

Structure of distributions generated by the scenery flow

Antti Käenmäki, Tuomas Sahlsten and Pablo Shmerkin

ABSTRACT

We expand the ergodic theory developed by Furstenberg and Hochman on dynamical systems that are obtained from magnifications of measures. We prove that any fractal distribution (FD) in the sense of Hochman is generated by a USM, which provides a converse to a regularity theorem on the structure of distributions generated by the scenery flow. We further show that the collection of FDs is closed under the weak topology and, moreover, is a Poulsen simplex, that is, extremal points are dense. We apply these to show that a Baire generic measure is as far as possible from being uniformly scaling: at almost all points, it has all FDs as tangent distributions.

1. Introduction

1.1. Historical background

A central theme in analysis over the years has been the study of ‘tangents’ of possibly complicated objects, in order to take advantage of the regularity arising in the limiting structures through the metamorphosis of magnification. For example, a differentiable function looks locally like an affine map, which is more regular than, and gives information about, the original function. In [21], Preiss introduced the more general notion of tangent measure and employed it to solve some outstanding open problems in the theory of rectifiability. Tangent measures are useful because, again, they are more regular than the original measure (for example, tangent measures of rectifiable measures are flat) but one can still pass from information about the tangent measure to the original measure. As another example of the general idea, for certain non-conformal repellers the tangent sets and measures turn out to have a regular product structure which is absent in the more complicated original object; see [1, 7]. The process of taking blow-ups of a measure or a set around a point in fact induces a natural dynamical system consisting in ‘zooming in’ around the point. This opens a door to ergodic-theoretic methods, which were pioneered by Furstenberg [8] and then in more developed form in [9], with a comprehensive theory developed by Hochman [11].

It turns out that for some geometric problems, notably those involving some notion of dimension, the ‘correct’ class of tangent objects to consider is not tangent measures, but the empirical distributions that appear by magnifying around a typical point. That is, the tangent objects are *measures on measures*, which we call *tangent distributions* (precise definitions are given in Section 2). The reason for this is that tangent measures are defined as weak limits of magnifications around a point, but the sequence along which a tangent measure arises can be very sparse, and for many problems only the behavior on a positive proportion of scales is

Received 20 March 2014; revised 17 September 2014.

2010 *Mathematics Subject Classification* 37A10, 28A80 (primary); 28A33, 28A75 (secondary).

Tuomas Sahlsten acknowledges the support from the University of Bristol, the Finnish Centre of Excellence in Analysis and Dynamics Research, the Emil Aaltonen Foundation and the European Union (ERC grant #306494). Pablo Shmerkin was partially supported by a Leverhulme Early Career Fellowship and by Project PICT 2011-0436 (ANPCyT).

significant. Tangent distributions are supported on tangent measures that reflect precisely the structure of the original measure on a positive density set of scales.

Furstenberg's key innovation was the introduction of a Markov process on the b -adic scaling sceneries of a measure, which he called a *conditional probability (CP) chain*. Since then, CP chains have proved to be a key tool several important problems in fractal geometry, probability theory and ergodic theory. In [9], Furstenberg applied this technology to understand dimension conservation of homogeneous measures. Then Hochman and Shmerkin used it to study projections of fractal measures [13] and the behavior of measures with respect to normal numbers [14]. Furthermore, recently Orponen [20], and Ferguson, Fraser and Sahlsten [7] found connections to the distance set conjecture for several dynamically defined sets. However, CP chains are defined in a discrete dyadic (or b -adic) fashion, and as a result the point that is being zoomed in upon is not 'in the center of the frame' which is often a disadvantage.

An alternative approach is to consider *scenery flows*, in which the magnification is carried out continuously with the point in the center of the frame. Scenery flows have been studied (sometimes with this name and sometimes under different names) by many authors, both for special classes of sets and measures, and in general. We refer the reader to [11] for a historical discussion and references. Mörters and Preiss [18] proved the surprising fact that, when dealing with Ahlfors regular measures, the tangent distributions are Palm distributions, which are distributions with a strong degree of symmetry and translation invariance. Hochman [11] then showed that a similar phenomenon holds for all Radon measures: he proved that tangent distributions for any measure are almost everywhere quasi-Palm distributions, which is a weaker notion than Palm but still represents a strong spatial invariance. Hochman named distributions which are scale-invariant and enjoy the quasi-Palm property as *fractal distributions* (FDs). He also proved the remarkable fact that distributions of CP chains give rise to FDs in a natural way and, reciprocally, any FD can be obtained from the distribution of a CP chain. The main definitions and results from [11] are recalled below, in Section 2.

1.2. Summary of main results

In this work, we continue developing the theory of CP chains and FDs. We state our main results in somewhat informal fashion; precise definitions and statements are postponed to the later sections.

Since FDs are the cornerstone of the theory developed by Hochman [11], a natural problem is to study the topological structure of the family of fractal distributions.

THEOREM A (see Theorem 3.1). *The family of FDs is closed with respect to the weak topology.*

At first sight this may appear rather surprising, since the scaling flow is not continuous, its support is not closed and, more significantly, the quasi-Palm property is not a closed property. Thus, this theorem is another manifestation of the general principle that, although FDs are defined in terms of seemingly strong and discontinuous geometric properties, they are in fact very robust. Besides its intrinsic interest, Theorem A has a number of applications in classical problems in the realm of geometric measure theory, which we develop elsewhere; see [15]. Indeed, these applications were our initial motivation to continue developing the general theory of FDs.

Recall that a *Choquet simplex* Δ in a locally convex topological vector space is a compact convex set with the property that each $x \in \Delta$ can be expressed in a unique way (up to measure zero sets) as an integral $\int y dP(y)$ for some probability distribution P on the extremal points of Δ . It follows from Theorem A and results in [11] that the family of FDs is in fact a Choquet simplex, so another question arises: what kind of Choquet simplex is it?

THEOREM B (see Theorem 4.1). *The family of FDs is a Poulsen simplex.*

A *Poulsen simplex* is a non-trivial Choquet simplex in which extremal points are dense. A classical result of Lindenstrauss, Olsen and Sternfeld [16] states that there is in fact a unique Poulsen simplex up to affine homeomorphism. In that paper, two other striking properties of the Poulsen simplex are established: any affine homeomorphism between two proper faces of the Poulsen simplex may be extended to a homeomorphism of the whole simplex (homogeneity), and any metrizable simplex is affinely homeomorphic to a face of the Poulsen simplex. The Poulsen simplex is a common object in ergodic theory, as the space of invariant measures for many dynamical systems is Poulsen; this is often a manifestation of some kind of hyperbolic behavior.

In our case, the set of extremal points is precisely the collection of ergodic fractal distributions (EFDs) with respect to the scenery flow. We remark that the dense set we exhibit consists of distributions of random self-similar measures, where the self-similarity is with respect to a b -adic grid. This potentially allows us to prove certain statements for arbitrary measures or distributions by reducing it to this fairly concrete and well-behaved class. The construction of these self-similar measures is a special case of what we term the *splicing* of scales. Roughly speaking, this consists in pasting together a sequence of measures along dyadic scales; see Subsection 4.1 for more details. Splicing is often employed to construct sets or measures with a given property based on properties of the component measures. For example, in [23], splicing was used to investigate the dimensions of iterated sums of a Cantor set, and Hochman [11, Section 8.3] employed it (under the name of *discretization*) to construct examples of uniformly scaling measures (USMs) with non-ergodic limit geometry and bad projection properties.

Given a measure μ , we can study its geometric properties via its tangent distributions. The situation is especially nice when at μ almost all points there is a single tangent distribution, and even nicer when all these tangent distributions coincide. This leads us to the concepts of USMs and *generated* distributions; these concepts were first defined by Gavish [10] and investigated further by Hochman [11]; see Subsection 2.2 for more details.

USMs are geometrically much more regular than arbitrary measures, for example in the behavior of their projections [11, 13] and the distance sets of their supports [7]. Examples of USMs are many conformal and non-conformal constructions, both deterministic and random [7, 10, 11], measures invariant under $x \rightarrow px \bmod 1$ on the circle [12] and the occupation measure of Brownian motion in dimension $d \geq 3$ [10].

It seems natural to ask what kind of distributions can arise as the (unique) distribution generated by a USM. Hochman [11] proved the striking fact that generated distributions are always FDs. We provide a converse to this.

THEOREM C (see Theorem 5.1). *Every FD is generated by some uniformly scaling measure.*

Again, our motivation for this result arose from our applications to problems in geometric measure theory; see [15]. Roughly speaking, our approach there is to study families of measures (for example, measures satisfying a certain porosity condition) through the family of tangent distributions to those measures at typical points. A key last step is then to pass from the information gleaned on the FD side back to information about measures; this is where Theorem C comes in.

Recall that a property is *Baire generic* if it is satisfied everywhere except possibly in a set of first category, that is, a countable union of sets whose closure has empty interior. A recurrent topic in geometric measure theory and analysis is the behavior of Baire generic objects, such as sets, measures or functions. For example, in recent years many authors have explored the fractal and multifractal behavior of generic Borel measures; see, for example, [2, 3, 5, 22] and references therein. In this context, it seems very natural to study the tangent structures of

generic measures. O’Neil [19] and Sahlsten [22] proved that a Baire generic measure has *all* Borel measures as tangent measures at almost every point. Even though Theorem C concerns measures which have a single tangent distribution at typical points, perhaps surprisingly, it gives us an application which shows that the exact opposite holds for a generic measure.

THEOREM D (see Theorem 6.1). *For a Baire generic Radon measure μ on \mathbb{R}^d , the set of tangent distributions is the set of all FDs at μ almost every x .*

We also obtain an analogous result for CP distributions (CPDs); see Proposition 6.2. These results are in some sense expected, since there is a heuristic principle that says that Baire generic objects behave ‘as wildly as possible’.

The rest of the paper is organized as follows. In Section 2, we recall the main elements and results of Hochman’s theory. Precise versions of Theorems A–D are stated and proved in Sections 3–6, respectively. In the appendix, we discuss the independence of our results from the chosen norm.

2. Scenery flow, FDs and CPDs

In this section, we recall the main definitions and results from Hochman’s work [11], and provide some minor extensions to the theory developed there. We use much of Hochman’s notation but we also introduce some new terms, such as tangent distribution and micromasure distribution.

NOTATION 2.1. Equip \mathbb{R}^d with the norm $\|x\| = \max_i |x_i|$ and the induced metric. The closed ball centered at x with radius $r > 0$ is denoted by $B(x, r)$. In particular, we write $B_1 = B(0, 1) = [-1, 1]^d$. Given a metric space X , we denote the family of all Borel probability measures on X by $\mathcal{P}(X)$, and the family of all Radon measures on X by $\mathcal{M}(X)$. When X is locally compact, $\mathcal{M}(X)$ and $\mathcal{P}(X)$ are endowed with the weak topology. Recall that $\mu_n \rightarrow \mu$ weakly if $\int f d\mu_n \rightarrow \int f d\mu$ for all continuous functions $f: X \rightarrow \mathbb{R}$ of compact support.

Whenever we consider convergence in a space of probability measures, it is implicitly understood that we are considering the weak convergence. When $X = \mathbb{R}^d$, we write $\mathcal{M} = \mathcal{M}(\mathbb{R}^d)$. The space \mathcal{M} is metrizable, complete and separable. If X is compact, then also $\mathcal{P}(X)$ is compact.

Following terminology from [11], we refer to elements of $\mathcal{P}([-1, 1]^d)$ or \mathcal{M} as *measures*, and to elements of $\mathcal{P}(\mathcal{P}([-1, 1]^d))$ and $\mathcal{P}(\mathcal{M})$ as *distributions*. Measures are denoted by lowercase Greek letters μ, ν , etc. and distributions by capital letters P, Q , etc. We use the notation $x \sim \mu$ if a point x is chosen randomly according to a measure μ . Moreover, write $\mu \sim \nu$ if the measures μ and ν are *equivalent*, that is, they have the same null-sets. If X and Y are metric spaces, $\mu \in \mathcal{P}(X)$ and $f: X \rightarrow Y$ is a Borel map, then the *push-down* $f\mu$ is the measure defined via $f\mu(A) = \mu(f^{-1}A)$.

2.1. Ergodic-theoretic preliminaries

In this article, we make use of many standard definitions and facts from ergodic theory which we briefly recall here for the reader’s convenience. Good general references are the books of Einsiedler and Ward [6] and Walters [24].

Let (X, \mathcal{B}, μ) be a probability space. We say that a transformation $T: X \rightarrow X$ preserves μ if it is \mathcal{B} -measurable and $T\mu = \mu$; the set of all such transformations is a semigroup under composition. A *measure-preserving system* (m.p.s.) is a tuple (X, \mathcal{B}, μ, T) , where (X, \mathcal{B}, μ) is a probability space and T is an action of a semigroup by transformations that preserve μ . That is, there is a semigroup S and for each $s \in S$ there is a map $T_s: X \rightarrow X$ that preserves μ ,

such that $T_{s+s'} = T_s \circ T_{s'}$. In this article, the underlying semigroup is always one of \mathbb{N} , \mathbb{Z} (in which case, we speak of measure-preserving maps, since the action is determined by $T := T_1$) or \mathbb{R}^+ , \mathbb{R} (in which case, we speak of flows). Moreover, for us X is always a metric space and \mathcal{B} is the Borel σ -algebra on X (thus no explicit reference is made to it). In the following, we always assume that we are in this setting to avoid unnecessarily technical assumptions.

An m.p.s. is *ergodic* if any set $A \in \mathcal{B}$ with $\mu(T_s^{-1}A \Delta A) = 0$ for all s has either zero or full μ -measure. For a given action T on a space X , ergodic measures are the extremal points of the convex set of all probability measures which are preserved by T . The *ergodic theorem* for discrete actions says that if $f \in L^1(\mu)$ and the system is ergodic, then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int f d\mu \quad \text{for } \mu \text{ almost all } x.$$

For flows, the same holds replacing the left-hand side by $\lim_{t \rightarrow \infty} (1/t) \int_0^t f \circ T_s ds$.

Given two measure-preserving systems (X, μ, T) and (X', μ', T') , a map $\pi: X \rightarrow X'$ is called a *factor map* if the underlying semigroups coincide, $\pi\mu = \mu'$ and π intertwines the actions of the semigroups: $\pi T_s = T'_s \pi$ for all s . In this case, we also say that (X', μ', T') is a factor of (X, μ, T) . The factor of an ergodic system is ergodic. When π is a measure-theoretical isomorphism, we say that the systems (X, μ, T) and (X', μ', T') are *isomorphic*.

