Mathematische Annalen Manuscript-Nr.
(will be inserted by hand later)

Tangent measure distributions of fractal measures

Peter Mörters[1] and David Preiss[2⋆]

[1] Universität Kaiserslautern, Fachbereich Mathematik, 67663 Kaiserslautern, Germany.
[2] University College London, Department of Mathematics, London WC1E 6BT, U.K.

February 26, 1998

Summary. Tangent measure distributions provide a natural tool to study the local geometry of fractal sets and measures in Euclidean spaces. The idea is, loosely speaking, to attach to every point of the set a family of random measures, called the α-dimensional tangent measure distributions at the point, which describe asymptotically the α-dimensional scenery seen by an observer zooming down towards this point. This tool has been used by Bandt [Ba] and Graf [G] to study the regularity of the local geometry of self similar sets, but in this paper we show that its scope goes much beyond this situation and, in fact, it may be used to describe a strong regularity property possessed by *every* measure: We show that, for every measure μ on a Euclidean space and any dimension α, at μ-almost every point, all α-dimensional tangent measure distributions are Palm measures. This means that the local geometry of every dimension of general measures can be described —like the local geometry of self similar sets— by means of a family of statistically self similar random measures. We believe that this result reveals a wealth of new and unexpected information about the structure of such general measures and we illustrate this by pointing out how it can be used to improve or generalize recently proved relations between ordinary and average densities.

Mathematics Subject Classification (1995): 28A80, 28A75, 60G57.

1. Introduction

The study of the regularity of the local behaviour of a measure μ on a Euclidean space is much simplified by the fact that, under the presence of mild a priori

⋆ Research was supported by the ARC-Programme of British Council and DAAD.

estimates, the weak compactness of bounded sets of measures allows one to define the measures representing the local behaviour of μ at a point x as limit points of enlargements of μ about x. In various disguises, this idea has been used for a long time: Examples include Young measures (see e.g. [V]), tangent cones to currents (see [Fe, 4.3.16]), Bonnet's recent approach to problems in image segmentation (see [Bo]) or tangent measures (see [Pr]). The applicability of this extremely valuable tool stems from the observation that taking this particular limit of measures often improves their geometric behaviour; this behaviour then can be analyzed and, by approximation, information may be obtained about the original measure.

We illustrate this idea further by looking at the concept of tangent measures, which was introduced in [Pr] for the study of sets and measures of integer dimension and their rectifiability questions. Recall that a measure μ is called α-rectifiable if it is absolutely continuous with respect to α-Hausdorff measure on a countable family of smooth α-manifolds. For a measure μ on $\mathbf{R}^d$ and $x \in \mathbf{R}^d$ we define for every $r > 0$ the $(1/r)$-enlargement $\mu_{x,r}$ by $\mu_{x,r}(A) = \mu(x + rA)$. The *$\alpha$-dimensional tangent measures* of μ at x are the vague limit points of $\mu_{x,r}/r^\alpha$ as $r \downarrow 0$. The key to most applications of tangent measures is the fact that on the one hand many weak local regularity properties of μ at x translate into strong almost everywhere regularity properties regularity properties of the tangent measures at x and on the other hand μ is α-rectifiable if and only if at μ-almost every point the set of α-dimensional tangent measures is a singleton consisting of a positive multiple of Lebesgue measure on a linear α-subspace $V \subseteq \mathbf{R}^d$, the approximate tangent space at x. The strength of this method becomes apparent in [Mat] and interesting new results in this direction may be found in [O2] and [H].

The limitations of the tangent measure approach, however, become clear in investigations of non-rectifiable sets and measures. Simple examples show that the set of tangent measures may be rather large and contain measures that do not have much better geometric properties than the original one. Therefore for the investigation of non-rectifiable measures more refined tools seem to be necessary.

One class of tools is based on an averaging idea of Bedford and Fisher in [BeFi]. Instead of looking directly at all limit points of the enlargements of μ about x as the enlargement factor goes to infinity, we take a natural family of probability distributions on the set of positive reals whose mass tends to infinity and study the limiting distributions of the induced family of distributions on the set of enlargements. The limiting distributions define "random tangent measures" which are called tangent measure distributions. This concept first appeared in a weaker form in [BeFi] and then in its full strength in [Ba] and [G]. Bandt and Graf used tangent measure distributions to study the particular case of self similar sets and found that they have unique tangent measure distributions at almost all points, which they described explicitly as scaling invariant Palm measures. Scaling invariant Palm measures were studied and termed self similar random measures by U. Zähle in [ZU]. We give some more details about these results in the next section.

