Self-similar random measures and tangent measure distributions[1]

Peter Mörters[2]
Technische Universität Berlin, Fachbereich Mathematik, Ma7-5
Straße des 17. Juni 136, 10623 Berlin.
E-Mail: peter@mathematik.uni-kl.de

1 Self-similar random measures

Let $\mathcal{M}(\mathbb{R}^d)$ be the space of locally finite positive measures on $\mathbb{R}^d$ equipped with the vague toplogy. To motivate our results I describe in this section a natural definition of self-similarity for random elements of $\mathcal{M}(\mathbb{R}^d)$, which is due to U. Zähle (1988).

The heuristic idea of self-similarity of a random measure is the following: if an observer looks at the measure through a circular window and renormalizes size and mass of his observation, he cannot decide from the distribution of this renormalized observation about the original size and position of the window. A first idea to define self-similarity of a random measure therefore suggests that the distribution of the random measure should be invariant with respect to the group of rescaling operators

$$S_\lambda^\alpha : \mathcal{M}(\mathbb{R}^d) \to \mathcal{M}(\mathbb{R}^d)\ ,\ S_\lambda^\alpha \nu(E) = \frac{\nu(\lambda E)}{\lambda^\alpha}\ ,$$

where $\alpha \geq 0$ is a fixed renormalization parameter, which should be linked to the dimension of the random measure, and with respect to the group of shift operators

$$T^u : \mathcal{M}(\mathbb{R}^d) \to \mathcal{M}(\mathbb{R}^d)\ ,\ T^u\nu(E) = \nu(u+E)\ .$$

Unfortunately, it turns out that there are no nontrivial examples of probability distributions on $\mathcal{M}(\mathbb{R}^d)$ that are invariant with respect to these operators. To obtain a meaningful definition one has to use non-finite distributions on $\mathcal{M}(\mathbb{R}^d)$.

Definition A σ-finite distribution $\mathbb{Q}$ on $\mathcal{M}(\mathbb{R}^d)$ with locally finite intensity measure $\Lambda = \int \nu\, \mathbb{Q}(d\nu)$ is called *α-scale invariant* if

- for every vector $u \in \mathbb{R}^d$ we have $\mathbb{Q} = \mathbb{Q} \circ T^u$,
- for every $\lambda > 0$ there is $\gamma(\lambda) > 0$ such that $\mathbb{Q} = \gamma(\lambda)\mathbb{Q} \circ S_\lambda^\alpha$.

Then we automatically have $\gamma(\lambda) = \lambda^{d-\alpha}$.

Let us look at an **example**: Let $\{B_t\}$ be a two-sided Brownian motion with $B_0 = 0$ and L_t^x a continuous version of the local time at level x. As L_t^x is an increasing function in t

[1]This is the outline of an invited talk at the 5th Workshop on Probability Theory and its Applications, Bielefeld, February 1999. It is based on a joint research project with David Preiss, London, which was supported by the ARC-Programme of British Council and DAAD.

[2]Currently on leave from: Fachbereich Mathematik, Universität Kaiserslautern, 67663 Kaiserslautern.

it can be interpreted as a locally finite random measure on the real line, supported by the set $\{t : B_t = x\}$. This measure is the local time measure at level x. A 1/2-scale invariant distribution $\mathbb{Q}$ may defined by $\mathbb{Q}(M) = \mathbb{E}\int_{-\infty}^{\infty} 1_M(L^x)\,dx$ for Borel sets $M \subset \mathcal{M}(\mathbb{R}^d)$. In other words, choose a level x according to Lebesgue measure, independently from the motion, and study the local time measure at the random level x.

