to a well chosen subset A is s-Frostman. It is clear that this gives the desired bound on $\dim_2 \mu$.

Thus let s be such that $\iint |x-y|^{-s}\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) < \infty$. Let $M > 0$ be sufficiently large so that

$$A := \left\{x : \int |x-y|^{-s} d\mu(y) \leq M\right\}$$

has $\mu(A) > 0$. Then for $x \in A$ is fixed,

$$M \geq \int_{B(x,r)} |x-y|^{-s} d\mu \geq r^{-s}\mu(B(x,r)).$$

In the second inequality, we have used that $|x-y| \leq r$ so that $|x-y|^{-s} \geq r^{-s}$. ◻

Remark 2.10. By Frostman's lemma, we have established the following:

$$\dim_{\mathrm{H}} A = \sup\{s : \text{there exists } \mu \in M(A) \text{ with } I_s(\mu) < \infty\}. \tag{2.1}$$

The energy $I_s\mu$ is sometimes written in terms of the Riesz kernel $k_s(x) = |x|^{-s}$. Namely,

$$I_s(\mu) = \iint |x-y|^{-s} d\mu(y) d\mu(x) = \int k_s * \mu(x) d\mu(x). \tag{2.2}$$

If $d/2 < s < d$, then $k_s \in L^1 + L^2$. However, this is no longer the case for $s \leq d/2$, so to make formal sense of the Fourier transform of k_s one must use the machinery of tempered distributions as introduced in Section 1.6. The following is [3, Theorem 3.6].

Lemma 2.11. *Let $0 < s < d$. Then there exists a positive and finite constant $\gamma(d,s)$ such that $\widehat{k_s} = \gamma(d,s)k_{d-s}$ as a tempered distribution.*

Using this identity, one can establish the following alternative formula for the energy.

Theorem 2.12. *Let $\mu \in M(\mathbb{R}^d)$ and $0 < s < d$. Then*

$$I_s(\mu) = \gamma(d,s) \int |\xi|^{s-1} |\widehat{\mu}(\xi)|^2 \, \mathrm{d}\xi.$$

One can provide a highly questionable justification of this formula as follows:

$$I_s(\mu) = \int k_s * \mu d\mu = \int \widehat{k_s * \mu}\widehat{\overline{\mu}} = \int \widehat{k_s} |\widehat{\mu}|^2 = \gamma(n,s) \int |\widehat{\mu}(x)|^2 |x|^{s-n} dx. \tag{2.3}$$

The main difficulty is that the Riesz kernel only has a Fourier transform in the distributional sense, so all of these operations must be made carefully rigorous. The precise details can be found in [3, Theorem 3.10].

This representation of the energy indicates that $|\widehat{\mu}(x)|$ must decay a certain amount "on average". In particular, if $\mu \in M(\mathbb{R}^d)$ and $I_s(\mu) < \infty$, then

$$|\widehat{\mu}(\xi)| \leq |\xi|^{-s/2}$$

for 'most' ξ with large norm. This can be quantified somewhat: one can show that the set

$$\left\{\xi \in \mathbb{R}^d : |\xi| \leq R \text{ and } |\widehat{\mu}(\xi)| \geq |\xi|^{-s/2}\right\}$$

can be covered by $o(R^d)$ balls of radius 1. In some sense, the set of ξ where $\widehat{\mu}$ does not decay like $|\xi|^{-s/2}$ has measure 0.

This numerology motivates the following definition.

Definition 2.13. We then define the *Fourier dimension* of μ by

$$\dim_{\mathrm{F}} \mu = \sup\{s \leq d : |\widehat{\mu}(\xi)| \lesssim |\xi|^{-s/2} \text{ for all } \xi \in \mathbb{R}^d\}.$$

Recalling Lemma 2.9, it always holds that $\dim_{\mathrm{F}} \mu \leq \dim_2 \mu$, and equality in particular implies that the measure μ enjoys optimal decay properties relative to its dimension. As we recall, for instance, from measures supported on hyperplanes, it can easily happen that $\dim_{\mathrm{F}} \mu$ is much smaller than $\dim_2 \mu$.

2.3 Power Fourier decay and absolute continuity

To conclude this introduction on power Fourier decay, we show that some Fourier decay combined with dimension sufficiently close to d and convolution structure is sufficient to establish absolute continuity. This is, for instance, [4, Lemma 2.1].