Starting point of our research was the question to what extent the strong results on connections between rectifiability of a measure and smoothness of its tangent measures have their analogy in the fractal world. Are there analogous connections between self similarity of a measure and statistical self similarity of its tangent measure distributions? More precisely, we were looking for a converse of the results of Bandt and Graf: Is self similarity of a set equivalent to all its tangent measure distributions being Palm measures?

This conjecture has turned out to be wrong. Instead we found a surprising and strong result of opposite nature: We prove here that for *every* measure μ on a Euclidean space and every dimension α at μ-almost every point all α-dimensional tangent measure distributions are Palm measures. As all tangent measure distributions are scaling invariant, this means that for every measure the local geometry of every dimension is, from the statistical point of view, given by a family of self similar random measures and hence not different from the local geometry of a self similar set.

To sum up, one can say that tangent measure distributions are a widely applicable example of the effect described at the beginning of this introduction: The localization process leads to objects of good geometric regularity even if the original objects are highly irregular. We believe that this improves the understanding of the geometry of non-rectifiable sets and measures and we illustrate the strength of this result by means of several examples of its application to questions studied previously in [BeFi], [MöPr], [Mö2] and [FaSp], [Ma2].

We also remark that our result is new even on the line. However, in the important case of measures on the line with positive lower and finite upper densities it is possible to deduce the Palm property of tangent measure distributions from very detailed information about the structure of these measures which has been obtained in the thesis [Mö1] and will appear in [Mö2].

2. The main result and its applications

In this section we give the full definitions and basic properties of the main geometric measure theoretical notions describing the α-dimensional part of the behaviour of a Radon measure μ on $\mathbf{R}^d$ about a point $x \in \mathbf{R}^d$ and we formulate the precise statement of our main result. We also prove the aforementioned applications, but delay the proof of the main theorem until the Sections 3 and 4. We shall frequently refer to some measure theoretical preliminaries, some of them not belonging to standard knowledge, which we have collected in the appendix of this paper.

Tangent measure distributions, definition and remarks

We denote by $\mathcal{M}(\mathbf{R}^d)$ the Polish space of all nonnegative Radon measures on $\mathbf{R}^d$ endowed with the vague topology. Finite Radon measures on the Polish space $\mathcal{M}(\mathbf{R}^d)$ will be termed *measure distributions* on $\mathbf{R}^d$; depending on the context we may also call them random measures. The measure distributions we are interested in will always be sub-probability measures and often, but not always, they will be probability measures. We say that a sequence (P_n) of measure distributions *converges vaguely* to a measure distribution P if the following two conditions hold

1. $P(K) \geq \limsup_{k\to\infty} P_k(K)$ for every compact set $K \subseteq \mathcal{M}(\mathbf{R}^d)$, and

2. $P(L) \le \liminf_{k\to\infty} P_k(L)$ for every open set $L \subseteq \mathcal{M}(\mathbf{R}^d)$.

The vague convergence is weaker than the usual weak convergence of probability distributions and this enables us to give a stronger formulation of our main theorem. The relevant properties of the space $\mathcal{M}(\mathbf{R}^d)$ and the vague convergence may be found in the appendix.

Definition 1. *Suppose μ is a Radon measure on $\mathbf{R}^d$ and $x \in \mathbf{R}^d$. We define the family of measures $(\mu_{x,r})_{r>0} \subseteq \mathcal{M}(\mathbf{R}^d)$, the enlargements of μ about x, by $\mu_{x,r}(A) = \mu(x + rA)$ for all Borel sets $A \subseteq \mathbf{R}^d$ and probability distributions P_δ^x on $\mathcal{M}(\mathbf{R}^d)$ by*

$$P_\delta^x(M) = (|\log\delta|)^{-1} \int_\delta^1 1_M\Big(\frac{\mu_{x,r}}{r^\alpha}\Big)\,\frac{dr}{r} \quad \textit{for Borel sets } M \subseteq \mathcal{M}(\mathbf{R}^d).$$

$\mathcal{P}^\alpha(\mu,x)$ is defined as the set of all limit points of $(P_\delta^x)_{\delta>0}$ as $\delta \downarrow 0$ in the vague convergence, The elements of $\mathcal{P}^\alpha(\mu,x)$ are the α-dimensional tangent measure distributions of μ at x.

This approach also allows us to study the local geometry of sets $E \subseteq \mathbf{R}^d$ by means of natural measures on the set, like for example Hausdorff measures.