The drawback of the definition above is of course the non-finiteness of the distributions $\mathbb{Q}$. One can, however, go back to probability distributions by conditioning: Heuristically, if we choose a measure X according to the distribution $\mathbb{Q}$ and, given X, choose a point x according to X, then x is Λ-distributed and we may ask for the distribution of the measure X given the point x. More formally, disintegration yields the existence of a family $\{\mathbb{P}^x : x \in \mathbb{R}^d\}$ of probability distributions on $\mathcal{M}(\mathbb{R}^d)$ such that for all $M \subset \mathcal{M}(\mathbb{R}^d)$ and $A \subset \mathbb{R}^d$ Borel,

$$\int_M X(A)\,\mathbb{Q}(dX) = \int_A \mathbb{P}^x(M)\,\Lambda(dx)\,.$$

Using the translation invariance of Q we observe that Λ is a constant multiple of Lebesgue measure and, more importantly, that there exists a probability distribution $\mathbb{P}$ on $\mathcal{M}(\mathbb{R}^d)$ such that

$$\mathbb{P}^x = \mathbb{P} \circ T^x \text{ for Lebesgue-almost all } x\,.$$

$\mathbb{P}$ is called the *Palm distribution associated with* $\mathbb{Q}$. A direct formula for $\mathbb{P}$ is easy to deduce: Choose a Borel set $B \subset \mathbb{R}^d$ of positive and finite Lebesgue measure, then

$$\mathbb{P}(M) = \frac{1}{\Lambda(B)} \int\int_B 1_M(T^x X)\,X(dx)\,\mathbb{Q}(dX) \text{ for } M \subset \mathcal{M}(\mathbb{R}^d) \text{ Borel}\,.$$

Palm distributions appear in many areas of probability, for example in queuing theory or in the theory of point processes, where they play the rôle of conditional distributions of stationary point processes given a point at the origin. The heuristics in our case is that Palm distributions are probability distributions on $\mathcal{M}(\mathbb{R}^d)$ that have the origin as a typical point of their samples.

The Palm distribution $\mathbb{P}$, being the probabilistic object, is the natural object to study and hence we have to ask, whether the α-scale invariance of the underlying distribution $\mathbb{Q}$ can be checked for $\mathbb{P}$ without reference to $\mathbb{Q}$ itself.

Theorem 1.1 *(i) (Mecke (1967)) A probability distribution $\mathbb{P}$ on $\mathcal{M}(\mathbb{R}^d)$ is the Palm distribution of a translation invariant distribution $\mathbb{Q}$ with nonzero intensity measure $\Lambda \in \mathcal{M}(\mathbb{R}^d)$ if and only if $\mathbb{P}$ does not charge the zero measure and the Palm formula*

$$\int\int G(\nu,u)\,\nu(du)\,\mathbb{P}(d\nu) = \int\int G(T^u\nu,-u)\,\nu(du)\,\mathbb{P}(d\nu) \tag{1}$$

holds for all Borel functions $G : \mathcal{M}(\mathbb{R}^d) \times \mathbb{R}^d \longrightarrow [0,\infty)$.

(ii) (U. Zähle (1988)) $\mathbb{P}$ is the Palm distribution of an α-scale invariant measure $\mathbb{Q}$ if and only if in addition to the above we have $\mathbb{P} = \mathbb{P} \circ S^\alpha_\lambda$ for every $\lambda > 0$.

(iii) (U. Zähle (1988)) If $\mathbb{P}$ is the Palm distribution of an α-scale invariant measure $\mathbb{Q}$ with non-zero intensity measure, then α is the carrying dimension of $\mathbb{P}$-almost every measure.

The following definition is now natural, see U. Zähle (1988).

Definition A probability distribution $\mathbb{P}$ on $\mathcal{M}(\mathbb{R}^d)$ is the distribution of an *α-self similar random measure* if

(i) $\mathbb{P}$ satisfies the Palm formula (1) for all Borel functions $G : \mathcal{M}(\mathbb{R}^d) \times \mathbb{R}^d \longrightarrow [0, \infty)$,

(ii) $\mathbb{P}$ satisfies $\mathbb{P} = \mathbb{P} \circ S^\alpha_\lambda$ for every $\lambda > 0$.

In our **example** it is easy to check that the Palm distribution $\mathbb{P}$ associated with $\mathbb{Q}$ is just the distribution of the local time measure L^0 at level 0 and hence L^0 is a (1/2)-self-similar random measure.