Let (X, μ) be a metric probability space, and consider the product spaces $(X^{\mathbb{N}}, \mu^{\mathbb{N}})$ and $(X^{\mathbb{Z}}, \mu^{\mathbb{Z}})$. The shift map T defined by $T((x_i)_i) = (x_{i+1})_i$ acts on both spaces. This map preserves the product measure and the resulting system is always ergodic. In general, there may be many other measures on $X^{\mathbb{N}}$ or $X^{\mathbb{Z}}$ that are preserved by the shift. If $(X^{\mathbb{N}}, \mu, T)$ is an m.p.s., then there is always an m.p.s. $(X^{\mathbb{Z}}, \hat{\mu}, T)$ such that the former is a factor of the latter under the natural projection map (on the natural extension the σ -algebra is not the Borel σ -algebra but the smallest σ -algebra that makes the projection $x \mapsto x_0$ measurable, but this exception to our convention causes no trouble). This is called the *two-sided extension* or natural extension for shift spaces. The two-sided extension is ergodic if and only if the one-sided version is ergodic. Furthermore, any discrete m.p.s. (X, μ, T) can always be represented as a shift space via the identification $x \rightarrow (T^i x)_{i=0}^{\infty}$ (that is, the measure on $X^{\mathbb{N}}$ is the push-down of μ under this map).

A general m.p.s. (X, \mathcal{B}, μ, T) can be decomposed into (possibly uncountably many) ergodic parts according to the *ergodic decomposition* theorem: there exists a Borel map $x \rightarrow \mu_x$ from X to $\mathcal{P}(X)$ such that for μ almost all x it holds that each $(X, \mathcal{B}, \mu_x, T)$ is measure-preserving and ergodic, and $\mu = \int \mu_x d\mu(x)$. Moreover, this map is unique up to sets of zero μ -measure. The measures μ_x are called the *ergodic components* of μ .

A standard way to build measure-preserving flows from discrete systems is via *suspensions*; we only consider the case of a constant roof function. Let (X, μ, T) be a discrete m.p.s. (the base) and let $r > 0$ (the height). Write $\hat{X} = X \times [0, r)$, $\hat{\mu} = \mu \times \mathcal{L}$, where \mathcal{L} is normalized Lebesgue measure on $[0, r)$, and set $\hat{T}_t(x, s) = (x, s + t)$ if $s + t < r$ and $\hat{T}_t(x, r - t) = (Tx, 0)$. By iterating, this defines a flow (called *suspension flow*) for all $t > 0$, which indeed preserves the measure $\hat{\mu}$. Moreover, the m.p.s. (X, μ, T) is ergodic if and only if $(\hat{X}, \hat{\mu}, \hat{T})$ is ergodic.

2.2. Normalizations and the scenery flow

If $\mu \in \mathcal{M}$ and $\mu(A) > 0$, then $\mu|_A$ is the restriction of μ to A and, provided also $\mu(A) < \infty$, we denote by μ_A the restriction normalized to be a probability measure, that is

$$\mu_A(B) = \frac{1}{\mu(A)} \mu(A \cap B).$$

For any $\mu \in \mathcal{M}$ for which $\mu(B_1) > 0$ we define the normalization operations $*$, \square in \mathcal{M} by

$$\begin{aligned}\mu^* &:= \frac{1}{\mu(B_1)}\mu, \\ \mu^\square &:= \mu|_{B_1} = \mu^*|_{B_1}.\end{aligned}$$

We define the translation and scaling actions on measures by

$$\begin{aligned}T_x\mu(A) &= \mu(A - x), \\ S_t\mu(A) &= \mu(e^{-t}A).\end{aligned}$$

The reason for the exponential scaling is to make S_t into a partial action of \mathbb{R} into \mathcal{M} . Whenever R is an operator on \mathcal{M} , we write R^* , R^\square for the corresponding operator obtained by postcomposition with the respective normalizations. So, for example, $T_x^\square\mu = (T_x\mu)^\square$. We also write

$$\mathcal{M}^* = \{\mu \in \mathcal{M} : 0 \in \text{spt } \mu\}$$

and $\mathcal{M}^\square = \mathcal{P}(B_1)$.

We note that the actions S_t^* and S_t^\square are discontinuous and fully defined only on the (Borel but not closed) set \mathcal{M}^* . Nevertheless, the philosophy behind many of the results in [11] is that in practice they behave in a very similar way to a continuous action on a complete metric space (compact in the case of S_t^\square).

DEFINITION 2.2 (Scenery flow and tangent measures). We call the flow S_t^\square acting on \mathcal{M}^* the *scenery flow* at 0. Given $\mu \in \mathcal{M}$ and $x \in \text{spt } \mu$, we consider the one-parameter family

$$\mu_{x,t} := \mu_{x,t}^\square = S_t^\square(T_x\mu)$$

generated by the action of S_t^\square on $T_x\mu$ and call it the *scenery of μ at x* . Accumulation points of this scenery are called *tangent measures* of μ at x , and the family of tangent measures of μ at x is denoted by $\text{Tan}(\mu, x)$.

REMARK 2.3. We deviate slightly from the usual definition of tangent measures, which corresponds to taking weak limits of $S_t^*(T_x\mu)$ instead, that is, without restricting the measures.

One of the main ideas of [11], which we further pursue in [15], is that, as far as certain properties of a measure are concerned (including their dimensions), the ‘correct’ tangent structure to consider is not a single limit of μ_{x,t_k} along some subsequence, but the whole statistics of the scenery $\mu_{x,t}$ as $t \rightarrow \infty$.

DEFINITION 2.4 (Scenery and tangent distributions). The *scenery distribution* of μ up to time T at x is defined by

$$\langle \mu \rangle_{x,T} := \frac{1}{T} \int_0^T \delta_{\mu_{x,t}} dt$$

for all $0 \leq T < \infty$. Any weak limit of $\langle \mu \rangle_{x,T}$ in $\mathcal{P}(\mathcal{M}^\square)$ for $T \rightarrow \infty$ is called a *tangent distribution* of μ at x . The family of tangent distributions of μ at x is denoted $\mathcal{TD}(\mu, x)$.

Here the integration makes sense since we are on a convex subset of a topological linear space. Since the space of distributions $\mathcal{P}(\mathcal{M}^\square)$ is compact, $\mathcal{TD}(\mu, x)$ is always a non-empty compact set at $x \in \text{spt } \mu$. Note also that every $P \in \mathcal{TD}(\mu, x)$ is supported on $\text{Tan}(\mu, x)$.

DEFINITION 2.5 (Generated distributions and USMs). We say that a measure μ *generates* a distribution $P \in \mathcal{P}(\mathcal{M}^\square)$ at x if

$$\mathcal{TD}(\mu, x) = \{P\}.$$

Furthermore, μ generates P if it generates P at μ almost every point. In this case, we say that μ is a *uniformly scaling measure* (USM).

If μ is a USM, then, intuitively, it means that the collection of views $\mu_{x,t}$ have well-defined statistics as we zoom-in upon smaller and smaller neighborhoods of x .

2.3. Fractal distributions

It has been observed in various forms that tangent measures and distributions have some kind of additional spatial invariance (the simplest form of this is perhaps the well-known fact that ‘tangent measures to tangent measures are tangent measures’; see [17, Theorem 14.16]). A very sharp and powerful formulation of this principle was obtained in [11]. In order to state it, we need some additional definitions.

DEFINITION 2.6 (Fractal distributions). Let $P \in \mathcal{P}(\mathcal{M})$. We say that the distribution P has the following properties.

(i) The distribution P is *scale-invariant* if it is supported on \mathcal{M}^* , and is invariant under the action of the semigroup S_t^* , that is,

$$P((S_t^*)^{-1}\mathcal{A}) = P(\mathcal{A})$$

for all Borel sets $\mathcal{A} \subset \mathcal{P}(\mathcal{M}^*)$ and all $t > 0$.

(ii) The distribution P is *quasi-Palm* if a Borel set $\mathcal{A} \subset \mathcal{M}$ satisfies $P(\mathcal{A}) = 1$ if and only if P almost every ν satisfies $T_z^*\nu \in \mathcal{A}$ for ν almost every z .

(iii) The distribution P is a *fractal distribution* (FD) if it is scale-invariant and quasi-Palm.

(iv) The distribution P is an *ergodic fractal distribution* (EFD) if it is an FD and it is ergodic under the action of S_t^* .

Write \mathcal{FD} and \mathcal{EFD} for the family of all FDs and EFDs, respectively.

REMARK 2.7. (i) Hochman [11] used an alternative definition for the quasi-Palm property. The requirement was that any bounded open set U containing the origin satisfies

$$P \sim \iint_U \delta_{T_x^*\mu} d\mu(x) dP(\mu).$$

Both definitions are easily seen to agree, since two distributions are equivalent if and only if they have the same sets of full measure.

(ii) A distribution P is called *Palm* if for any open set U containing the origin, we have

$$P = \iint_U \delta_{T_x\mu} d\mu(x) dP(\mu)$$

with finite intensity,

$$\int \mu(B_1) dP(\mu) < \infty.$$

Note that in the above definitions P is a distribution on $\mathcal{P}(\mathcal{M})$, that is, on measures with unbounded support. Most of the time we need to deal with distributions supported on $\mathcal{P}(B_1)$ instead (the main advantage being that this is a compact metrizable space).

NOTATION 2.8 (Restricted distributions). Given P , we write P^\square as the push-down of P under $\mu \rightarrow \mu^\square$. Slightly abusing notation, whenever P is an FD/EFD, we also refer to P^\square as an FD/EFD. Note that in this case P^\square is S^\square -invariant, but the quasi-Palm is not properly defined

for P^\square . When we want to emphasize whether we are talking about P or P^\square , we call the former the *extended version* of P and the latter the *restricted version* of P .

This abuse of notation is justified by the following result; see [11, Lemma 3.1].

LEMMA 2.9. *The action $P \rightarrow P^\square$ induces a 1–1 correspondence between S^* -invariant and S^\square -invariant distributions.*

We can now state the key result of Hochman [11, Theorem 1.7] asserting the additional spatial invariance enjoyed by typical tangent distributions.

THEOREM 2.10. *For any $\mu \in \mathcal{M}$ and μ almost every x , all tangent distributions of μ at x are FDs.*

Given an FD P , as an invariant measure for dynamical systems defined by the flow S_t^* we can consider the *ergodic decomposition* $\{P_\alpha\}$ of P with respect to S_t^* :

$$P = \int P_\alpha dP(\alpha).$$

Hochman [11, Theorem 1.3] also proved that the quasi-Palm property is also preserved when passing to the ergodic components.

THEOREM 2.11. *Almost all ergodic components of an FD are EFDs.*

To conclude our discussion of FDs, we note the following consequence of the Besicovitch Density Point Theorem. A particular case of this is [11, Proposition 3.7]; the proof is the same, so is omitted.

PROPOSITION 2.12. *If $\mu \in \mathcal{M}$ and $\mu(A) > 0$, then, for μ almost all $x \in A$, we have*

$$\mathcal{TD}(\mu, x) = \mathcal{TD}(\mu_A, x).$$

2.4. CP distributions

CP processes, introduced by Furstenberg [9] (though in embryonic form go back to [8]), are analogous to FDs, except that the zooming in is done through b -adic cubes rather than cubes centered at the reference point.

NOTATION 2.13 (Dyadic systems). For simplicity, we restrict ourselves to dyadic CP processes. Let \mathcal{D} be the partition of B_1 into 2^d cubes of the form $I_1 \times \cdots \times I_d$, where $I_i \in \{[-1, 0), [0, 1]\}$. Given $x \in B_1$, let $D(x)$ be the only element of \mathcal{D} containing it. More generally, for $k \geq 1$, we let \mathcal{D}_k be the collection of cubes of the form $I_1 \times \cdots \times I_d$, where

$$I_i \in \{[-1, -1 + 2^{1-k}), \dots, [1 - 2 \cdot 2^{1-k}, 1 - 2^{1-k}), [1 - 2^{1-k}, 1]\}.$$

We refer to elements of \mathcal{D}_k as dyadic cubes of step k (or size 2^{1-k}). Further, if D is any cube, then write T_D for the orientation-preserving homothety mapping from \overline{D} onto B_1 .

DEFINITION 2.14 (CP magnification operator). With this notation, we define the (dyadic) *CP magnification operator* M on $\mathcal{P}(B_1) \times B_1$ by

$$M(\mu, x) = M^\square(\mu, x) := (T_{D(x)}^\square \mu, T_{D(x)}(x)).$$

This is defined whenever $\mu(D(x)) > 0$.

Note that, unlike FDs, here it is important to keep track of the orbit of the point that is being zoomed in upon. Note also that M acts on $\Xi := \mathcal{P}(B_1) \times B_1$.

REMARK 2.15. In Hochman’s work, CP processes are defined via dyadic partitions of the half-open cube $[-1, 1)^d$. This creates some technical issues with measures which give positive mass to the set $\{(x_1, \dots, x_n) : x_i = 1 \text{ for some } i\}$. Here we have followed Furstenberg’s original definition from [9]. Because ultimately we deal only with measures which give zero mass to the boundaries of cubes, this is just a matter of convenience.

The analogue of the quasi-Palm in this context is the *adaptedness* of distributions.

DEFINITION 2.16 (CPDs). A distribution Q on Ξ is *adapted* if there is a disintegration

$$\int f(\nu, x) dQ(\nu, x) = \iint f(\nu, x) d\nu(x) d\bar{Q}(\nu) \quad \text{for all } f \in C(\Xi), \quad (2.1)$$

where \bar{Q} is the projection of Q onto the measure component. Given a distribution Q on Ξ , its *intensity measure* is given by

$$[Q](A) := \int \mu(A) d\bar{Q}(\mu), \quad A \subset B_1.$$

A distribution on Ξ is a CPD if it is M -invariant (that is, $MQ = Q$), adapted, and its intensity measure is normalized Lebesgue measure on B_1 , which we denote by \mathcal{L} . The family of all CPDs is denoted by \mathcal{CPD} , and the ergodic ones by \mathcal{ECPD} .

Note that adaptedness can be interpreted in the following way: in order to sample a pair (μ, x) from the distribution Q , we have to first sample a measure μ according to \bar{Q} , and then sample a point x using the chosen distribution μ . This interpretation highlights the connection with the quasi-Palm property.

REMARK 2.17. The condition that $[Q]$ is Lebesgue is not part of the definition of a CP process given in [9, 11], and indeed there are important examples of adapted, M -invariant distributions with non-Lebesgue intensity. However, all distributions we consider do have this property, which is required repeatedly in the proofs. As a first useful consequence, note that if $[Q] = \mathcal{L}$, then for any fixed cube $B(x, r)$, Q almost all measures give zero mass to the boundary of $B(x, r)$. In particular, Q almost every measure gives zero mass to the boundary of the elements of \mathcal{D} . This helps us in dealing with the discontinuities inherent to the dyadic partition.

The usefulness of this condition was already implicit in Hochman’s work [11], where a random translation is often applied to ensure that the resulting CP processes have Lebesgue intensity. We also remark that, for us, CP processes are a tool towards the study of the scenery flow and FDs, so we have adopted the definition that happens to be most useful with this goal in mind.

We can define concepts similar to the scenery and tangent distributions (Definition 2.4) for CP processes.