Remarks:

a) Note that P_δ^x is the push-forward under the mapping

$$(1,\infty) \to \mathcal{M}(\mathbf{R}^d)\ ,\ t \mapsto \frac{\mu_{x,1/t}}{(1/t)^\alpha}$$

of the measure p_δ given as the normalized restriction of Haar measure on the group $(0,\infty)$ with multiplication to the interval $(0, 1/\delta)$,

$$p_\delta(A) = (\log(1/\delta))^{-1} \int_1^{1/\delta} 1_A(t)\,\frac{dt}{t} \quad \text{for } A \subseteq \mathbf{R} \text{ Borel}.$$

Observe that, for every bounded set $A \subseteq (1,\infty)$ we have $\lim_{\delta\to 0} p_\delta(A) = 0$ and hence the mass escapes to infinity as $\delta \to 0$. $(p_\delta)_{\delta>0}$ is a natural choice of probability distributions on the set of scales $(1,\infty)$ because of the scaling invariance property $\lim_{\delta\to 0} p_\delta(sA) = \lim_{\delta\to 0} p_\delta(A)$ for all Borel sets $A \subseteq (1,\infty)$ and $s > 0$. Another motivation for this choice comes from the intimate relationship of the family (p_δ) to the order two averaging method of Hardy and Riesz. Although the behaviour of tangent measure distributions we describe in this paper is sensitive to the choice of (p_δ) our results also hold for tangent measure distributions with respect to the families (p_δ) coming from higher order averaging methods, this follows from the general theory of averaging procedures, see [BeFi] and [Fi] for more details.

b) No a priori relation between μ, α and x has been assumed. Observe that, for every μ and all choices of dimension $\alpha \ge 0$ we always have $\mathcal{P}^\alpha(\mu,x) \ne \emptyset$; indeed, Lemma A4 of the appendix shows that for every sequence $\delta_k \downarrow 0$ the sequence $P_{\delta_k}^x$ has a vaguely convergent subsequence. However, although the P_δ^x are probability measures, the $P \in \mathcal{P}^\alpha(\mu,x)$ need only be sub-probability measures. Hence a tangent measure distribution $P \in \mathcal{P}^\alpha(\mu,x)$ may be trivial in two ways: Either P is concentrated on the zero-measure, this case corresponds, roughly speaking, to the situation when the dimension of μ at x is in fact bigger than α, or P is the zero measure distribution, this case corresponds to the situation when the dimension of μ at x is smaller than α. We give some easy conditions for the nontriviality of tangent measure distributions in Proposition 2.

c) Several close concepts have been used to describe the behaviour of a Radon measure μ on $\mathbf{R}^d$ about a point. The most classical are the *lower and upper α-dimensional densities* of μ at x defined by

$$\liminf_{r\downarrow 0} \frac{\mu(U(x,r))}{r^\alpha} \quad \text{and} \quad \limsup_{r\downarrow 0} \frac{\mu(U(x,r))}{r^\alpha} ,$$

respectively, where $U(x,r)$ denotes the open Euclidean ball of radius r with centre at x. More recently Bedford and Fisher introduced in [BeFi] the notion of *lower and upper average, or order two, α-dimensional densities* of μ at x defined by

$$\liminf_{\delta\downarrow 0} (|\log\delta|)^{-1} \int_\delta^1 \frac{\mu(U(x,r))}{r^\alpha}\,\frac{dr}{r} \quad \text{and} \quad \limsup_{\delta\downarrow 0} (|\log\delta|)^{-1} \int_\delta^1 \frac{\mu(U(x,r))}{r^\alpha}\,\frac{dr}{r} ,$$

respectively. The *α-dimensional tangent measures* of μ at x are defined as elements of the set

$$\mathrm{Tan}_\alpha(\mu,x) = \{\nu = \lim_{n\to\infty} \frac{\mu_{x,r_n}}{r_n^\alpha} \text{ in the vague topology for some } r_n \downarrow 0 \ \} .$$

Again, no relation between α, μ and x has been assumed. Note also that α-dimensional tangent measures may form a proper subset of the set of tangent measures as considered originally in [Pr]. A further natural concept, which was introduced in [G], are the *α-dimensional average tangent measures* of μ at x defined as vague limit points of the Radon measures

$$(|\log\delta|)^{-1} \int_\delta^1 \frac{\mu_{x,r}}{r^\alpha}\,\frac{dr}{r}$$

as $\delta \downarrow 0$. Under certain conditions, formulated in Proposition 8, the average tangent measures are the expectations of the tangent measure distributions and the average densities are the values of the average tangent measures at the unit ball. The relation of tangent measure distributions and tangent measures is given in the following proposition.