Let us finally see, how deterministc self-similar sets may be modelled in our framework with a very sparse randomization. Recall the definition of a deterministic self-similar set: Let $S_1, \dots, S_n$ be a family of contracting similtudes with contracion ratios $r_1, \dots, r_n$ and define α to be the unique solution of $\sum_{i=1}^n r_i^\alpha = 1$. Then there is a unique nonzero compact set K such that

$$K = S_1(K) \cup \ldots \cup S_n(K) \, .$$

We are assuming that there is no overlap of the different parts $S_i(K)$, so that the so-called open-set condition holds. Then α is the Hausdorff dimension of K and a canonical probability measure μ is defined on $K \subset \mathbb{R}^d$ by

$$\mu\Big(S_{x_1} \circ \cdots \circ S_{x_k}(K)\Big) = r^\alpha_{x_1} \cdots r^\alpha_{x_k} \text{ for all } x_1, \dots, x_k \in \{1, \dots, n\} \, .$$

Definition Let $\mu \in \mathcal{M}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$. Define the family of measures $\{\mu_{x,r} \, : \, r > 0\} \subseteq \mathcal{M}(\mathbb{R}^d)$, the *enlargements of μ about x*, by $\mu_{x,r}(A) = \mu(x + rA)$ for all Borel sets $A \subset \mathbb{R}^d$. For every $x \in \mathbb{R}^d$ and $\alpha \geq 0$ we define probability distributions $\mathbb{P}_\varepsilon(x, \alpha) = \mathbb{P}_\varepsilon$ on $\mathcal{M}(\mathbb{R}^d)$ by

$$\mathbb{P}_\varepsilon(M) = \frac{1}{\log(1/\varepsilon)} \int_\varepsilon^1 1_M\Big(\frac{\mu_{x,r}}{r^\alpha}\Big) \, \frac{dr}{r} \quad \text{for Borel sets } M \subset \mathcal{M}(\mathbb{R}^d).$$

Note that the choice of the random scales according to dr/r is well-adapted to the geometry of self-similar sets. The following theorem is due to Graf (1995).

Theorem 1.2 *Let μ be the canonical measure on a deterministic self-similar set as above and α its dimension. Then, for μ-almost every x, the distributions $\mathbb{P}_\varepsilon$ converge weakly to the distribution $\mathbb{P}$ of an α-self-similar random measure, which is the canonical randomization of μ.*

2 Tangent measure distributions

The definition of the probability distributions $\mathbb{P}_\varepsilon$ makes sense for arbitrary locally finite measures μ. Of course, we cannot generally expect the sequence $\{\mathbb{P}_\varepsilon\}$ to converge as $\varepsilon \downarrow 0$ as in Theorem 1.2, but mild assumptions ensure that for some critical α the family is tight and nontrivial limit points exist. These limit points are called the α-dimensional *tangent measure* *distributions* of μ at x. This concept was first introduced by Christoph Bandt, the word *tangent measure* is justified by an intimate relationship with the measure theoretic concept of tangent measures, see Mattila (1995).

Tangent measure distributions generalize the concept of tangent spaces, indeed if μ is the surface measure on a differentiable α-manifold M, the only nontrivial tangent measure distribution at $x \in M$ is the Dirac distribution concentrated at a constant multiple of Lebesgue measure on the tangent space at x. To make such a generalization valuable, it is necessary that for a large class of measures, the tangent measure distributions retain some of the beautiful properties of tangent spaces. In fact, if we want to include the canonical measure on self-similar sets in our class, by Theorem 1.2, the best property we can hope for is α-self similarity. Very surprisingly, it turns out that such an optimal result holds for *every* measure. The main result of this talk is from Mörters-Preiss (1998).

Theorem 2.1 *Let $\alpha \geq 0$ and let $\mu \in \mathcal{M}(\mathbb{R}^d)$ be any measure. Then at μ-almost all points x every α-dimensional tangent measure distribution is the distribution of an α-self similar random measure.*

In the remainder of this section I give a very rough sketch of parts of the proof of this theorem. The scaling invariance property $\mathbb{P} = \mathbb{P} \circ S_\lambda^\alpha$ follows almost directly from the definition. The surprising fact is, that the limit points $\mathbb{P}$ satisfy the Palm formula (1).