DEFINITION 2.18 (CP scenery and micromasure distributions). Given a measure $\mu \in \mathcal{M}^{\square}$, $x \in B_1$ and $N \in \mathbb{N}$, we define the *CP scenery distribution* of μ at x along the scales $1, \dots, N$ by

$$\langle \mu, x \rangle_N := \frac{1}{N} \sum_{k=0}^{N-1} \delta_{M^k(\mu, x)}.$$

Any accumulation point of $\langle \mu, x \rangle_N$ in $\mathcal{P}(\Xi)$, as $N \rightarrow \infty$, is called a *micromasure distribution*, and the set of them is denoted by $\mathcal{MD}(\mu, x)$. We say that a measure $\mu \in \mathcal{M}$ *CP-generates* Q if

$$\langle \mu, x \rangle_N \longrightarrow Q, \quad \text{as } N \longrightarrow \infty,$$

that is, $\mathcal{MD}(\mu, x) = \{Q\}$, at μ almost every x .

REMARK 2.19. By compactness of $\mathcal{P}(\Xi)$, the set $\mathcal{MD}(\mu, x)$ is always non-empty and compact.

Again similarly as in Proposition 2.12, a consequence of the Besicovitch Density Point Theorem (in its version for dyadic cubes, which is simpler and can be seen from a martingale argument) yields the following proposition.

PROPOSITION 2.20. *If $\mu \in \mathcal{M}^\square$ and $\mu(A) > 0$, then, for μ almost all $x \in A$, we have*

$$\mathcal{MD}(\mu, x) = \mathcal{MD}(\mu_A, x).$$

Just as tangent distributions at typical points are FDs (Theorem 2.10), micromasure distributions at typical points are CPDs, but only after we randomly translate the measure. However, the only role of the random translation is to ensure that all micromasure distributions have Lebesgue intensity.

THEOREM 2.21. *Let $\mu \in \mathcal{M}$. The following holds for μ almost all x .*

- (i) *All distributions in $\mathcal{MD}(\mu, x)$ are adapted.*
- (ii) *If $Q \in \mathcal{MD}(\mu, x)$ has Lebesgue intensity, then Q is a CPD.*
- (iii) *For Lebesgue almost all $\omega \in B(0, \frac{1}{2})$, all distributions in $\mathcal{MD}(\mu + \omega, x + \omega)$ are CPDs.*

Proof. The first claim is [11, Proposition 5.4]. The second follows from the proof of [11, Proposition 5.5(2)]: although [11, Proposition 5.5] is stated for random translations of a fixed measure, the second part uses only the fact that the intensity measure of Q gives zero mass to all the boundaries of dyadic cubes. Finally, the last claim is precisely the content of [11, Proposition 5.5], except that there μ is assumed to be supported on $B(0, \frac{1}{2})$, but after rescaling the measure we can extend the result to arbitrary $\mu \in \mathcal{M}$. \square

Just as the ergodic components of FDs are again FDs, ergodic components of CPDs are again CPDs.

PROPOSITION 2.22. *Let Q be a CPD.*

- (i) *For \overline{Q} almost all μ and μ almost all x , we have $\mathcal{MD}(\mu, x) = \{Q_{(\mu, x)}\}$, where $Q_{(\mu, x)}$ is the ergodic component of (μ, x) .*
- (ii) *Almost all ergodic components of Q are CPDs.*

Proof. Let $Q = \int Q_\alpha dQ(\alpha)$ be the ergodic decomposition of Q . By the ergodic theorem, for a fixed $f \in C(\Xi)$ and Q_α almost all (μ, x) ,

$$\lim_{n \rightarrow \infty} \int f d\langle \mu, x \rangle_n = \int f dQ_\alpha.$$

Hence the same holds simultaneously for all f in a uniformly dense countable subset of $C(\Xi)$, and therefore for all $f \in C(\Xi)$. This yields the first claim.

For the second claim, note that from the first part of Theorem 2.21 and adaptedness of Q , it follows that for Q almost all (μ, x) , all the elements of $\mathcal{MD}(\mu, x)$ are adapted distributions. Thus, by the first part, Q_α is adapted for Q almost all α .

It remains to show $[Q_\alpha] = \mathcal{L}$ for Q almost all α . For this, we use some well-known facts on measure-theoretical entropy; [24, Chapters 4 and 8] contains all the definitions and facts we need.

Let $\Phi: B_1 \rightarrow B_1$ be equal to T_D^{-1} on D for each dyadic cube $D \in \mathcal{D}$. Let Δ be the union of the boundaries of dyadic cubes of first level. Note that, since Q is adapted, $[Q_\alpha](\Delta) = 0$ for Q almost all α . Also,

$$\iint f(\Phi(x)) d\mu(x) d\bar{Q}_\alpha(\mu) = \iint f(x) d\mu(x) d\bar{Q}_\alpha(\mu).$$

Thus each $[Q_\alpha]$ is Φ -invariant. The map Φ on $B_1 \setminus \Delta$ is naturally conjugated to the full shift on 2^d symbols. It follows that the system $(B_1, [Q_\alpha], \Phi)$ is measure-theoretically isomorphic to an m.p.s. on the full shift on 2^d symbols, and the invariant measure on the latter is the measure of maximal entropy if and only if $[Q_\alpha] = \mathcal{L}$. In particular, denoting measure-theoretical entropy by $h_\nu(\Phi)$, we have $h_{[Q_\alpha]}(\Phi) \leq d \log 2$, with equality if and only if $[Q_\alpha] = \mathcal{L}$. On the other hand, by the affinity of entropy,

$$h_{[Q]}(\Phi) = \int h_{[Q_\alpha]}(\Phi) dQ(\alpha).$$

We conclude that $[Q_\alpha] = \mathcal{L}$ for Q almost all α , as claimed. □

LEMMA 2.23. *The family \mathcal{CPD} is a convex subset of $\mathcal{P}(\Xi)$, and the set of extremal points is exactly \mathcal{ECPD} .*

Proof. The properties of adaptedness and having Lebesgue intensity are checked from definitions to be convex, so \mathcal{CPD} is indeed a convex set. Ergodic CPDs are extremal points of \mathcal{CPD} since they are extremal points for the larger set of M^\square -invariant measures, and conversely, since we know from Proposition 2.22 that the ergodic components of CPDs are CPDs. □

2.5. Extended CPDs

The operator M has an extended version M^* , defined on $\mathcal{M} \times B_1$ via

$$M^*(\mu, x) = (T_{D(x)}^* \mu, T_{D(x)} x).$$

We have the following analogue of Lemma 2.9.

LEMMA 2.24. *Given a CPD Q , there is an M^* -invariant distribution \hat{Q} on $\mathcal{M}^* \times B_1$ such that $\hat{Q}^\square = Q$, where \hat{Q}^\square is the push-down of \hat{Q} under $(\mu, x) \rightarrow (\mu^\square, x)$.*

Proof. The lemma follows from [11, Section 3.2], but we give a complete proof as the construction is used later. We can realize the system (Ξ, M, Q) as a process $(\xi_n)_{n \in \mathbb{N}}$, where $\xi_1 \sim Q$ and $M\xi_n = \xi_{n+1}$. By definition, this process is stationary with marginal Q (note that given ξ_n , the future of the process $\{\xi_m : m \geq n\}$ is deterministic). Any stationary one-sided process has an extension to a stationary two-sided process with the same finite-dimensional marginals; thus, let $(\xi_n)_{n \in \mathbb{Z}}$ be the two-sided extension of the above process, and denote its distribution by \mathbb{P}_Q . Then, in particular, \mathbb{P}_Q almost surely it holds that $M\xi_n = \xi_{n+1}$ for all $n \in \mathbb{Z}$.

Suppose that a \mathbb{P}_Q -typical sequence $\xi_n = (\mu_n, x_n)$ is given. For each $n \geq 0$ and $x \in B_1$, let $T_{x,n}$ be the orientation-preserving homothety that maps the dyadic square of size $2 \cdot 2^{-n}$ containing x onto B_1 (this is well defined for x not in the boundary of a dyadic cube). Let $E_n = T_{x_{-n}, n} B_1$. The sequence $(E_n)_{n \geq 0}$ is then an increasing sequence of compact sets, starting with B_1 . Furthermore, $M^k(\mu_{-n}, x_{-n}) = (\mu_{-n+k}, x_{-n+k})$ for all $n, k \in \mathbb{N}$. It follows from these considerations that the limit

$$\nu = \lim_{n \rightarrow \infty} \nu_n = \lim_{n \rightarrow \infty} T_{x_{-n}, n}^* \mu_{-n}$$

exists, in the sense that the measures ν_n are supported on E_n and are compatible: $\nu_{n+k}|_{E_n} = \nu_n$. Moreover, $\nu_0 = \mu_0$.

Note that $(\nu, x_0) \in \mathcal{M}^* \times B_1$ is a function of the sequence (μ_n, x_n) ; let \widehat{Q} be the push-down of \mathbb{P}_Q under this map. One can check from the definitions that this is the desired extension: \widehat{Q} is M^* -invariant, and $\widehat{Q}^\square = Q$. \square

The distribution \widehat{Q} is also called the extended version of Q . We note a consequence of the construction.

COROLLARY 2.25. *Let Q be a CPD, and let \widehat{Q} be its extended version. If $f: B_1 \rightarrow \mathbb{R}$ is a Borel function, then*

$$\int f(x) d\widehat{Q}(\mu, x) = \int f(x) dx.$$

Proof. By the construction of the extended version,

$$\int f(x) d\widehat{Q}(\mu, x) = \mathbb{E}_{\mathbb{P}_Q} f(x_0) = \int f(x) dQ(\mu, x) = \int f(x) dx,$$

using that the sequence (μ_n, x_n) has marginal Q and that Q is adapted and has Lebesgue intensity. \square

2.6. Weak convergence

To conclude this section, we collect a number of standard facts on weak convergence. We often have to prove weak convergence of distributions on Ξ . The following lemma shows that when considering convergence of CP sceneries, it is enough to establish convergence of the measure component.

LEMMA 2.26. *Let $\overline{\langle \mu, x \rangle}_N$ be the projection of $\langle \mu, x \rangle_N$ onto the measure part, that is,*

$$\overline{\langle \mu, x \rangle}_N := \frac{1}{N} \sum_{k=0}^{N-1} \delta_{T_{D^k(x)}^\square \mu},$$

where $D^k(x)$ is the dyadic cube of side length $2 \cdot 2^{-k}$ containing x . If for some measure μ and a CPD Q it holds that

$$\overline{\langle \mu, x \rangle}_N \longrightarrow \overline{Q} \quad \text{as } N \longrightarrow \infty$$

at μ almost every x , then μ CP-generates Q .

Proof. By Theorem 2.21(i) and the hypothesis, for μ , almost all x any subsequential limit of $\overline{\langle \mu, x \rangle}_N$ is an adapted distribution with measure marginal \overline{Q} , hence it equals Q . \square

The above lemma is repeatedly used without further reference in the later sections. We often need to use a metric which induces the weak topology on probability measures.

DEFINITION 2.27. For any compact metric space X , we define a distance $d_X(\mu, \nu)$ between two finite measures μ and ν on X by

$$d_X(\mu, \nu) = \sup_{f \in \text{Lip}_1(X)} \int f d(\mu - \nu),$$

where $\text{Lip}_1(X)$ is the class of Lipschitz functions $X \rightarrow \mathbb{R}$ with Lipschitz constant 1.

This metric is known by different names in the literature (bounded Lipschitz metric, Prokhorov distance, first Wasserstein metric). It induces the weak topology on $\mathcal{P}(X)$ (see, for example, [17, Chapter 14]; the statement there is for \mathbb{R}^d but the proof works on any locally compact metric space). We slightly abuse notation and denote by d both the metric above on the space of measures $\mathcal{P}(B_1)$ and on the space of distributions $\mathcal{P}(\mathcal{P}(B_1))$ as it is clear from the context which space we are dealing with.

REMARK 2.28. Even though the metric d_X is defined with arbitrary test functions in $\text{Lip}_1(X)$, it is enough to consider 1-Lipschitz functions uniformly bounded by some constant C_X . Indeed, since subtracting a constant does not change the value of $\int f d(\mu - \nu)$, we may assume $f(x_0) = 0$ for some fixed point $x_0 \in X$. But then, since $f \in \text{Lip}_1(X)$,

$$|f(x)| = |f(x) - f(x_0)| \leq d(x, x_0) \leq \text{diam}(X).$$

Even though weak convergence is defined in terms of continuous functions, it still holds for functionals whose discontinuity set is null for the limiting measure.

LEMMA 2.29. *Let X be a locally compact metric space, and let $\mu_n, \mu \in \mathcal{P}(X)$. If $\mu_n \rightarrow \mu$ weakly and $f: X \rightarrow \mathbb{R}$ is a function such that*

$$\mu(\{x \in X : f \text{ is discontinuous at } x\}) = 0,$$

then $\int f d\mu_n \rightarrow \int f d\mu$.

See, for example, [4, Theorem 2.7] for a stronger statement.

To finish this section, we show that to prove convergence of distributions in $\mathcal{P}(B_1)$, it is enough to consider test functions ‘with a finite resolution’. Recall that \mathcal{D}_k is the family of dyadic cubes of level k , and let \mathcal{F}_k be the class of functions $f: \mathcal{P}(B_1) \rightarrow \mathbb{R}$ such that $f(\mu)$ depends only on the values of $\mu(D)$, $D \in \mathcal{D}_k$.

LEMMA 2.30. *Let $Q_n, Q \in \mathcal{P}(B_1)$. If $\int f dQ_n \rightarrow \int f dQ$ for all $k \in \mathbb{N}$ and all $f \in \mathcal{F}_k$, then $Q_n \rightarrow Q$.*

Proof. If the functions in \mathcal{F}_k were continuous, then this would be a direct application of the Stone–Weierstrass Theorem. It would be possible to still rely on Stone–Weierstrass by approximating elements of \mathcal{F}_k by continuous functions in a suitable way, but we give a direct argument.

Given $\mu \in \mathcal{P}(B_1)$, let $\mu_k = \sum_{D \in \mathcal{D}_k} \mu(D) \delta_{z_D}$, where z_D is the center of D . Note that if $\varphi: B_1 \rightarrow \mathbb{R}$ is 1-Lipschitz, then $|\varphi(z) - \varphi(z_D)| \leq \sqrt{d} 2^{-k}$ for any $z \in D$, and therefore

$$\begin{aligned} \left| \int \varphi d\mu - \int \varphi d\mu_k \right| &= \left| \sum_{D \in \mathcal{D}_k} \mu(D) \left(\frac{\int_D \varphi d\mu}{\mu(D)} - \varphi(z_D) \right) \right| \\ &\leq \sum_{D \in \mathcal{D}_k} \mu(D) \sqrt{d} 2^{-k} = \sqrt{d} 2^{-k}. \end{aligned}$$

This shows $d(\mu, \mu_k) \leq \sqrt{d} 2^{-k}$.