Proposition 1. *Suppose that $\mu \in \mathcal{M}(\mathbf{R}^d)$ and $\alpha \geq 0$. For every $x \in \mathbf{R}^d$ all $P \in \mathcal{P}^\alpha(\mu,x)$ are concentrated on* $\mathrm{Tan}_\alpha(\mu,x)$.

Proof. If the sequence $P^x_{\delta_k}$ vaguely converges to P and $P(\mathcal{M}(\mathbf{R}^d)\setminus \mathrm{Tan}_\alpha(\mu,x))$ > 0, we find a $\delta > 0$ such that $P(\mathcal{M}(\mathbf{R}^d)\setminus M_\delta) > 0$, where M_δ is the closure of $\{\mu_{x,r}/r^\alpha, 0 < r \leq \delta\}$. Hence $\liminf P^x_{\delta_k}(\mathcal{M}(\mathbf{R}^d)\setminus M_\delta) \geq P(\mathcal{M}(\mathbf{R}^d)\setminus M_\delta) > 0$, which contradicts the definition of P^x_δ. □

Basic properties of tangent measure distributions

We shall give some statements concerning compactness, localization and scaling invariance properties of tangent measure distributions.

Proposition 2. *Suppose that $\mu \in \mathcal{M}(\mathbf{R}^d)$ and $\alpha \geq 0$.*

(a) If μ has finite upper average α-dimensional density at x, then for every sequence $\delta_k \downarrow 0$ the sequence $P^x_{\delta_k}$ has a weakly convergent subsequence. In particular, all $P \in \mathcal{P}^\alpha(\mu, x)$ are probability measures and the set $\mathcal{P}^\alpha(\mu, x)$ is compact in the weak topology.

(b) If μ has positive lower α-dimensional density at x, then $P(\{\phi\}) = 0$ for every $P \in \mathcal{P}^\alpha(\mu, x)$. (Here ϕ denotes the zero measure.)

Proof. If μ has finite upper average α-dimensional density at x, we infer from

$$(|\log \delta|)^{-1} \int_\delta^1 \frac{\mu(U(x,nr))}{r^\alpha} \frac{dr}{r} = n^\alpha (|\log \delta|)^{-1} \int_{n\delta}^n \frac{\mu(U(x,r))}{r^\alpha} \frac{dr}{r}$$

that the numbers

$$D_n = \sup \left\{ (|\log \delta|)^{-1} \int_\delta^1 \frac{\mu(U(x,nr))}{r^\alpha} \frac{dr}{r} : 0 < \delta < 1/2 \right\}$$

are finite. Thus, for any $\eta > 0$, we may choose $c(n) \in (0, \infty)$ such that $\sum_{n=1}^\infty \frac{D_n}{c(n)} < \eta$ and observe that the set

$$K = \{\nu \in \mathcal{M}(\mathbf{R}^d) : \nu U(0,n) \leq c(n) \text{ for } n = 1, 2, \dots\}$$

is compact and that by Chebyshev's inequality $P^x_\delta(\mathcal{M}(\mathbf{R}^d) \setminus K) < \eta$. Hence the existence of a weakly convergent subsequence and the compactness statement follows from Lemmas A4 and A6 and the fact that $\mathcal{P}^\alpha(\mu, x)$ consist of probability measures follows from Lemma A6. Finally, if μ has positive lower α-dimensional density at x, then $\mathrm{Tan}_\alpha(\mu, x)$ does not contain the zero measure. Hence (b) follows from Proposition 1. □

Remark. Observe that by Proposition 2, whenever μ has positive lower and finite upper α-dimensional densities at x, all tangent measure distributions $P \in \mathcal{P}^\alpha(\mu, x)$ are nontrivial in the strict sense that $P(\mathcal{M}(\mathbf{R}^d) \setminus \{\phi\}) = 1$.

The following lemma is the key to the localization properties of tangent measure distributions.

Lemma 1. *Suppose that $\mu, \nu \in \mathcal{M}(\mathbf{R}^d)$ are Radon measures, $\alpha \geq 0$ and $x \in \mathbf{R}^d$ is such that*

$$\lim_{r \downarrow 0} \frac{\sup\{|\mu(E) - \nu(E)| : E \subseteq U(x,r)\}}{(\mu + \nu)(U(x,r))} = 0$$

and that $\delta_k \downarrow 0$. Then, if for either of μ or ν the sequence $P^x_{\delta_k}$ converges vaguely, then both sequences converge and they have the same limit.