Approximate the left hand side in the Palm formula (1) by

$$\int \int G(\nu, u)\, \nu(du)\, \mathbb{P}_\varepsilon(d\nu) = \frac{1}{\log(1/\varepsilon)} \int_\varepsilon^1 \int G\Big(\frac{\mu_{x,r}}{r^\alpha}, \frac{y-x}{r}\Big)\, \frac{\mu(dy)}{r^\alpha}\, \frac{dr}{r}$$

and the right hand side by

$$\begin{aligned} &\int \int G(T^u \nu, -u)\, \nu(du)\, \mathbb{P}_\varepsilon(d\nu) \\ = \; & \frac{1}{\log(1/\varepsilon)} \int_\varepsilon^1 \int G\Big(T^{\frac{y-x}{r}}\Big(\frac{\mu_{x,r}}{r^\alpha}\Big), \frac{x-y}{r}\Big)\, \frac{\mu(dy)}{r^\alpha}\, \frac{dr}{r} \\ = \; & \frac{1}{\log(1/\varepsilon)} \int_\varepsilon^1 \int G\Big(\frac{\mu_{y,r}}{r^\alpha}, \frac{x-y}{r}\Big)\, \frac{\mu(dy)}{r^\alpha}\, \frac{dr}{r}\,. \end{aligned}$$

Hence, writing $H(x,y,r) = G\big(\frac{\mu_{x,r}}{r^\alpha}, \frac{y-x}{r}\big)$, we have to prove that

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\log(1/\varepsilon)} \int_\varepsilon^1 \int \frac{H(x,y,r) - H(y,x,r)}{r^\alpha}\, \mu(dy)\, \frac{dr}{r} = 0\,.$$

The key step in our proof is an "ergodic theorem", which proves this fact under fairly general conditions on the mapping $H : \mathbb{R}^d \times \mathbb{R}^d \times (0,\infty) \to (0,\infty)$. We hope that this is of some independent interest.

Lemma 2.2 *Whenever a measure $\mu \in \mathcal{M}(\mathbb{R}^d)$, a Borel measurable function $H : \mathbb{R}^d \times \mathbb{R}^d \times (0,\infty) \longrightarrow [0,1]$ and $0 < \alpha < d$ satisfy, for some constants $R, A, B \in (0,\infty)$, the conditions*

(i) $H(x,y;r) = 0$, *whenever* $\|x - y\| \geq Rr$,

(ii) $H(x,y;r)=0$, *whenever* $\mu B(x,5Rr)\geq Ar^\alpha$, *and*

(iii) *for all* $x,\tilde{x},y,\tilde{y}\in\mathbb{R}^d$ *and* $0<r<\infty$,

$$|H(x,y;r)-H(\tilde{x},\tilde{y};r)|\leq\frac{B}{r}\Big(\|x-\tilde{x}\|+\|y-\tilde{y}\|\Big),$$

then, for μ*-almost every* $x\in\mathbb{R}^d$,

$$\lim_{T\uparrow\infty}\frac{1}{T}\int_0^T\int\frac{H(x,y;e^{-t})-H(y,x;e^{-t})}{e^{-t\alpha}}\,\mu(dy)\,dt$$
$$=\lim_{\varepsilon\downarrow 0}\frac{1}{\log(1/\varepsilon)}\int_\varepsilon^1\int\frac{H(x,y;r)-H(y,x;r)}{r^\alpha}\,\mu(dy)\,\frac{dr}{r}=0\,. \tag{2}$$

Note that the statement of the lemma becomes false if we insert any absolute value signs in our integral: the correctness of the statement depends on some quite subtle cancellation properties. To get some flavour of the proof, we illustrate some of its main steps.

We first discretize the problem in a dyadic grid and show that it suffices to prove that

$$\lim_{n\to\infty}\frac{1}{n}\int\sum_{k=0}^{n-1}H_k(x,y)\,d\mu(y)=0\text{ for }\mu\text{-almost all }x\,, \tag{3}$$

where

$$H_k(x,y)=1_{Q_k^x}(y)\frac{H(x,y;2^{-k})-H(y,x;2^{-k})}{2^{-k\alpha}}$$

and Q_k^x is the dyadic cube of sidelength 2^{-k} containing x. An interesting subtlety here is that, in order to avoid a concentration of μ-mass near the boundary of the cubes, we have to include an averaging procedure with respect to shifts of the measure in the proof of the approximation.