Now let $f \in C(\mathcal{P}(B_1))$, and write $f_k(\mu) = f(\mu_k)$. Then $f_k \in \mathcal{F}_k$ by definition, and hence, by the hypothesis,

$$\lim_{n \rightarrow \infty} \int f_k dQ_n = \int f_k dQ.$$

Since f is continuous, it is uniformly continuous; hence given $\varepsilon > 0$ there is $k > 0$ such that $|f(\mu) - f(\nu)| < \varepsilon$ if $d(\mu, \nu) < \sqrt{d} 2^{-k}$. In particular, by the above, $|f(\mu) - f_k(\mu)| < \varepsilon$ for all

$\mu \in \mathcal{P}(B_1)$. By writing $f = (f - f_k) + f_k$, it follows from the above that

$$\limsup_{n \rightarrow \infty} \int f dQ_n \leq \int f dQ + \varepsilon,$$

and likewise with \liminf . Since $\varepsilon > 0$ was arbitrary, this completes the proof. \square

3. FDs form a closed set

In this section, we prove the following precise version of Theorem A.

THEOREM 3.1. *The family \mathcal{FD} is a closed set in the weak topology.*

The proof relies on the results in [11] relating FDs to CPDs, which we first recall. The main tool is the *centering* operation that provides a way to map CPDs onto FDs, and vice versa.

DEFINITION 3.2 (Centering operation). Let $C: \mathcal{M}^* \times B_1 \times [0, \log 2] \rightarrow \mathcal{M}^*$ be given by

$$C(\mu, x, t) = S_t^* T_x \mu.$$

If Q is a distribution on $\mathcal{M}^* \times B_1$, then its (continuous) *centering* $\text{cent}(Q)$ is the push-down of $Q \times \lambda$ under C , that is,

$$\text{cent}(Q) := C(Q \times \lambda),$$

where λ is normalized Lebesgue measure on $[0, \log 2]$. For illustration, see Figure 1.

THEOREM 3.3. *Let Q be a CPD, and let \widehat{Q} be its extended version (given by Lemma 2.24). Then $\text{cent}(\widehat{Q})$ is an extended FD.*

Conversely, given an extended FD P , there exists a CPD Q such that $P = \text{cent}(\widehat{Q})$, with \widehat{Q} the extended version of Q .

Proof. See [11, Theorems 1.14, 1.15 and Proposition 1.16]. We remark that although in [11] CPDs are not required to have Lebesgue intensity measure, [11, Proposition 1.16] states that one can find an appropriate CPD with this additional property. \square

This correspondence theorem between CPDs and FDs allows us to reduce the investigation back to CPDs. In this ‘discrete’ setting, the analogue of Theorem 3.1 is quite straightforward.

LEMMA 3.4. *The family \mathcal{CPD} is closed in $\mathcal{P}(\Xi)$.*

Proof. Write \mathcal{U} for the distributions in $\mathcal{P}(\Xi)$ with Lebesgue intensity measure. Note that $Q \in \mathcal{U}$ if and only if

$$\iint f d\nu d\bar{Q}(\nu) = \int f(x) dx \quad \text{for all } f \in C(B_1).$$

The left-hand side defines a continuous function of Q , so \mathcal{U} is a closed set. For fixed $f \in C(\Xi)$, both sides of (2.1) are continuous as a function of Q , so the family of adapted distributions is also closed.

It remains to show that if Q_n are CPDs and $Q_n \rightarrow Q$, then Q is M -invariant. Note that M is discontinuous in general, however, it is discontinuous only at pairs (μ, x) , where x has some coordinate equal to 0 (that is, x is in the boundary of two dyadic cubes of first level). Since we already know $Q \in \mathcal{U}$ and that Q is adapted, M is continuous off a set of Q -measure zero. Lemma 2.29 then tells us $MQ = \lim_{n \rightarrow \infty} MQ_n = Q$, as desired. \square

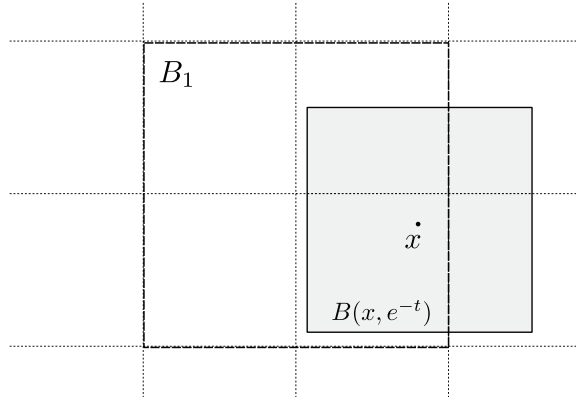


FIGURE 1. The centering $\text{cent}(Q)$ provides a distribution of magnifications $C(\mu, x, t)$, where the scale e^{-t} is chosen uniformly between $\frac{1}{2}$ and 1 and (μ, x) according to the distribution Q .

Proof of Theorem 3.1. Let $\{P_n\}$ be a sequence of restricted FDs such that $P_n \rightarrow P$ in $\mathcal{P}(B_1)$. For each n , the second part of Theorem 3.3 provides us with a CPD Q_n such that $\text{cent}(\widehat{Q}_n) = \widehat{P}_n$, where $\widehat{P}_n, \widehat{Q}_n$ are the corresponding extended versions. By Lemma 3.4 and compactness, there exists a CPD Q which is an accumulation point of the Q_n . Let \widehat{Q} be the extended version of Q . It is enough to prove $\text{cent}(\widehat{Q})^\square = P$, since then the first part of Theorem 3.3 yields that P is a restricted FD. In turn, this follows if we can prove that the map $Q \rightarrow \text{cent}(\widehat{Q})^\square$ is continuous on CPDs.

Note that $C(\mu, x, t)$ depends only on the restriction of μ to $B_2 = B(0, 2) = [-2, 2]^d$. Let

$$\mu^\diamond := \mu_{B_2},$$

with $Q^\diamond, \mathcal{M}^\diamond$, etc. being defined in the usual way. The desired continuity then follows if we can establish the following two claims.

- (i) The map $Q \rightarrow \widehat{Q}^\diamond$ is continuous from the set of CPDs to $\mathcal{P}(\mathcal{M}^\diamond \times B_1)$.
- (ii) The map $Q \rightarrow \text{cent}(Q)^\square$ is continuous from $\mathcal{P}(\mathcal{M}^\diamond \times B_1)$ to $\mathcal{P}(\mathcal{M})$.

These claims are proved in the following two lemmas. □

LEMMA 3.5. *The map $Q \rightarrow \widehat{Q}^\diamond$ is continuous on the set of CPDs.*

Proof. Let $f \in C(\mathcal{M}^\diamond \times B_1)$. We have to show that if Q_k, Q are CPDs, and $Q_k \rightarrow Q$, then

$$\int f(\nu^\diamond, x) d\widehat{Q}_k(\nu, x) \longrightarrow \int f(\nu^\diamond, x) d\widehat{Q}(\nu, x).$$

We may and will assume that $|f|$ is uniformly bounded by 1. Fix $n \geq 2$, and write $\Delta_n = \{y : \text{dist}(y, \partial B_1) \geq 2 \cdot 2^{-n}\}$. Decompose

$$\int f(\nu^\diamond, x) d\widehat{Q}(\nu, x) = \int_{x \in B_1 \setminus \Delta_n} f(\nu^\diamond, x) d\widehat{Q}(\nu, x) + \int_{x \in \Delta_n} f(\nu^\diamond, x) d\widehat{Q}(\nu, x).$$

Using $\|f\|_\infty \leq 1$ and Corollary 2.25, it follows that

$$\left| \int_{x \in B_1 \setminus \Delta_n} f(\nu^\diamond, x) d\widehat{Q}(\nu, x) \right| \leq \int \mathbf{1}_{B_1 \setminus \Delta_n}(x) d\widehat{Q}(\nu, x) \leq C_d 2^{-n},$$

and likewise for \widehat{Q}_k , where $C_d > 0$ depends on the dimension d only. Thus it is enough to show, for a fixed n ,

$$\lim_{k \rightarrow \infty} \int_{x \in \Delta_n} f(\nu^\diamond, x) d\widehat{Q}_k(\nu, x) = \int_{x \in \Delta_n} f(\nu^\diamond, x) d\widehat{Q}(\nu, x).$$

We use the notation of the proof of Lemma 2.24. Let (μ_n, x_n) be a \mathbb{P}_Q -typical point, and let (ν, x_0) be the resulting element of $\mathcal{M} \times B_1$. Note that if $x_{-n} \in \Delta_n$, then $T_{x_{-n}, n} B_1 \supset B_2$, and it follows that $\nu^\diamond = T_{x_{-n}, n}^\diamond \mu_{-n}$. Hence, keeping in mind, the construction of the extended version given in Lemma 2.24,

$$\begin{aligned} \int_{x \in \Delta_n} f(\nu^\diamond, x) d\widehat{Q}(\nu, x) &= \mathbb{E}_{\mathbb{P}_Q} \mathbf{1}_{\Delta_n}(x_{-n}) f(T_{x_{-n}, n}^\diamond \mu_{-n}, T_{x_{-n}, n} x_{-n}) \\ &= \int \mathbf{1}_{\Delta_n}(x) f(T_{x, n}^\diamond \mu, T_{x, n} x) dQ(\mu, x), \end{aligned}$$

where in the last step we used that the stationary sequence (μ_n, x_n) has marginal Q . Likewise, the same holds for Q_k in place of Q . The function $(\mu, x) \rightarrow \mathbf{1}_{\Delta_n}(x) f(T_{x, n}^\diamond \mu, T_{x, n} x)$ is continuous except for some pairs (μ, x) with x at the boundary of two dyadic cubes of side length 2×2^{-n} . Since Q is adapted and has Lebesgue intensity, it gives zero mass to this discontinuity set. We are done, thanks to Lemma 2.29. \square

LEMMA 3.6. *The map $Q \rightarrow \text{cent}(Q)^\square$ is continuous from $\mathcal{P}(\mathcal{M}^\diamond \times B_1)$ to $\mathcal{P}(B_1)$.*

Proof. Let $f \in C(\mathcal{M}^\square)$ and $Q \in \mathcal{P}(\mathcal{M}^\diamond \times B_1)$. By the definition of centering and Fubini,

$$\int f d(\text{cent}(Q)^\square) = \int \int_0^{\log 2} f(S_t^\square T_x^* \mu) d\lambda(t) dQ(\mu, x).$$

Even though each S_t^\square may be discontinuous (when the boundary of $B(0, e^{-t})$ has positive mass), for a given measure, there can be discontinuities only for countably many values of t . It follows from the bounded convergence theorem that the inner integral is a continuous function of $T_x^* \mu$, which in turn is a continuous function of (μ, x) , and the lemma follows. \square

4. Splicing and the simplex of FDs

The goal of this section is to establish Theorem B, which is precisely stated as follows.

THEOREM 4.1. *The convex set \mathcal{FD} is a Poulsen simplex (as a subset of the locally convex space of finite Radon measures). In other words, extremal points of \mathcal{FD} are weakly dense in \mathcal{FD} .*

Again, invoking the centering operation all we need to prove is the following proposition.

PROPOSITION 4.2. *The family \mathcal{CPD} is a Poulsen simplex.*

We show how to deduce Theorem 4.1 from this proposition; the remainder of this section is devoted to the proof of the proposition.

Proof of Theorem 4.1 (Assuming Proposition 4.2). By Theorem 3.3, the centering operation maps CPDs onto FDs. The centering map is in fact a factor map from the suspension flow with base $(\mathcal{M} \times B_1, M)$ and height $\log 2$ onto \mathcal{FD} ; see [11] for the details. It follows that if

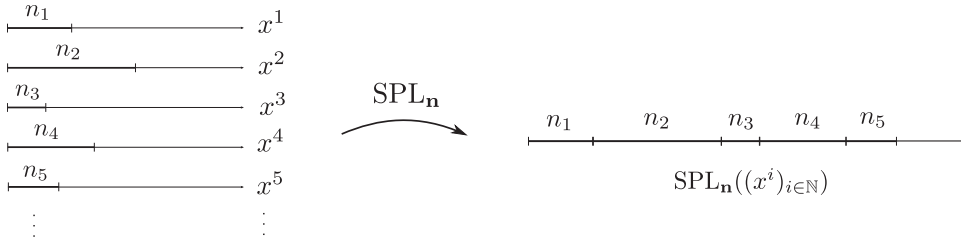


FIGURE 2. Given $\mathbf{n} = (n_i)_{i \in \mathbb{N}}$, the splicing map $\text{SPL}_{\mathbf{n}}$ takes the first n_i coordinates of the word $x^i \in B_1$ and concatenates the obtained finite words into a point in B_1 .

Q is an ergodic extended CPD, then its centering is an ergodic FD. Moreover, if a CPD is ergodic, then so is its extended version. Indeed, the construction in Lemma 2.24 shows that the extended version is a factor of the two-sided extension of Q , which is ergodic if and only if Q is ergodic.

Finally, we recall from Lemmas 3.5 and 3.6 that the map $Q \rightarrow \text{cent}(\widehat{Q})^\square$ is continuous. Thus the image of the dense set of ergodic CPDs under this map is dense in the set of restricted FDs and, by the above observations, consists of ergodic distributions, concluding the proof. \square

4.1. The splicing operation

The construction of the ergodic CPD which approximates a given CPD is done via an operation which we term the *splicing* of scales. To introduce the notation, it is convenient to identify points in B_1 with dyadic sequences.

NOTATION 4.3 (Coding dyadic cubes). Write $\mathcal{A} = \{0, 1, \dots, 2^d - 1\}$. Enumerate \mathcal{D} (the dyadic subcubes of $B = B_1$ of first level; recall Subsection 2.4) as $\{B_i : i \in \mathcal{A}\}$. Each $x \in \mathcal{A}^k$ then corresponds to a dyadic cube B_x of generation k and side length 2^{1-k} . We silently identify x with B_x whenever there is no possibility of confusion. Moreover, we also identify each $x \in B$ with $x \in \mathcal{A}^{\mathbb{N}}$ such that

$$\{x\} = \bigcap_{k \in \mathbb{N}} B_{x|_k}.$$

If x is any sequence of length at least b (possibly infinite), then we write $x_a^b = (x_{a+1}, x_{a+2}, \dots, x_b)$. Also, if x is an infinite word, then we write $x_n^\infty = (x_{n+1}, x_{n+2}, \dots)$; geometrically, $x_n^\infty = T_{B_{x_n}} x$. In the case $a = 0$, we also write $x|_b = x_0^b$. We allow the empty word \emptyset , which in our identification corresponds to B_1 ; we note $x_a^a = \emptyset$. If $\mu \in \mathcal{P}(B)$ and $x \in \mathcal{A}^k$, then we write

$$\mu(\cdot|x) = T_{B_x} \mu_{B_x}.$$

In symbolic notation, if $y \in \mathcal{A}^\ell$, then $\mu(y|x) = \mu(xy)/\mu(x)$.