Proof. Denote by P_k and Q_k the sequence $P^x_{\delta_k}$ with respect to μ, respectively ν. Because of the possibility of passing to convergent subsequences, it suffices to assume that P_k and Q_k converge vaguely to, say P and Q, respectively. If $\eta > 0$ and $L \subseteq \mathcal{M}(\mathbf{R}^d)$ is open, we use Lemma A2(b) to find a compact set $K \subseteq L$ such that $\limsup_{k\to\infty} P_k(K) > P(L) - \eta$. Since L is a vague neighbourhood of the compact set K, there is a positive integer m such that, with a suitable $0 < \sigma < 1$, L contains every measure γ which satisfies

$$\sup\{|\gamma(E) - \beta(E)| : E \subseteq U(0,m)\} \leq \sigma \text{ for some } \beta \in K. \tag{1}$$

Let $c(n) \in (1,\infty)$ be such that $\beta U(0,n) < c(n)$ for all $\beta \in K$ and let $0 < R < 1$ be such that

$$\frac{\sup\{|\mu(E) - \nu(E)| : E \subseteq U(x,r)\}}{(\mu+\nu)(U(x,r))} < \sigma/(3c(m))$$

for $0 < r < R$; using this with $E = U(x,r)$, we infer that $\nu U(x,r) \leq 2\mu U(x,r)$ and, consequently,

$$\frac{\sup\{|\mu(E) - \nu(E)| : E \subseteq U(x,r)\}}{\mu U(x,r)} < \sigma/c(m)$$

for $0 < r < R$ and $E \subseteq U(x,r)$.

Let K_0 be the set of all measures γ which have the property (1) and, in addition, satisfy

$$\gamma U(0,n) \leq 2c(n) + (n/R)^\alpha \nu U(x,n) \text{ for every } n \geq m\ .$$

Then $K \subseteq K_0 \subseteq L$ and K_0 is compact. Moreover, if $0 < r < R/m$ and $\frac{\mu_{x,r}}{r^\alpha} \in K$, we show that $\frac{\nu_{x,r}}{r^\alpha} \in K_0$. Indeed, if $nr \geq R$, then

$$\frac{\nu_{x,r}}{r^\alpha}(U(0,n)) \leq \frac{\nu U(x,n)}{r^\alpha} \leq (n/R)^\alpha \nu U(x,n)\ ,$$

if $nr < R$, then

$$\frac{\nu_{x,r}}{r^\alpha}(U(0,n)) \leq \frac{\nu U(x,nr)}{\mu U(x,nr)} \frac{\mu_{x,r}}{r^\alpha}(U(0,n)) \leq 2c(n)$$

and, finally,

$$\begin{aligned} &\sup\left\{\left|\frac{\mu_{x,r}}{r^\alpha}(E) - \frac{\nu_{x,r}}{r^\alpha}(E)\right| : E \subseteq U(0,m)\right\} \\ \leq\ &\frac{\sup\{|\mu(E) - \nu(E)| : E \subseteq U(x,mr)\}}{\mu U(x,mr)} \frac{\mu_{x,r}}{r^\alpha}(U(0,m)) < \sigma. \end{aligned}$$

Consequently, $Q(L) \geq Q(K_0) \geq \limsup_{k\to\infty} Q_k(K_0) \geq \limsup_{k\to\infty} P_k(K) > P(L) - \eta$, which, as $\eta > 0$ is arbitrary, implies that $Q(L) \geq P(L)$ for every open set $L \subseteq \mathcal{M}(\mathbf{R}^d)$. Exchanging the roles of P and Q, we finally conclude that $P = Q$. □

Proposition 3. *Suppose that $\alpha \geq 0$ and that $E \subseteq \mathbf{R}^d$ is μ-measurable. Then the restriction of μ to E, which is the Radon measure $\mu|_E$ on $\mathbf{R}^d$ defined by $\mu|_E(A) = \mu(E \cap A)$, satisfies $\mathcal{P}^\alpha(\mu,x) = \mathcal{P}^\alpha(\mu|_E,x)$ for μ-almost every $x \in E$.*

Proof. For μ-almost every $x \in E$ we infer from the Besicovitch-Morse Density Theorem (see [Fe, 2.9.11]) that

$$\lim_{r \downarrow 0} \frac{\mu(B(x,r) \setminus E)}{\mu(B(x,r))} = 0 \,.$$

For any such $x \in E$ the required statement holds because the measure μ and $\nu = \mu|_E$ satisfy the assumptions of Lemma 1. $\square$