To prove the convergence we use an L^2-estimate. Writing, for every x,

$$\varphi_x(E)=\sum_{k=0}^{n-1}\frac{\mathcal{L}^d(E\cap Q_k^x\setminus Q_{k+1}^x)}{\mathcal{L}^d(Q_k^x\setminus Q_{k+1}^x)}\cdot\int_{Q_k^x}H_k(x,y)\,\mu(dy)\,,$$

enables us to express the term in question as

$$\int\Big[\frac{1}{n}\int\sum_{k=0}^{n-1}H_k(x,y)\,\mu(dy)\Big]^2\mu(dx)=\frac{1}{n^2}\int\Big(\varphi_x(\mathbb{R}^d)\Big)^2\mu(dx)\,. \tag{4}$$

The fact that φ_x is a signed measure allows us to use a suitable a partition of $\mathbb{R}^d$ into sets $\{G(I)\}_{I\in\mathcal{I}}$ and write

$$\frac{1}{n^2}\int\Big(\varphi_x(\mathbb{R}^d)\Big)^2\mu(dx)=\frac{1}{n^2}\sum_{I\in\mathcal{I}}\sum_{J\in\mathcal{I}}\int\varphi_x(G(I))\varphi(G(J))\,\mu(dx)\,. \tag{5}$$

The partition is chosen in such a way that the index set I is a family of good cubes (where we have some lower control on the mass) and every set $G(I)$ is the finite union of those dyadic subcubes of I of sidelength $2^{-(n-1)}$, which are not contained in any good strict subcube of I.

The proof now proceeds by giving careful individual estimates for the summands on the right hand side of (5).

We now become even sketchier and pick a typical example of such an individual estimate. Look at the following situation: Assume $J \in \mathcal{I}$ is a large cube and K is a cube, much smaller than J and situated far away from J realtive to its own size. Decompose K in a collection of cubes $I \in \mathcal{I}$ of very small sidelength. Now $\varphi_x(G(I))$ is small if x is far from I relative to its sidelength and, by continuity, $\varphi_x(G(J))$ is almost constant for all $x \in K$. Hence

$$\begin{aligned} & \sum_{I \subset K} \int \varphi_x(G(I))\varphi_x(G(J))\,\mu(dx) \\ \approx & \sum_{I \subset K} \int_K \varphi_x(G(I))\varphi_x(G(J))\,\mu(dx) \\ \approx & \; \varphi_x(G(J)) \int_K \sum_{I \subset K} \varphi_x(G(I))\,\mu(dx) \\ = & \; \varphi_x(G(J)) \int_K \varphi_x(K)\,\mu(dx). \end{aligned}$$

But $\int_K \varphi_x(K)\,\mu(dx)$ is small, as $\int \varphi_x(\mathbb{R}^d)\,\mu(dx) = 0$. This allows a favourable estimate in the present situation. The full proof of Lemma 2.2 can be found in Mörters-Preiss (1998).

3 Further results

Some measure theoretic applications of Theorem 2.1 are given in Mörters-Preiss (1998) and in Mörters (1997). Examples of tangent measure distributions were studied in Graf (1995), Arbeiter-Patzschke (1996), Krieg-Mörters (1998) and, for the case of planar Brownian paths, in Mörters (1998).

References

M. Arbeiter, N. Patzschke. Random self-similar multifractals. *Math. Nachr.*, 181:5–42, 1996.

S. Graf. On Bandt's tangential distribution for self-similar measures. *Mh. Math.* 120:223–246, 1995.

D. Krieg, P. Mörters. Tangent measure distributions of hyperbolic Cantor sets. *Mh. Math.* 126:313-328, 1998.

P. Mattila. *The Geometry of Sets and Measures in Euclidean Spaces.* Cambridge University Press, 1995.

J. Mecke. Stationäre zufällige Maße auf lokalkompakten abelschen Gruppen. *Zeitsch. Wahrsch.*, 9:36–58, 1967.

P. Mörters. Average densities and linear rectifiability of measures. *Mathematika.* 44:313–324, 1997.

P. Mörters. Pathwise Kallianpur-Robbins laws for Brownian motion in the plane. *Preprint* , 1998.

P. Mörters, D. Preiss. Tangent measure distributions of fractal measures. *Math. Ann.* 312:53-93, 1998.

U. Zähle. Self-similar random measures I. *Probab. Theory Rel. Fields*, 80:79–100, 1988.