We can now give the definition of the splicing map.

DEFINITION 4.4 (Splicing map). Given a sequence $\mathbf{n} = (n_i)_{i \in \mathbb{N}}$ of integers, we define $\text{SPL}_{\mathbf{n}}: B_1^{\mathbb{N}} \rightarrow B_1$ as

$$\text{SPL}_{\mathbf{n}}((x^i)_{i \in \mathbb{N}}) := (x^1|_{n_1} x^2|_{n_2} \dots),$$

where the notation on the right-hand side indicates concatenation of words. For illustration, see Figure 2.

Of particular interest to us is the action of the splicing map on product measures $\times_{i=1}^{\infty} \mu^i$ on $B_1^{\mathbb{N}}$, and also on product distributions. Let

$$\nu = \text{SPL}_{\mathbf{n}} \left(\times_{i=1}^{\infty} \mu^i \right).$$

By the definition of splicing, the ν -mass of a finite word y (or equivalently, a dyadic cube D) is built from the μ_i -masses of consecutive subwords of y whose length comes from the sequence \mathbf{n} . To make this statement precise, for $k \in \mathbb{N}$, define the sum

$$S_k = S_k(\mathbf{n}) := n_1 + n_2 + \cdots + n_k.$$

LEMMA 4.5. *If $k \in \mathbb{N}$, $S_k \leq N \leq S_{k+1}$ and $y \in \mathcal{A}^N$, then*

$$\nu(y) = \mu^1(y_0^{S_1}) \mu^2(y_{S_1}^{S_2}) \cdots \mu^{k-1}(y_{S_{k-1}}^{S_k}) \mu^k(y_{S_k}^N). \quad (4.1)$$

Proof. If $x \in \mathcal{A}^{S_k+i}$ and $y \in \mathcal{A}^j$, with $0 \leq i \leq i+j \leq n_{k+1}$, then

$$\nu(y|x) = \mu^k(y|x_{S_k}^{S_k+i}). \quad (4.2)$$

In particular, if $x \in \mathcal{A}^{S_k}$ and $y \in \mathcal{A}^j$ with $0 \leq j \leq n_{k+1}$, then $\nu(y|x) = \mu^k(y)$. By iterating (4.2), if $S_k \leq N \leq S_{k+1}$ and $y \in \mathcal{A}^N$, then

$$\nu(y) = \mu^1(y_0^{S_1}) \mu^2(y_{S_1}^{S_2}) \cdots \mu^{k-1}(y_{S_{k-1}}^{S_k}) \mu^k(y_{S_k}^N)$$

as claimed. □

Relying on this lemma, by choosing a suitable sequence \mathbf{n} , we can now control the frequency of occurrences of the measures μ^i in the CP scenery $\langle \nu, x \rangle_N$.

4.2. Proof of Proposition 4.2

We are now ready to establish Proposition 4.2. We know from Lemma 3.4 that \mathcal{CPD} is compact. Moreover, by Lemma 2.23 and the existence and uniqueness of the ergodic decomposition, \mathcal{CPD} is a Choquet simplex. Thus, for Proposition 4.2, we need to show only that \mathcal{ECPD} is dense in \mathcal{CPD} .

PROPOSITION 4.6. *Ergodic CPDs are dense in \mathcal{CPD} .*

Proof. The density of ergodic CPDs is implied by the Krein–Milman Theorem if we are able to prove that, given a rational probability vector $(t_1/q, \dots, t_k/q)$, and given ergodic CPDs R_1, \dots, R_k , there is a sequence of ergodic CPDs Q^N converging to $(1/q) \sum_{i=1}^k t_i R_i$ as $N \rightarrow \infty$.

To find the sequence Q^N , let $\text{SPL} = \text{SPL}_{\mathbf{n}}$ be the splicing map corresponding to the k -periodic sequence

$$\mathbf{n} = (Nt_1, Nt_2, \dots, Nt_k)^{\infty} = (Nt_1, \dots, Nt_k, Nt_1, \dots, Nt_k, \dots).$$

Note that, as t_i/q is rational, Nt_i is an integer for all $i \in \{1, \dots, k\}$. Write $\tilde{R} = \times_{i=1}^k R_i$ and define an adapted distribution $P = P^N$ by setting

$$\bar{P} = \text{SPL}(\tilde{R}^{\mathbb{N}}),$$

that is, we take the distributions R_i in the product k -periodically. Then define Q^N by setting the measure marginal

$$\bar{Q}^N = \frac{1}{N} \sum_{j=0}^{N-1} \bar{Q}_j,$$

where \bar{Q}_j is the push-forward of (Ξ, P) under the map $(\mu, x) \rightarrow \mu(\cdot | x_0^j)$. Then Q^N is an ergodic CPD with $Q^N \rightarrow (1/q) \sum_{i=1}^k t_i R_i$ as $N \rightarrow \infty$. These two facts are verified in Lemmas 4.7 and 4.8. \square

LEMMA 4.7. *For fixed N , the distribution $Q = Q^N$ is an ergodic CPD.*

Proof. The outline of the proof is simple: we show that Q has Lebesgue intensity by definition, and then that $\mathcal{MD}(\mu, x) = \{Q\}$ for Q almost all μ and μ almost all x . Then it follows from Theorem 2.21 that Q is a CPD (here we need to know that Q has Lebesgue intensity), and then from Proposition 2.22 that Q is ergodic. We proceed to the details.

We start by showing that P has Lebesgue intensity. Since each R_i has Lebesgue intensity, using (4.1) and the k -periodicity of the sequence \mathbf{n} , we find that if $K = S_{\ell k}$ for some $\ell \in \mathbb{N}$, then

$$\begin{aligned} [P](x_0^K) &= \int \nu(x_0^K) d\bar{P}(\nu) = \prod_{j=1}^{\ell} \prod_{i=1}^k \int \mu(x_{S_{j+i-1}}^{S_{j+i}}) d\bar{Q}_{j+i}(\mu) \\ &= \prod_{j=1}^{\ell} \prod_{i=1}^k \int \mu(x_{S_{i-1}}^{S_i}) d\bar{R}_i(\mu) \\ &= \left(\prod_{i=1}^k 2^{-dNt_i} \right)^{\ell} = 2^{-dK}. \end{aligned}$$

Since cubes of the form x_0^K generate the Borel σ -algebra, P has Lebesgue intensity as claimed. Now

$$\begin{aligned} [Q_j](x_0^{\ell}) &= \int \mu(x_0^{\ell}) d\bar{Q}_j(\mu) = \int \mu(x_0^{\ell} | y_0^j) d\mu(y) d\bar{P}(\mu) \\ &= \int \frac{\mu((y_0^j)(x_0^{\ell}))}{\mu(y_0^j)} d\mu(y) d\bar{P}(\mu) \\ &= \sum_{z \in \mathcal{A}^j} \int \mu((zx)_0^{j+\ell}) d\bar{P}(\mu) = \mathcal{L}(x_0^{\ell}), \end{aligned}$$

showing that Q_j and therefore the average Q also has Lebesgue intensity.

Next, we claim

$$\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{i=0}^{L-1} \delta_{\mu(\cdot | x_0^{iN})} = \bar{P} \quad \text{for } P \text{ almost all } (\mu, x). \quad (4.3)$$

In essence, this is a consequence of the ergodic theorem for product measures under the shift. After re-indexing, the splicing map $\text{SPL}: B_1^{\mathbb{N}} \rightarrow \mathbb{N}$ induces a map $\text{SPL}: (B_1^k)^{\mathbb{N}} \rightarrow B_1$ on the space of k -tuples $(B_1^k)^{\mathbb{N}}$ by

$$\text{SPL}(\eta) = (\Phi(\eta_1)\Phi(\eta_2)\cdots), \quad \eta = (\eta_1, \eta_2, \dots) \in (B_1^k)^{\mathbb{N}},$$

where for a given $\zeta = (x^1, \dots, x^k) \in B_1^k$, we define

$$\Phi(\zeta) = (x^1|_{Nt_1} \cdots x^k|_{Nt_k}).$$

Using this description, we can reformulate (4.3) as

$$\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{i=0}^{L-1} \delta_{\text{SPL}(\sigma^i \eta)} = \text{SPL}(\tilde{R}^{\mathbb{N}}) \quad \text{for } \tilde{R}^{\mathbb{N}} \text{ almost all } \eta,$$

where σ is the shift on the sequence space $(B_1^k)^{\mathbb{N}}$. But this is a consequence of the ergodic theorem applied to the ergodic system $((B_1^k)^{\mathbb{N}}, \tilde{R}^{\mathbb{N}})$. Indeed, we need to show that for any $f \in C(B_1)$ and $\tilde{R}^{\mathbb{N}}$ almost all η ,

$$\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{i=0}^{L-1} f(\text{SPL}(\sigma^i \eta)) = \int f d\text{SPL}(\tilde{R}^{\mathbb{N}}).$$

In turn, it is enough to verify this for f in a countable dense subset of $C(B_1)$, and hence for a fixed $f \in C(B_1)$. But this holds by the ergodic theorem applied to the function $f \circ \text{SPL}$.

Next, we claim

$$\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{i=0}^{L-1} \delta_{\mu(\cdot | x_0^{iN+j})} = \bar{Q}_j \quad \text{for } P \text{ almost all } (\mu, x). \quad (4.4)$$

We start by noting

$$\mu(\cdot | x_0^{iN+j}) = \mu(\cdot | x_0^{iN})(\cdot | x_{iN}^{iN+j}).$$

(The notation on the right-hand side means $\nu(\cdot | x_{iN}^{iN+j})$, where $\nu = \mu(\cdot | x_0^{iN})$.) Indeed, it is straightforward to check the equality for cubes $[z]$ which form a basis of the σ -algebra.

Given $\eta \in (B_1^k)^{\mathbb{N}}$, let

$$\Psi(\eta) = \text{SPL}(\eta)(\cdot | \Phi(\eta_1)_0^j).$$

Using our previous notation, the last observation and the definitions of \bar{P} and \bar{Q}_j , we find that (4.4) is equivalent to

$$\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{i=0}^{L-1} \delta_{\Psi(\sigma^i \eta)} = \Psi(\tilde{R}^{\mathbb{N}}) \quad \text{for } \tilde{R}^{\mathbb{N}} \text{ almost all } \eta.$$

Just as before, this follows from the ergodic theorem. Averaging over j , we conclude that

$$\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{i=0}^{L-1} \delta_{\mu(\cdot | x_0^i)} = \bar{Q} \quad \text{for } P \text{ almost all } (\mu, x).$$

Now from Theorem 2.21(2), we deduce that Q is a CPD. Since a full P -measure set has positive Q -measure, the second part of Proposition 2.22 shows that Q is ergodic, finishing the proof. \square

LEMMA 4.8. *It holds that*

$$\lim_{N \rightarrow \infty} Q^N = \frac{1}{q} \sum_{i=1}^k t_i R_i.$$

Proof. By Lemma 2.30, we only have to prove that if $f \in \mathcal{F}_p$ for some p , then

$$\lim_{N \rightarrow \infty} \int f dQ^N = \frac{1}{q} \sum_{i=1}^k t_i \int f dR_i.$$

Recall that \mathcal{F}_p is the class of functions $f: \mathcal{P}(B_1) \rightarrow \mathbb{R}$ such that $f(\mu)$ depends only on the values $\mu(D)$, $D \in \mathcal{D}_p$. Define the set of indices

$$\mathcal{G}_{N,i} = \{j \in \{0, \dots, N-1\} : N(t_1 + \dots + t_{i-1}) \leq j \leq j+p \leq N(t_1 + \dots + t_i)\}.$$

Note that (since we are keeping p fixed)

$$\lim_{N \rightarrow \infty} \frac{\#\mathcal{G}_{N,i}}{t_i N} = 1 \quad \text{for all } i \in \{1, \dots, k\}.$$

On the other hand, if $j \in \mathcal{G}_{N,i}$, then it follows from (4.2) and the definition of \mathcal{F}_p that

$$\int f dQ_j^N = \int f(\mu(\cdot|y_0^j)) d\mu(y) dP(\mu) = \int f(\mu(\cdot|y_0^j)) d\mu(y) dR_j(\mu) = \int f dR_j,$$

using that R_j is a CPD in the last equality. We conclude that

$$\left| \int f dQ^N - \frac{1}{q} \sum_{i=1}^k t_i \int f dR_i \right| \leq \|f\|_\infty \frac{\sum_{i=1}^k |t_i N - \#\mathcal{G}_{N,i}|}{N} \rightarrow 0$$

as $N \rightarrow \infty$. □

5. Every FD is generated by a USM

In this section, we establish Theorem C, which we restate as follows.

THEOREM 5.1. *For any $P \in \mathcal{FD}$ there is a USM μ which generates P . In other words, there is a Radon measure μ such that $\mathcal{TD}(\mu, x) = \{P\}$ for μ almost all x .*

First, we note that for EFDs this is a consequence of the ergodic theorem.

LEMMA 5.2. *Let P be an FD. Then P almost all μ are USM generating the ergodic component of P_μ . In particular, if P is an EFD, then P almost all measures generates P .*

Proof. This follows from [11, Theorem 3.9] and the ergodic decomposition. □

In particular, if P is an EFD, then there exists at least one measure generating P . If P is not ergodic, then this is still true, but requires a more involved argument using the splicing operation introduced in the previous section.

Yet again, the corresponding statement for CPDs is easier to prove, and implies Theorem 5.1 by invoking the centering operation.

PROPOSITION 5.3. *If $\mu \in \mathcal{M}$ and Q is a CPD, then at μ almost every x where $Q \in \mathcal{MD}(\mu, x)$, also $P = \text{cent}(\widehat{Q})^\square \in \mathcal{TD}(\mu, x)$, where \widehat{Q} is the extended version of Q . In fact, if μ CP-generates Q , then μ is a USM generating P .*

Proof. This is essentially proved in the course of the proof of [11, Proposition 5.5(3)]. Although in that proposition the setting is that of an arbitrary measure that has been translated by a random vector, in the proof of the third part what really gets proved is that if μ is a measure such that for μ typical x , the sequence $\langle \mu, x \rangle_{N_i}$ converges to a CPD Q along some sequence (N_i) , and $\langle \mu \rangle_{x, T_i} \rightarrow P$ as $i \rightarrow \infty$ for $T_i := N_i \log 2$, $i \in \mathbb{N}$, then

$$P = \text{cent}(\widehat{Q})^\square.$$

The point of the first two parts of [11, Proposition 5.5] is that a random translation of a fixed measure does satisfy these conditions. This yields the first claim.