Remark. Similarly, Lemma 1 may be used to describe tangent measure distributions of a measure absolutely continuous with respect to μ: If the Radon-Nikodym derivative of ν with respect to μ is f, and if x is a Lebesgue point of f, i.e. $\lim_{r\to 0} \int_{U(x,r)} |f(x) - f(y)|\, d\mu(y)/\mu U(x,r) = 0$, we can first transform tangent measure distributions of μ to its constant multiple by $f(x)$, and then infer from Lemma 1 that these are precisely the tangent measure distributions of ν at x. Since μ-almost every point is a Lebesgue point of f, this applies in μ, hence ν, almost every point of $\mathbf{R}^d$.

In the following proposition we see that tangent measure distributions have an automatic scaling invariance property. This is largely due to the choice of the measures (p_δ) on the set of scales.

Proposition 4. *For every $\lambda > 0$ we define the rescaling operator $S_\lambda^\alpha : \mathcal{M}(\mathbf{R}^d) \to \mathcal{M}(\mathbf{R}^d)$ by $S_\lambda^\alpha \nu(E) = (1/\lambda^\alpha) \cdot \nu(\lambda E)$. For all $x \in \mathbf{R}^d$ and $P \in \mathcal{P}^\alpha(\mu, x)$ we have*

$$P = P \circ (S_\lambda^\alpha)^{-1} \quad \textit{for all } \lambda > 0.$$

Proof. Let $\lambda > 0$ and observe that S_λ^α is a homeomorphism of $\mathcal{M}(\mathbf{R}^d)$ onto itself, hence it preserves compact and open subsets of $\mathcal{M}(\mathbf{R}^d)$. Suppose that $\delta_k \downarrow 0$ and that P is the vague limit of $P_{\delta_k}^x$. For every compact set $K \subseteq \mathcal{M}(\mathbf{R}^d)$ and every open set $L \subseteq \mathcal{M}(\mathbf{R}^d)$ we calculate

$$\begin{aligned}
(P \circ (S_\lambda^\alpha)^{-1})(K) &= P((S_\lambda^\alpha)^{-1}(K)) \\
&\geq \limsup_{k\to\infty} (|\log \delta_k|)^{-1} \int_{\delta_k}^1 1_{(S_\lambda^\alpha)^{-1}(K)} \left(\frac{\mu_{x,r}}{r^\alpha}\right) \frac{dr}{r} \\
&= \limsup_{k\to\infty} (|\log \delta_k|)^{-1} \int_{\delta_k}^1 1_K \left(\frac{\mu_{x,\lambda r}}{(\lambda r)^\alpha}\right) \frac{dr}{r} \\
&= \limsup_{k\to\infty} (|\log \delta_k|)^{-1} \int_{\delta_k/\lambda}^{1/\lambda} 1_K \left(\frac{\mu_{x,\lambda r}}{(\lambda r)^\alpha}\right) \frac{dr}{r} \\
&= \limsup_{k\to\infty} (|\log \delta_k|)^{-1} \int_{\delta_k}^1 1_K \left(\frac{\mu_{x,r}}{r^\alpha}\right) \frac{dr}{r} \\
&= \limsup_{k\to\infty} P_{\delta_k}^x(K)
\end{aligned}$$

and, similarly, $(P \circ (S_\lambda^\alpha)^{-1})(L) \leq \liminf_{k\to\infty} P_{\delta_k}^x(L)$. Thus $P_{\delta_k}^x$ converges vaguely to $P \circ (S_\lambda^\alpha)^{-1}$, which, because of the unique determination of the limit, proves that $P = P \circ (S_\lambda^\alpha)^{-1}$. $\square$

Examples of tangent measure distributions

Using Proposition 2 we can find the α-dimensional tangent measure distributions in several trivial situations.

Proposition 5. *For every measure $\mu \in \mathcal{M}(\mathbf{R}^d)$, at μ-almost every $x \in \mathbf{R}^d$,*

(a) the only 0-dimensional tangent measure distribution is given by the Dirac measure concentrated at the $\mu(\{x\})$ multiple of the Dirac measure concentrated at the origin (in fact, this holds at every x),
(b) the only d-dimensional tangent measure distribution is given either by the zero measure or by the Dirac measure concentrated at a multiple of Lebesgue measure on $\mathbf{R}^d$ depending on whether $\limsup_{r\to 0} \mu U(x,r)/\mathcal{L}^d U(x,r)$ is infinite or finite (and, in the latter case, this limit defines the multiple),
(c) if $\alpha > d$, then the only α-dimensional tangent measure distribution is the zero measure distribution.