Theorem 2.14. *Let $\mu, \nu \in M(\mathbb{R}^d)$ and suppose $\dim_{\mathrm{F}} \nu = s > 0$.*

1. *If $\dim_2 \mu > d - s$, then $\nu * \mu$ is absolutely continuous with density in L^2.*
2. *If $\underline{\dim}_{\mathrm{H}} \mu > d - s$, then $\nu * \mu$ is absolutely continuous.*

Proof. We first see Item 1. By Theorem 2.6 it suffices to prove that

$$\int |\widehat{\mu}(\xi)|^2 |\widehat{\nu}(\xi)|^2 d\xi = \int |\widehat{\mu * \nu}(\xi)|^2 d\xi < \infty.$$

Since $\dim_{\mathrm{F}} \nu = s$, there is a constant $C > 0$ so that for all $\xi \in \mathbb{R}^d$,

$$|\widehat{\nu}(\xi)| \leq C|\xi|^{-s/2}.$$

Moreover, since $\dim_2 \mu > d - s$, taking $t = d - s$, by Lemma 2.9 and Theorem 2.12,

$$\int |\xi|^{t-d} |\widehat{\mu}(\xi)|^2 d\xi < \infty.$$

Combining these two inequalities, it follows that

$$\int |\widehat{\mu}(\xi)|^2 |\widehat{\nu}(\xi)|^2 d\xi C \leq \int |\widehat{\mu}(\xi)|^2 |\xi|^{-s} d\xi < \infty$$

as required.

Now we prove Item 2. Recalling the definition of Hausdorff dimension and the assumption, get $\delta > 0$ so that

$$\underline{\dim}_{\mathrm{H}}(\mu) = \operatorname{essinf}_{x \sim \mu} \liminf_{r \to 0} \frac{\log(\mu(B(x,r)))}{\log r} > d - s + \delta,$$

and therefore for μ almost every $x \in \operatorname{supp} \mu$,

$$\liminf_{r \to 0} \frac{\log(\mu(B(x,r)))}{\log r} > d - s + \delta.$$

Therefore, by Egorov's theorem, for all $\varepsilon > 0$ there exists a set A_ε such that $\mu(A_\varepsilon^c) < \varepsilon$ and a constant $C_\varepsilon > 0$ such that, defining $\mu_\varepsilon = \mu|_{A_\varepsilon}/\mu(A_\varepsilon)$,

$$\mu_\varepsilon(B(x,r)) \leq C_\varepsilon r^{d-s+\delta}.$$

In particular, $\dim_2 \mu_\varepsilon \geq d - s + \delta > d - s$. Therefore applying Item 1, $\nu * \mu|_{A_\varepsilon}$ is absolutely continuous for all $\varepsilon > 0$.

To conclude the proof, suppose E has Lebesgue measure 0. Then

$$\mu * \nu(E) = \mu|_{A_\varepsilon} * \nu(E) + \mu|_{A_\varepsilon^c} * \nu(E) \leq \varepsilon$$

since $\mu|_{A_\varepsilon^c} * \nu$ is absolutely continuous with respect to Lebesgue measure and $\mu(A_\varepsilon^c) \leq \varepsilon$. Since $\varepsilon > 0$ was arbitrary, $\mu * \nu(E) = 0$ so by the Radon–Nikodym theorem, $\mu * \nu$ is absolutely continuous. □

References

[1] Gerald B. Folland. *Real analysis. Modern techniques and their applications.* Pure Appl. Math., Wiley-Intersci. Ser. Texts Monogr. Tracts. New York, NY: Wiley, 2nd ed. edition, 1999.

[2] Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Grad. Texts Math.* New York, NY: Springer, 3rd ed. edition, 2014.

[3] Pertti Mattila. *Fourier analysis and Hausdorff dimension*, volume 150 of *Camb. Stud. Adv. Math.* Cambridge: Cambridge University Press, 2015.

[4] Pablo Shmerkin. On the exceptional set for absolute continuity of Bernoulli convolutions. *Geom. Funct. Anal.*, 24(3):946–958, 2014.

[5] Thomas H. Wolff. *Lectures on harmonic analysis. Edited by Izabella Łaba and Carol Shubin*, volume 29 of *Univ. Lect. Ser.* Providence, RI: American Mathematical Society (AMS), 2003.