For the latter statement, if μ CP-generates Q , then by definition for μ typical x , we have $\langle \mu, x \rangle_{N_i} \rightarrow Q$ for all sequences $N_i \rightarrow \infty$. Hence any accumulation point of the scenery distributions $\langle \mu \rangle_{x, T_i}$ must equal $\text{cent}(\widehat{Q})^\square$, and we conclude that μ generates $\text{cent}(\widehat{Q})^\square$, as claimed. □

In light of this proposition and the equivalence between FDs and CPDs given in Theorem 3.3, Theorem 5.1 is established once we prove the following proposition.

PROPOSITION 5.4. *For any CPD Q there exists a measure $\mu \in \mathcal{M}$ which CP-generates Q .*

Proof. Write

$$\mathcal{G} := \{Q \in \mathcal{CPD} : \text{there exists } \mu \in \mathcal{M} \text{ which CP-generates } Q\}.$$

We need to prove $\mathcal{G} = \mathcal{CPD}$. If Q is ergodic, then we know from Proposition 2.22 that Q almost all μ do generate Q . On the other hand, we saw in Proposition 4.2 that ergodic CPDs are dense. Hence it is enough to show that \mathcal{G} is closed under the weak topology.

In the course of the proof, we use notation from Subsection 4.1. Let $Q^i \in \mathcal{G}$ and suppose there exists

$$Q = \lim_{i \rightarrow \infty} Q^i.$$

Since each Q^i is a CPD, we know from Lemma 3.4 that Q is a CPD, so in order to show $Q \in \mathcal{G}$ we are required to construct a measure μ which CP-generates Q . For each $i \in \mathbb{N}$, let μ^i be a measure CP-generating Q^i . Fix $0 < \varepsilon < 1$ and choose a sequence $\varepsilon_i \downarrow 0$ such that

$$\prod_{i=1}^{\infty} (1 - \varepsilon_i) = \varepsilon.$$

Since μ^i generates Q^i we can find $m_i \in \mathbb{N}$ such that $\mu^i(U_i) > 1 - \varepsilon_i$ for the set

$$U_i = \{x \in B_1 : d(\langle \mu^i, x \rangle_N, \overline{Q^i}) < \varepsilon_i \text{ for every } N \geq m_i\}.$$

We use the sequence (m_i) to construct a sequence (n_i) as follows: let $n_1 = \max\{e^{m_1}, e^{m_2}\}$ and for $i > 1$ put

$$k_i := \max\{e^{n_{i-1}}, e^{m_i}, e^{m_{i+1}}\} \quad \text{and} \quad n_i := m_i + k_i.$$

We let $\text{SPL} = \text{SPL}_{\mathbf{n}}$ be the splicing map associated to the sequence $\mathbf{n} = (n_i)_{i \in \mathbb{N}}$; recall Subsection 4.1. Write

$$\mu := \text{SPL} \left(\times_{i=1}^{\infty} \mu^i \right) \quad \text{and} \quad U := \text{SPL} \left(\times_{i=1}^{\infty} U_i \right).$$

By the definition of product measure,

$$\mu(U) = \lim_{M \rightarrow \infty} \prod_{i=1}^M \mu^i(U_i) \geq \lim_{M \rightarrow \infty} \prod_{i=1}^M (1 - \varepsilon_i) = \varepsilon > 0.$$

Hence μ_U is well defined and

$$\mathcal{MD}(\mu_U, x) = \mathcal{MD}(\mu, x)$$

for μ almost every $x \in U$ by Proposition 2.20. Hence if we can prove for a fixed $z \in U$ that $\langle \mu, z \rangle_N \rightarrow \overline{Q}$ as $N \rightarrow \infty$, then the normalized restriction μ_U CP-generates Q by definition. This is what we do in Lemma 5.5. \square

LEMMA 5.5. *Given $z \in U$, we have*

$$\lim_{N \rightarrow \infty} \langle \mu, z \rangle_N = \overline{Q}.$$

Proof. Choose $x^i \in U_i$, $i \in \mathbb{N}$, such that $z = \text{SPL}(x^1, x^2, \dots)$. Define the sum

$$S_i = S_i(\mathbf{n}) = n_1 + n_2 + \dots + n_i.$$

For each $N \in \mathbb{N}$ choose $i = i(N) \in \mathbb{N}$ such that $S_i \leq N < S_{i+1}$. Note that the sequence $(i(N))_{N \in \mathbb{N}}$ increases to infinity as the numbers S_i increase to infinity. We do not write the dependence of i on N explicitly, but it is important to keep in mind that $i \rightarrow \infty$ as $N \rightarrow \infty$. Write the proportions

$$p_N = \frac{S_i}{N} \quad \text{and} \quad q_N = 1 - p_N.$$

Then $0 \leq p_N, q_N \leq 1$ for any $N \in \mathbb{N}$.

Fix a control parameter $K \in \mathbb{N}$, which we do not touch until the end of the proof after we have let $N \rightarrow \infty$. Assume that N is so large (depending on K) that $S_{i-1} + K < S_i$. This is possible since $n_i \rightarrow \infty$ as $N \rightarrow \infty$.

We make use of the metric d from Definition 2.27. The aim is to estimate $d(\overline{\langle \mu, z \rangle}_N, \overline{Q})$ and for this we need to find suitable distribution decompositions of the difference $\overline{\langle \mu, z \rangle}_N - \overline{Q}$. We provide a different decomposition of this difference depending on the position of $N - S_i$ with respect to m_i . In both cases, we obtain a representation for the average $\overline{\langle \mu, z \rangle}_N - \overline{Q}$ and we see that when N is very large, this representation is close to 0. For this purpose, fix $f \in \text{Lip}_1(X)$ and let $M := \|f\|_\infty$.

(1) Suppose $N - S_i > m_{i+1}$. Then split the average $\overline{\langle \mu, z \rangle}_N$ into

$$p_N \left(\frac{1}{S_i} \sum_{k=S_{i-1}}^{S_i-K} \delta_{\mu(\cdot|z_0^k)} \right) + q_N \left(\frac{1}{N - S_i} \sum_{k=S_i+K}^{N-1} \delta_{\mu(\cdot|z_0^k)} \right) + \frac{1}{N} \sum_{\text{rest of } k} \delta_{\mu(\cdot|z_0^k)}$$

and denote this sum by $p_N F_N + q_N G_N + E_N$. Then, as $p_N + q_N = 1$, we obtain

$$\int f d(\overline{\langle \mu, z \rangle}_N - \overline{Q}) = p_N \int f d(F_N - \overline{Q}) + q_N \int f d(G_N - \overline{Q}) + \int f dE_N.$$

Moreover, we continue splitting:

$$F_N - \overline{Q} = (F_N - \overline{Q}^i) + (\overline{Q}^i - \overline{Q}) \quad \text{and} \quad G_N - \overline{Q} = (G_N - \overline{Q}^{i+1}) + (\overline{Q}^{i+1} - \overline{Q}).$$

Thus, in this case, we obtain an estimate for $\int f d(\overline{\langle \mu, z \rangle}_N - \overline{Q})$ in terms of

$$p_N d(F_N, \overline{Q}^i) + p_N d(\overline{Q}^i, \overline{Q}) + q_N d(G_N, \overline{Q}^{i+1}) + q_N d(\overline{Q}^{i+1}, \overline{Q}) + \int f dE_N$$

as the first four terms can be estimated from above with their d -distance. We claim

$$d(F_N, \overline{Q}^i) \leq \sqrt{d} 2^{-dK} + \varepsilon_i, \tag{5.1}$$

$$d(G_N, \overline{Q}^{i+1}) \leq \frac{S_i - K - S_{i-1}}{S_i} \cdot \sqrt{d} 2^{-dK} + \varepsilon_{i+1}, \tag{5.2}$$

$$\int f dE_N \leq M \cdot \frac{S_{i-1} + K}{S_i}. \tag{5.3}$$

In order not to interrupt the flow of the proof, these are shown in Lemma 5.6.

(2) Suppose $N - S_i \leq m_{i+1}$. Then split the average $\overline{\langle \mu, z \rangle}_N$ into

$$\frac{1}{N} \sum_{k=S_{i-1}}^{S_i-K} \delta_{\mu(\cdot|z_0^k)} + \frac{1}{N} \sum_{\text{rest of } k} \delta_{\mu(\cdot|z_0^k)} =: F'_N + E'_N.$$

In this case, we write

$$\begin{aligned} \int f d(\overline{\langle \mu, z \rangle}_N - \overline{Q}) &= \int f d(F'_N - \overline{Q}) + \int f dE'_N \\ &= \int f d(F'_N - \overline{Q}^i) + \int f d(\overline{Q}^i - \overline{Q}) + \int f dE'_N. \end{aligned}$$

Thus we obtain an estimate for $d(\overline{\langle \mu, z \rangle}_N, \overline{Q})$ in terms of

$$d(F'_N, \overline{Q}^i) + d(\overline{Q}^i, \overline{Q}) + \int f dE'_N.$$

Now we have the estimates

$$d(F'_N, \overline{Q}^i) \leq \frac{S_i - K - S_{i-1}}{S_i} \cdot \sqrt{d} 2^{-dK} + \varepsilon_i, \quad (5.4)$$

$$\int f dE'_N \leq M \cdot \frac{S_{i-1} + K + m_{i+1}}{S_i}. \quad (5.5)$$

The proof of these is deferred to Lemma 5.7.

We are left to analyze the estimates above. Note that when $N \rightarrow \infty$, the fractions $S_{i-1}/S_i \rightarrow 0$ so we have

$$\frac{S_i - K - S_{i-1}}{S_i} \rightarrow 1 \quad \text{and} \quad \varepsilon_i \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

Thus we see that these upper bounds in (5.1), (5.2) and (5.4) tend to 0 if we first let $N \rightarrow \infty$ and then $K \rightarrow \infty$. Moreover, by the choices of n_i , the numbers $m_{i+1}/S_i \rightarrow 0$ as $N \rightarrow \infty$, so the upper bounds in (5.3) and (5.5) tend to 0 as $N \rightarrow \infty$ for any K . Since by our assumption $Q^i \rightarrow Q$ as $N \rightarrow \infty$, the terms $d(\overline{Q}^i, \overline{Q}), d(\overline{Q}^{i+1}, \overline{Q}) \rightarrow 0$ as $N \rightarrow \infty$, so the proof of the proposition is complete. \square

We finish by proving the remaining estimates in the proof of Lemma 5.5.

LEMMA 5.6. *If $N - S_i > m_{i+1}$, then estimates (5.1)–(5.3) hold.*

Proof. Recall that $z = \text{SPL}(x^1, x^2, \dots)$, so by the triangle inequality,

$$d(F_N, \overline{Q}^i) \leq d(F_N, \overline{\langle \mu^i, x^i \rangle}_{S_i}) + d(\overline{\langle \mu^i, x^i \rangle}_{S_i}, \overline{Q}^i)$$

and

$$d(G_N, \overline{Q}^{i+1}) \leq d(G_N, \overline{\langle \mu^{i+1}, x^{i+1} \rangle}_{N-S_i}) + d(\overline{\langle \mu^{i+1}, x^{i+1} \rangle}_{N-S_i}, \overline{Q}^{i+1}).$$

Since $z \in U$, we know that the points $x^i \in U_i$ and $x^{i+1} \in U_{i+1}$. Moreover, $N - S_i > m_{i+1}$ and $S_i > m_i$, so by the definitions of U_i and U_{i+1} , we have

$$d(\overline{\langle \mu^i, x^i \rangle}_{S_i}, \overline{Q}^i) < \varepsilon_i \quad \text{and} \quad d(\overline{\langle \mu^{i+1}, x^{i+1} \rangle}_{N-S_i}, \overline{Q}^{i+1}) < \varepsilon_{i+1}.$$

Next, let us look at the term $d(G_N, \overline{\langle \mu^{i+1}, x^{i+1} \rangle}_{N-S_i})$. Note that if $S_i \leq k \leq N - 1$, then

$$d(\mu(\cdot | x_0^k), \mu^{i+1}(\cdot | x_0^{k-S_i})) \leq \sqrt{d} \cdot 2^{-K}.$$

This follows from (4.2), just like in the proof Lemma 2.30. Using this, we estimate

$$\begin{aligned} d(G_N, \overline{\langle \mu^{i+1}, x^{i+1} \rangle}_{N-S_i}) &\leq \frac{1}{N - S_i} \sum_{k=S_i}^{N-1} d(\mu(\cdot | x_0^k), \mu^{i+1}(\cdot | x_0^{k-S_i})) \\ &\leq \frac{1}{N - S_i} \sum_{k=S_i}^{N-1} \sqrt{d} 2^{-dK} = \sqrt{d} 2^{-dK}. \end{aligned}$$

Moreover,

$$d(F_N, \overline{\langle \mu^i, x^i \rangle}_{S_i}) \leq \frac{1}{S_i} \sum_{k=S_{i-1}}^{S_i-K} d(\mu(\cdot | x_0^k), \mu^i(\cdot | x_0^{k-S_{i-1}})) \leq \frac{S_i - K - S_{i-1}}{S_i} \cdot \sqrt{d} 2^{-dK}.$$

Moreover, for the distribution E_N , we see that the number of elements in its sum is exactly

$$N - [(S_i - K) - S_{i-1}] - [N - S_i] = S_{i-1} + K,$$

and so by $N \geq S_i$ and $M = \|f\|_\infty$ we obtain the claim

$$\int f dE_N \leq M \cdot \frac{S_{i-1} + K}{N} \leq M \cdot \frac{S_{i-1} + K}{S_i}. \quad \square$$

LEMMA 5.7. *If $N - S_i \leq m_{i+1}$, then estimates (5.4) and (5.5) hold.*

Proof. An argument symmetric to the proof of Lemma 5.6 when estimating G_N shows

$$d(F'_N, \bar{Q}^i) \leq \frac{S_i - K - S_{i-1}}{N} \cdot \sqrt{d}2^{-dK} + \varepsilon_i \leq \frac{S_i - K - S_{i-1}}{S_i} \cdot \sqrt{d}2^{-dK} + \varepsilon_i$$

as claimed. Moreover, the number of generations chosen in the sum over ‘rest of k ’ in E'_N is exactly

$$S_{i-1} + K + (N - S_i) \leq S_{i-1} + K + m_{i+1},$$

and so by $N \geq S_i$ and $M = \|f\|_\infty$, we obtain the claim

$$\int f dE'_N \leq M \cdot \frac{S_{i-1} + K + m_{i+1}}{N} \leq M \cdot \frac{S_{i-1} + K + m_{i+1}}{S_i}. \quad \square$$

6. Generic FDs

We now establish Theorem D, stated more precisely as follows.

THEOREM 6.1. *For a Baire generic measure $\mu \in \mathcal{M}$ (where in \mathcal{M} we are considering the weak topology), it holds that*

$$\mathcal{TD}(\mu, x) = \mathcal{FD} \quad \text{for } \mu \text{ almost all } x.$$

As usual, we prove first an auxiliary result for CPDs. This result is a consequence of the existence of measures that CP-generate a given CPD.