Proof. This is immediate from the fact that the α-dimensional tangent measure distributions are concentrated on the set of α-dimensional tangent measures and the form of these sets, see [Mat, Chap. 14–16]. □

In the case of α-rectifiable sets we see that tangent measure distributions are a natural generalization of the classical concept of a tangent space.

Proposition 6. *If μ is α-rectifiable then, at μ-almost every $x \in \mathbf{R}^d$, the α-dimensional tangent measure distribution at x is unique and given by a Dirac measure concentrated at the only tangent measure at this point, which is a constant multiple of Hausdorff measure on the approximate tangent space at the point.*

Proof. This is immediate from the fact that the α-dimensional tangent measure distributions are concentrated on the set of α-dimensional tangent measures. □

The tangent measure distributions of self similar sets were studied by Graf [G], who proved the following statement.

Proposition 7. *Suppose $K \subseteq \mathbf{R}^d$ is a self similar set in the sense that K is the unique compact set such that $K = \bigcup_{i=1}^n S_i K$ for a family $(S_1, \ldots, S_n)$ of contractive similitudes with the property that for the Hausdorff dimension α of K we have $\mathcal{H}^\alpha(S_i K \cap S_j K) = 0$ for all $i \neq j$. Then there is a measure distribution P such that for the α-Hausdorff measure μ on K the set of α-dimensional tangent measure distributions is $\mathcal{P}^\alpha(\mu, x) = \{P\}$ for μ-almost every x. Moreover, P is given by its marginals P_B on $\mathcal{M}(B)$, for every compact ball $B \subseteq \mathbf{R}^d$, by*

$$P_B(M) = \frac{1}{-\sum_{i=1}^n r_i^\alpha \log r_i} \int_G \int \int_{a(x)}^{b(x)} 1_M\Big(\frac{\mu_{x,r} \circ g}{r^\alpha}\Big)\, \frac{dr}{r}\, d\mu(x)\, d\nu(g)\,,$$

where $(r_1, \ldots, r_n)$ are the contraction ratios and G is the compact group generated by the rotational parts of $(S_1, \ldots, S_n)$, ν is normalized Haar measure on G and $a(x), b(x)$ are constants depending on x and B.

The measure distribution P in Graf's result has been described before as the natural randomization of μ by Patzschke and U. Zähle [PaZU].

Other explicit formulae for tangent measure distributions in self similar situations may be found in [APa] for the case of random self similar sets, in [KM] for the case of hyperbolic Cantor sets and in [Mö3] for the paths of Brownian motion.

The main result

While it is easy to prove the scaling invariance of tangent measure distributions, it is a more interesting and difficult task to formulate and prove their shift-invariance property. For every $u \in \mathbf{R}^d$ we define the shift operator $T^u : \mathcal{M}(\mathbf{R}^d) \to \mathcal{M}(\mathbf{R}^d)$ by $T^u\nu(E) = \nu(u+E)$. Our main theorem can be seen as a shift-invariance principle for tangent measure distributions.

Theorem 1. *Let $\alpha \geq 0$ and let $\mu \in \mathcal{M}(\mathbf{R}^d)$ be any measure. Then, at μ-almost all points x, every α-dimensional tangent measure distribution $P \in \mathcal{P}^\alpha(\mu, x)$ satisfies*

$$\iint G(\nu, u)\, d\nu(u)\, dP(\nu) = \iint G(T^u\nu, -u)\, d\nu(u)\, dP(\nu) \tag{2}$$

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

Remark. Because of the simple description of tangent measure distributions of rectifiable measures (see Proposition 6), the formula is easy whenever μ is rectifiable. In particular it holds trivially in the special cases $\alpha = 0$ and $\alpha \geq d$.

We point out that there is also a translational invariance property for tangent measures, see [Pr, 2.12], but the invariance property above is of a much more rigid nature and also provides additional information about the local symmetry of measures, which has no analogue for tangent measures. We give more details in the next subsection. Before that, since the formulation of the translational invariance property may not be obvious, we explain some reformulations of Theorem 1. The proof of the theorem is deferred to Sections 3 and 4.