PROPOSITION 6.2. *For a Baire generic $\mu \in \mathcal{M}^\square$, the set of micromasure distributions*

$$\mathcal{MD}(\mu, x) \supset \mathcal{CPD} \quad \text{at } \mu \text{ almost every } x \in B_1.$$

In particular, if μ is supported on $B(0, \frac{1}{2})$, then for Lebesgue almost every $\omega \in B(0, \frac{1}{2})$, we have

$$\mathcal{MD}(\mu + \omega, x + \omega) = \mathcal{CPD} \quad \text{at } \mu \text{ almost every } x.$$

The idea of the proof is to choose a suitable countable dense subset $\mathcal{S} \subset \mathcal{CPD}$ and prove that a Baire generic measure μ has \mathcal{S} as a subset of $\mathcal{MD}(\mu, x)$ at a μ typical x . Then the closedness of $\mathcal{MD}(\mu, x)$ guarantees the claim. By the countability of \mathcal{S} , we need just to verify the claim for a fixed $Q \in \mathcal{S}$, as the countable intersection of Baire generic sets is Baire generic. This in turn can be obtained by proving that the property of being close to a measure μ whose CP scenery distribution $\langle \mu, x \rangle_N$ is weakly close to Q in a set of large μ measure, is an open and dense property.

To deal with openness of measure-theoretical properties using Euclidean balls, the dense subset of CPDs must consist of measures which give zero mass to all boundaries of dyadic cubes. This is guaranteed by the Lebesgue intensity properties of CPDs.

DEFINITION 6.3. We define

$$B_1^\circ := B_1 \setminus \bigcup_{N=1}^{\infty} \bigcup_{D \in \mathcal{D}_N} \partial D.$$

That is, we remove the boundaries of dyadic cubes of all levels from B_1 . Further, we let

$$\mathcal{F} = \{\mu \in \mathcal{M}^\square : \mu(B_1^\circ) = 1\}.$$

LEMMA 6.4. *There exists a countable dense subset $\mathcal{S} \subset \mathcal{CPD}$ such that for any $Q \in \mathcal{S}$ there exists a measure $\mu \in \mathcal{F}$ that CP-generates Q .*

Proof. Since $\mathcal{ECPD} \subset \mathcal{CPD}$ is dense by the Poulsen property (Proposition 4.2) and \mathcal{CPD} is separable, there exists a countable subset $\mathcal{S} \subset \mathcal{ECPD}$ which is dense in \mathcal{CPD} . Each CPD Q has Lebesgue intensity measure, so $[Q](B_1^\circ) = 1$. This yields that Q almost every μ gives mass $\mu(B_1^\circ) = 1$. On the other hand, by the ergodic theorem, if Q is ergodic, then Q almost every μ CP-generates Q , so we in particular fix one such μ with $\mu(B_1^\circ) = 1$. This yields the desired set \mathcal{S} . \square

We state a simple lemma which explains why we work with measures in \mathcal{F} .

LEMMA 6.5. *For fixed $y \in B_1$ and $N \in \mathbb{N}$, the map $\nu \rightarrow \overline{\langle \nu, y \rangle}_N$ is continuous at all $\mu \in \mathcal{F}$ for which $\overline{\langle \mu, y \rangle}_N$ is defined.*

Proof. The claim follows if we can show that $\nu \rightarrow \nu(\cdot|y_0^j)$ is, for each fixed j , continuous at elements of \mathcal{F} for which $\nu(\cdot|y_0^j)$ is defined. Write $z = y_0^j$. Recall that by definition, $\mu(\cdot|z) = T_{B_z} \mu_{B_z}$. Fix a sequence (μ_n) with $\mu_n \rightarrow \mu$. Since $\mu \in \mathcal{F}$, we have $\mu(\partial B_z) = 0$, so by the weak convergence of measures (see, for example, Lemma 2.29 applied to $f = \chi_{B_z}$), we have $\mu_n(B_z) \rightarrow \mu(B_z)$. Hence the measures $(\mu_n)_{B_z} \rightarrow \mu_{B_z}$ as $n \rightarrow \infty$. The homothety map $T_{B_z} : B_1 \rightarrow B_1$ is continuous, so the continuity follows. \square

LEMMA 6.6. *Given any non-empty open set $\mathcal{O} \subset \mathcal{P}(\mathcal{M}^\square)$, $N \in \mathbb{N}$ and $c > 0$, let*

$$\mathcal{U} := \{\mu \in \mathcal{M}^\square : \mu(\{x : \overline{\langle \mu, x \rangle}_N \in \mathcal{O}\}) > c\}.$$

Then $\mathcal{F} \cap \mathcal{U}$ is contained in the interior of \mathcal{U} .

Proof. For any $\nu \in \mathcal{M}$, let

$$A_\nu = \{x \in B_1^\circ : \overline{\langle \nu, x \rangle}_N \in \mathcal{O}\}.$$

Fix $\mu \in \mathcal{F} \cap \mathcal{U}$, that is, $\mu(B_1^\circ) = 1$ and $\mu(A_\mu) > c$.

We claim that there is an open set $\mathcal{V} \ni \mu$ such $A_\mu \subset A_\nu$ for all $\nu \in \mathcal{V}$. Indeed, note that $\overline{\langle \nu, x \rangle}_N$ depends only on the dyadic cube of level N which contains x . Now let Z denote the (finite) collection of centers of dyadic cubes whose interior is contained in A_μ . Then $A_\mu \subset A_\nu$ whenever

$$\nu \in \bigcap_{z \in Z} \{\eta : \overline{\langle \eta, z \rangle}_N \in \mathcal{O}\},$$

which contains a neighborhood of μ by Lemma 6.5. This is where we used the fact that $\mu \in \mathcal{F}$.

Note that A_μ is open (as a union of open dyadic cubes). Since $\mu(A_\mu) > c$ and μ is Radon, there is a compact subset $K \subset A_\mu$ such that $\mu(K) > c$. If we let $\mathcal{W} = \{\nu \in \mathcal{M} : \nu(K) > c\}$, then \mathcal{W} is open: indeed, if η_n are in its complement and $\eta_n \rightarrow \eta$ then, using [17, Theorem 1.24],

$$\eta(K) \leq \liminf_{n \rightarrow \infty} \eta_n(K) \leq c,$$

so the complement of \mathcal{W} is closed.

We have seen that the set $\mathcal{V} \cap \mathcal{W}$ contains μ , is open and is contained in \mathcal{U} . Hence μ is in the interior of \mathcal{U} , as claimed. \square

LEMMA 6.7. *For any non-empty open sets $\mathcal{O} \subset \mathcal{P}(\mathcal{M}^\square)$ and $\mathcal{B} \subset \mathcal{M}^\square$, and any $\varepsilon > 0$, there exist $N_0 \in \mathbb{N}$ and a measure $\mu \in \mathcal{B} \cap \mathcal{F}$ such that*

$$\mu(\{x : \overline{\langle \mu, x \rangle}_N \in \mathcal{O} \text{ for all } N \geq N_0\}) > 1 - \varepsilon.$$

Proof. Recall the collection $\mathcal{S} \subset \mathcal{CPD}$ from Lemma 6.4. Since \mathcal{S} is dense, we can find a CPD $Q \in \mathcal{O} \cap \mathcal{S}$. By the definition of \mathcal{S} , there exists a measure $\nu \in \mathcal{F}$ that CP-generates Q . Pick $\tau \in \mathcal{B}$. Following the terminology of Hochman [11, Section 8.2], define the (ν, n) -discretization of τ by

$$\tau_n = \sum_{D \in D_n} \tau(D) T_D^{-1} \nu.$$

This is very similar to splicing, except that we use the measure τ for the first n dyadic generations and the measure ν for all the others. Since ν is a probability measure, we obtain $\tau_n \rightarrow \tau$ so there exists $n = n_\delta \in \mathbb{N}$ such that $\tau_n \in \mathcal{B}$. Let μ be this discretization τ_n . Note that $\mu \in \mathcal{F}$ since $\nu \in \mathcal{F}$.

Recall the symbolic coding of dyadic cubes introduced in Notation 4.3. It follows from the definition of μ that for any $x \in B_1$ such that $\tau(x_0^n) > 0$ and any $j \in \mathbb{N}$,

$$\mu(\cdot | x_0^{n+j}) = \nu(\cdot | x_n^{n+j}). \quad (6.1)$$

We employ the metric on measures d introduced in Subsection 2.6. Pick $\delta > 0$ such that $B_d(\overline{Q}, 2\delta) \subset \mathcal{O}$, where B_d denotes the open ball in this metric. Since ν CP-generates Q , there exists $m \in \mathbb{N}$ such that $\nu(U_\delta) > 1 - \varepsilon$ for

$$U_\delta = \{y \in B_1 : d(\overline{\langle \nu, y \rangle}_N, \overline{Q}) < \delta \text{ for all } N \geq m\}.$$

Now if $N \geq n$, $f: \mathcal{P}(\mathcal{M}^\square) \rightarrow \mathbb{R}$ is 1-Lipschitz, $x \in B_1$ is such that $\nu(x_0^N) > 0$, then we deduce from (6.1) that

$$\left| \int f d\overline{\langle \mu, x \rangle}_N - \int f d\overline{\langle \nu, x_n^\infty \rangle}_N \right| = \frac{1}{N} \left| \sum_{j=0}^{n-1} \int f d\mu(\cdot | x_0^j) - \sum_{j=N}^{N+n-1} \int f d\nu(\cdot | x_0^j) \right| \leq \frac{2n\|f\|_\infty}{N}.$$

Recall from Remark 2.28 that in the definition of the metric d we may restrict ourselves to 1-Lipschitz functions f with $\|f\|_\infty \leq C$ for some fixed constant $C > 0$. By taking m larger if needed, we may further assume $2nC/m < \delta$. The above calculation then shows

$$d(\overline{\langle \mu, x \rangle}_N, \overline{\langle \nu, x_n^\infty \rangle}_N) < \delta \quad \text{whenever } N \geq m.$$

Hence, if we define

$$V_\delta = \{x \in B_1 : x_n^\infty \in U_\delta\},$$

then we have $\overline{\langle \mu, x \rangle}_N \in \mathcal{O}$ for all $N \geq m$. Furthermore, by the definition of μ ,

$$\mu(V_\delta) = \sum_{D \in D_n} \tau(D) \nu(T_D V_\delta) \geq \sum_{D \in D_n} \tau(D) \nu(U_\delta) > 1 - \varepsilon.$$

This concludes the proof. \square

We can now conclude the proof of Proposition 6.2.

Proof of Proposition 6.2. Let $\mathcal{S} \subset \mathcal{CPD}$ be the countable dense subset from Lemma 6.4. For $Q \in \mathcal{S}$, $\varepsilon > 0$ and $K \in \mathbb{N}$, we define

$$\mathcal{U}_{Q,\varepsilon,K} := \bigcup_{N \geq K} \text{interior}\{\mu \in \mathcal{M}^\square : \mu(\{x : d(\overline{\langle \mu, x \rangle}_N, \overline{Q}) < \varepsilon\}) > 1 - \varepsilon\}$$

and

$$\mathcal{R} := \bigcap_{Q \in \mathcal{S}} \bigcap_{\varepsilon \in \mathbb{Q}, 0 < \varepsilon < 1} \bigcap_{K \in \mathbb{N}} \mathcal{U}_{Q,\varepsilon,K}.$$

The collection $\mathcal{U}_{Q,\varepsilon,K}$ is open as a union of open sets over $N \geq K$. Moreover, by Lemma 6.6 applied to the open set $\mathcal{O} = B_d(\overline{Q}, \varepsilon) \subset \mathcal{P}(\mathcal{M}^\square)$, the collection $\mathcal{U}_{Q,\varepsilon,K}$ contains the set

$$\mathcal{D}_{Q,\varepsilon,K} = \bigcup_{N \geq K} \{\mu \in \mathcal{F} : \mu(\{x : d(\overline{\langle \mu, x \rangle}_N, \overline{Q}) < \varepsilon\}) > 1 - \varepsilon\}.$$

By Lemma 6.7, the set $\mathcal{D}_{Q,\varepsilon,K}$ is dense: for a given non-empty open set $\mathcal{B} \subset \mathcal{M}^\square$, we can choose $\mu \in \mathcal{F} \cap \mathcal{B}$ such that $d(\overline{\langle \mu, x \rangle}_N, \overline{Q}) < \varepsilon$ happens for all $N \geq N_0$ in a set of μ measure greater than $1 - \varepsilon$. In particular, we can find $N \geq K$ with this property, showing that \mathcal{B} meets $\mathcal{D}_{Q,\varepsilon,K}$. Hence $\mathcal{D}_{Q,\varepsilon,K}$ is dense as claimed, and so is $\mathcal{U}_{Q,\varepsilon,K}$.

Since the set \mathcal{R} is a countable intersection of sets with dense interiors, its complement is a set of first category. Fix $\mu \in \mathcal{R}$ and let $\{Q_j\}$ be an enumeration of \mathcal{S} . Let $\varepsilon_i \searrow 0$ be a sequence of rational numbers. Then, for some $N_{i,j} \nearrow \infty$, we obtain $\mu(A_{i,j}) \geq 1 - \varepsilon_i$ for

$$A_{i,j} := \{x \in B_1 : d(\overline{\langle \mu, x \rangle}_{N_{i,j}}, \overline{Q_j}) < \varepsilon_i\}.$$

Write

$$A = \bigcap_{j \in \mathbb{N}} \limsup_{i \rightarrow \infty} A_{i,j}.$$

Then as $\varepsilon_i \searrow 0$, we have by the convergence of measures that $\mu(A) = 1$. Fix $x \in A$. For each j there are infinitely many i such that

$$d(\overline{\langle \mu, x \rangle}_{N_{i,j}}, \overline{Q_j}) < \varepsilon_i.$$

This shows $\mathcal{S} \subset \mathcal{MD}(\mu, x)$. Since $\mathcal{MD}(\mu, x)$ is closed in $\mathcal{P}(\Xi)$ we have $\mathcal{CPD} \subset \mathcal{MD}(\mu, x)$ at μ almost every x .

The second statement is immediate from the fact that $\mathcal{MD}(\mu + \omega, x + \omega) \subset \mathcal{CPD}$ for Lebesgue almost all ω (recall Theorem 2.21). \square

Proof of Theorem 6.1. Given $Q \in \mathcal{S}, \varepsilon > 0$ and $K \in \mathbb{N}$, let

$$\tilde{\mathcal{U}}_{Q,\varepsilon,K} = \text{interior}\{\mu \in \mathcal{M} : \mu^\square \in \mathcal{U}_{Q,\varepsilon,K}\},$$

where $\mathcal{U}_{Q,\varepsilon,K}$ is as in the proof of Proposition 6.2. These sets are open by definition. We claim they are also dense.

Let $\pi: \mathcal{M} \rightarrow \mathcal{M}^\square$ be the map $\mu \rightarrow \mu^\square$. This map is not continuous (and is not everywhere defined), but it is defined and continuous on the set $\mathcal{D} \subset \mathcal{M}$ of measures which give zero mass to the boundary of B_1 and positive mass to B_1 . It is easy to see that \mathcal{D} is in fact dense, and \mathcal{D}^\square is dense in \mathcal{M}^\square (where, as usual, $\mathcal{D}^\square = \{\mu^\square : \mu \in \mathcal{D}\}$). Hence, since $\mathcal{U}_{Q,\varepsilon,K}$ is open and dense, $\mathcal{D}^\square \cap \mathcal{U}_{Q,\varepsilon,K}$ is dense in \mathcal{M}^\square . Since measures in \mathcal{D} are continuity points of π , the set $\pi^{-1}(\mathcal{D}^\square \cap \mathcal{U}_{Q,\varepsilon,K})$ is contained in the interior of $\pi^{-1}(\mathcal{U}_{Q,\varepsilon,K})$. We see that it is also dense in \mathcal{M} .

Let D be a metric on \mathcal{M} inducing the weak topology (see, for example, [17, Remark 14.15] for an instance of such a metric). Fix $\tau \in \mathcal{M}$ and $\varepsilon > 0$. Pick $\nu \in \mathcal{D}$ such that $D(\nu, \tau) < \varepsilon/2$. Since $\nu \in \mathcal{D}$, $\varrho := \nu(B_1) > 0$ so in particular ν^\square is well defined. Now pick a sequence $\mu_n \rightarrow \nu^\square$ with $\mu_n \in \mathcal{D}^\square \cap \mathcal{U}_{Q,\varepsilon,K}$; this is possible by density. Even though μ_n is a measure on B_1 , we

identify it with a measure on \mathbb{R}^d . Let

$$\nu_n = \nu|_{\mathbb{R}^d \setminus B_1} + \varrho \mu_n.$$

Then it is easy to see that $\nu_n \rightarrow \nu$ weakly, so we have $D(\nu_n, \tau) < \varepsilon$ for some n . By construction $\nu_n^\square = \mu_n$ and therefore $\nu_n \in \pi^{-1}(\mathcal{D}^\square \cap \mathcal{U}_{Q, \varepsilon, K})$, showing that this set, and hence also $\tilde{\mathcal{U}}_{Q, \varepsilon, K}$, is dense as claimed.

Now, similarly to the proof of Proposition 6.2, define

$$\mathcal{R}_0 = \bigcap_{Q \in \mathcal{S}} \bigcap_{\varepsilon \in \mathbb{Q}, 0 < \varepsilon < 1} \bigcap_{K \in \mathbb{N}} \tilde{\mathcal{U}}_{Q, \varepsilon, K}.$$

Fix $P \in \mathcal{FD}$. Theorem 3.3 implies that there exists a CPD Q such that

$$P = \text{cent}(\hat{Q})^\square.$$

Since $\mu \in \mathcal{R}_0$, the proof of Proposition 6.2 implies $\mathcal{CPD} \subset \mathcal{MD}(\mu, x)$ at μ almost every $x \in B_1$, so in particular $Q \in \mathcal{MD}(\mu, x)$ at these x . Then, by Proposition 5.3, we know $P \in \mathcal{TD}(\mu, x)$ at μ almost every $x \in B_1$. We also know that at μ almost every $x \in B_1$ we have $\mathcal{TD}(\mu, x) \subset \mathcal{FD}$ by Theorem 2.10. We have shown $\mathcal{TD}(\mu, x) = \mathcal{FD}$ for any $\mu \in \mathcal{R}_0$ and μ almost all $x \in B_1$.

To finish the proof, define

$$\mathcal{R} = \bigcap_{\mathbf{n} \in \mathbb{Z}^d} \{T_{\mathbf{n}}\mu : \mu \in \mathcal{R}_0\},$$

that is, we intersect all integer translates of \mathcal{R}_0 . This set is a countable intersection of open and dense sets (note that for fixed \mathbf{n} , the map $T_{\mathbf{n}}$ is a homeomorphism of \mathcal{M}) and thus its complement is a set of first category. Moreover, if $\mu \in \mathcal{R}$, then for all integer vectors \mathbf{n} , $\mathcal{TD}(\mu, x) = \mathcal{FD}$ for μ almost all $x \in B(\mathbf{n}, 1)$. As the union of these balls covers \mathbb{R}^d , we are done. \square

Appendix. Remarks on the choice of norm

We have so far followed [11] in working with the L^∞ norm of \mathbb{R}^d . However, other than simplifying some proofs, there is nothing special about this norm, and all the results from [11] that we require work equally well with any other norm. Since in geometric measure theory one uses mainly the Euclidean norm, our geometric applications in [15] are simplified if we use Euclidean versions of the results presented here.

Rather than verifying that all the proofs from [11] can be made to work with any norm, we explain how to deduce the results from their L^∞ version. This is most clear for the extended version of FDs, since in this case choosing different norms amounts only to a different choice of normalization; see the discussion in [11, Section 3.1]. For restricted versions things also work, helped by the uniqueness of the extended version.

For the sake of completeness, we provide details of the modifications needed to carry the theory to the setting of an arbitrary norm. Note that the concepts of \mathcal{FD} and \mathcal{TD} are effectively tied to the L^∞ norm $\|\cdot\|$ which was implicit in their definition. Let $\|\cdot\|'$ be any other norm on \mathbb{R}^d . We use an apostrophe to denote the corresponding concepts defined in terms of this new norm. For example, B'_1 is the unit ball in the norm $\|\cdot\|'$, the set of FDs with respect to this norm is \mathcal{FD}' , etc. Further, for $\mu \in \mathcal{M}$ with $0 \in \text{supp } \mu$, we denote $\mu' = (1/\mu(B'_1))\mu$ and $\mu^\diamond = \mu_{B'_1}$. As with μ^*, μ^\square , we use the same notation to indicate postcomposition with these maps, as in Q', Q^\diamond , etc.

The proof of the one–one correspondence $Q \rightarrow Q^\diamond$ between restricted and extended versions [11, Lemma 3.1] is independent of the chosen norm. Thus we assume this.

PROPOSITION A.1. (A.1) The map $\mu \rightarrow \mu'$ is an isomorphism (that is, a bijective factor map) from (\mathcal{M}^*, S^*) to (\mathcal{M}', S') .

(A.2) If $Q \in \mathcal{FD}$, then $Q' \in \mathcal{FD}'$ and conversely for any $P \in \mathcal{FD}'$ there is $Q \in \mathcal{FD}$ with $Q' = P$.

(A.3) Theorem 2.10 continues to hold with the norm $\|\cdot\|'$, that is, if $\mu \in \mathcal{M}$, then for μ almost all x , $\mathcal{TD}'(\mu, x) \subset \mathcal{FD}'$.

(A.4) Theorem 2.11 holds for the norm $\|\cdot\|'$, that is, if $Q \in \mathcal{FD}'$, then the ergodic components of Q are also in \mathcal{FD}' .

(A.5) Theorem 3.3 holds for the norm $\|\cdot\|'$, that is, if Q is a CPD, then $\text{cent}'(Q) \in \mathcal{FD}'$ and, conversely, given any $P \in \mathcal{FD}'$, there is an extended CPD Q with $\text{cent}'(Q) = P$. Here $\text{cent}'(Q)$ is the push-down of $Q \times \lambda$ under $((\mu, x), t) \rightarrow S_t' T_x \mu$ (where λ is normalized Lebesgue measure on $[0, \log 2)$).

Proof. (A.1) This is routine from the definitions.

(A.2) Given the first part, we need to show only that if $Q \in \mathcal{FD}$, then Q' is quasi-Palm. Suppose $Q'(Y) = 1$. Since $\mu \rightarrow \mu'$ is a bijection, $Y = Z'$, where $Q(Z) = 1$. By the quasi-Palm property of Q , $\mu_{x,t} \in Z$ for Q almost all μ almost all x . But then $T_x \mu' \in Z'$ for Q almost all μ and μ almost all x , and whence also for Q' almost all μ' and μ' almost all x , since μ and μ' are equivalent.

(A.3) Fix $\mu \in \mathcal{M}$, and let x be a point such that $\mathcal{TD}(\mu, x) \subset \mathcal{FD}$; by Theorem 2.10, we know this happens for μ almost all x .

Assume first that the unit ball of $\|\cdot\|'$ is contained in B_1 . In this case, we have $S_t^\diamond T_x \mu = (S_t^\circ T_x \mu)^\diamond$. Suppose $\langle \mu \rangle'_{x, T_j} \rightarrow Q$; by passing to a subsequence we may assume $\langle \mu \rangle_{x, T_j} \rightarrow P$, where $P \in \mathcal{FD}$ by our choice of x . Then, writing \hat{P} for the extended version of P , we have $(\hat{P}')^\diamond = Q$; by the previous part, \hat{P}' is in \mathcal{FD}' and by uniqueness is therefore the extended version of Q , showing $Q \in \mathcal{FD}'$ as claimed. Note that this argument works in reverse: if $\mathcal{TD}'(\mu, x) \in \mathcal{FD}'$, then $\mathcal{TD}(\mu, x) \in \mathcal{FD}$.

The general case now follows by considering first the norm $\|\cdot\|'' = \max(\|\cdot\|, \|\cdot\|')$ (whose unit ball is contained in B_1) and then using the result for this norm to deduce the same for $\|\cdot\|'$ (which has a larger unit ball).

(A.4) This assertion follows from the first two: we know from Theorem 2.11 that if $Q = \int Q_\mu dQ(\mu)$ is the ergodic decomposition of Q , then $Q_\mu \in \mathcal{FD}$ for Q almost all μ . By the first part, the ergodic decomposition of Q' is $\int Q'_\mu dQ'(\mu)$, and by the second part $Q'_\mu \in \mathcal{FD}'$ for Q' almost all μ .

(A.5) Note that $\text{cent}'(Q) = (\text{cent}(Q))'$ so both statements are a consequence of the result for the L^∞ norm and the first part. \square

Now it becomes clear that Theorems 3.1, 4.1, 5.1 and 6.1 hold for any norm and in particular for the Euclidean norm. This can be deduced either from the fact that our proofs are norm-independent (once we have norm-independent versions of the results we assume), or by following simple arguments of the kind above to pass from the L^∞ norm to an arbitrary norm.

Acknowledgements. We thank M. Hochman for many enlightening discussions on topics related to this paper.

References

1. C. BANDT and A. KÄENMÄKI, ‘Local structure of self-affine sets’, *Ergodic Theory Dynam. Systems* 33 (2013) 1326–1337.
2. F. BAYART, ‘The multifractal box dimensions of typical measures’, *Fund. Math.* 219 (2012) 145–162.
3. F. BAYART, ‘Multifractal spectra of typical and prevalent measures’, *Nonlinearity* 26 (2013) 353–367.

4. P. BILLINGSLEY, *Convergence of probability measures* (John Wiley & Sons, New York, 1968).
5. Z. BUCZOLICH and S. SEURET, ‘Typical Borel measures on $[0, 1]^d$ satisfy a multifractal formalism’, *Nonlinearity* 23 (2010) 2905–2918.
6. M. EINSIEDLER and T. WARD, *Ergodic theory with a view towards number theory*, Graduate Texts in Mathematics 259 (Springer, London, 2011).
7. A. FERGUSON, J. FRASER and T. SAHLSTEN, ‘Scaling scenery of $(\times m, \times n)$ invariant measures’, *Adv. Math.* 268 (2015) 564–602.
8. H. FURSTENBERG, ‘Intersections of Cantor sets and transversality of semigroups’, *Problems in analysis (Symposium Salomon Bochner, Princeton University, Princeton, NJ, 1969)* (Princeton University Press, Princeton, NJ, 1970) 41–59.
9. H. FURSTENBERG, ‘Ergodic fractal measures and dimension conservation’, *Ergodic Theory Dynam. Systems* 28 (2008) 405–422.
10. M. GAVISH, ‘Measures with uniform scaling scenery’, *Ergodic Theory Dynam. Systems* 31 (2011) 33–48.
11. M. HOCHMAN, ‘Dynamics on fractals and fractal distributions’, Preprint, 2010, arXiv:1008.3731.
12. M. HOCHMAN, ‘Geometric rigidity of $\times m$ invariant measures’, *J. Eur. Math. Soc.* 14 (2012) 1539–1563.
13. M. HOCHMAN and P. SHMERKIN, ‘Local entropy averages and projections of fractal measures’, *Ann. of Math.* (2), 175 (2012) no. 3, 1001–1059.
14. M. HOCHMAN and P. SHMERKIN, ‘Equidistribution from fractal measures’, *Invent. Math.*, 2013, arXiv:1302.5792.
15. A. KÄENMÄKI, T. SAHLSTEN and P. SHMERKIN, ‘Dynamics of the scenery flow and geometry of measures’, *Proc. Lond. Math. Soc.*, 2013, arXiv:1401.0231.
16. J. LINDENSTRAUSS, G. OLSEN and Y. STERNFELD, ‘The Poulsen simplex’, *Ann. Inst. Fourier (Grenoble)* 28 (1978) vi, 91–114.
17. P. MATTLA, *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability* (Cambridge University Press, Cambridge, 1995).
18. P. MÖRTERS and D. PREISS, ‘Tangent measure distributions of fractal measures’, *Math. Ann.* 312 (1998) 53–93.
19. T. O’NEIL, ‘A measure with a large set of tangent measures’, *Proc. Amer. Math. Soc.* 123 (1995) 2217–2220.
20. T. ORPONEN, ‘On the distance sets of self-similar sets’, *Nonlinearity* 25 (2012) 1919–1929.
21. D. PREISS, ‘Geometry of measures in \mathbf{R}^n : distribution, rectifiability, and densities’, *Ann. of Math.* (2) 125 (1987) 537–643.
22. T. SAHLSTEN, ‘Tangent measures of typical measures’, *Real Anal. Exchange*, 40 (2014/2015) no. 1, 1–27.
23. J. SCHMELING and P. SHMERKIN, ‘On the dimension of iterated sumsets’, *Recent developments in fractals and related fields*, Applied and Numerical Harmonic Analysis (Birkhäuser, Boston, MA, 2010) 55–72.
24. P. WALTERS, *An introduction to ergodic theory* (Springer, New York–Berlin, 1982).

Antti Käenmäki
 Department of Mathematics and Statistics
 University of Jyväskylä
 P.O. Box 35 (MaD)
 FI-40014 Jyväskylä
 Finland

antti.kaenmaki@jyu.fi

Tuomas Sahlsten
 Einstein Institute of Mathematics
 The Hebrew University of Jerusalem
 Givat Ram
 Jerusalem 91904
 Israel

tuomas@sahlsten.org

Pablo Shmerkin
 Department of Mathematics and Statistics
 Torcuato Di Tella University
 Av. Figueroa Alcorta
 7350 Buenos Aires
 Argentina

pshmerkin@utdt.edu