Measure distributions satisfying (2) are well known in the theory of random measures and are called Palm measures. The meaning of (2) becomes clearer when we apply it to a function G satisfying the shift invariance condition $G(\nu, u) = G(T^u\nu, 0)$ for all $\nu \in \mathcal{M}(\mathbf{R}^d)$ and $u \in \mathbf{R}^d$. Then we get, for every ball B around the origin,

$$\iint_B G(\nu, u)\, d\nu(u)\, dP(\nu) = \iint_B G(\nu, 0)\, d\nu(u)\, dP(\nu)\,,$$

and we can interpret Palm measures as measure distributions which have the origin as a typical point of their realizations. Palm measures 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 (see [K, Chap.10]).

The original, slightly more restrictive, definition derives the Palm measure from a stationary, i.e. (T^u)-invariant, σ-finite nonzero measure Q on $\mathcal{M}(\mathbf{R}^d)$. We say such a measure has *finite intensity* if the measure $\int \nu dQ(\nu)$ is locally finite and hence, by shift-invariance, a positive multiple of d-dimensional Lebesgue measure. The multiplicity is called the *intensity* of Q. A measure distribution P on $\mathcal{M}(\mathbf{R}^d)$ is a *Palm measure* if there is a stationary σ-finite measure Q on $\mathcal{M}(\mathbf{R}^d)$ with finite intensity and

$$\int_M \nu(B)\,dQ(\nu) = \int_B P \circ T^u(M)\,du \quad \text{for all } M \subseteq \mathcal{M}(\mathbf{R}^d),\ B \subseteq \mathbf{R}^d \text{ Borel.}$$

By a theorem of Mecke (see [Me]) the formula (2) characterizes P as a Palm measure in this sense.

Lemma 2. *A measure distribution P on $\mathcal{M}(\mathbf{R}^d)$ is a Palm measure if and only if $P(\{\phi\}) = 0$, where ϕ is the zero-measure, and the Palm formula (2) holds.*

Clearly $P(\{\phi\}) = 0$ for all $P \in \mathcal{P}^\alpha(\mu, x)$ if the lower α-dimensional density of μ at x is positive. Thus our main result implies that

Theorem 2. *Let $\alpha \geq 0$ and $\mu \in \mathcal{M}(\mathbf{R}^d)$ be any measure. Then, at μ-almost every point x at which the lower α-dimensional density of μ is positive, every tangent measure distribution $P \in \mathcal{P}^\alpha(\mu, x)$ is a Palm measure.*

This statement establishes an interesting connection to the theory of self similar random measures developed by Ulrich Zähle in [ZU]. He suggested, heuristically speaking, that a random measure is statistically self similar if it is scaling invariant with respect to its “typical point” and we can interpret Palm measures as those measure distributions which have the origin as a “typical point”. More precisely, a measure distribution P on $\mathcal{M}(\mathbf{R}^d)$ is an *α-self similar random measure* if P is a Palm measure invariant under the rescaling group $(S^\alpha_\lambda)_{\lambda>0}$.
This concept has been subject of investigations by Patzschke, U. Zähle and M. Zähle, see for example [PaZU] and [PaZM]. In particular, they have studied the relation of the concept to statistically self similar measures in the constructive sense. We extend the notion of α-self similar random measures to allow, in the natural way, for the possibility that $P(\{\phi\}) > 0$ and/or that the total mass of P may be less than one. Of course, no such extension is needed if μ is supposed to have positive lower and finite upper α-dimensional densities. Using the notion of α-self similar random measures, we may reformulate Theorem 1 as

Theorem 3. *Let $\alpha \geq 0$ and $\mu \in \mathcal{M}(\mathbf{R}^d)$ be any measure. Then, at μ-almost all points x, every tangent measure distribution $P \in \mathcal{P}^\alpha(\mu, x)$ defines an α-self similar random measure.*

Applications to average densities

Without further assumptions on the measure, the connection between average densities and tangent measure distributions is not straightforward, since it involves passing to a limit of an integral of the unbounded function $\nu \mapsto \nu U(0,1)$. For our purposes, the following information will suffice.

Proposition 8. *Suppose that $\mu \in \mathcal{M}(\mathbf{R}^d)$, $\alpha \ge 0$, $\delta_n \downarrow 0$ and that $P^x_{\delta_n}$ converge vaguely to a measure distribution P. If*

$$\liminf_{n\to\infty} (|\log\delta_n|)^{-1} \int_{\delta_n}^1 \frac{\mu U(x,r)}{r^\alpha}\,\frac{dr}{r} < \infty, \tag{3}$$

then $P^x_{\delta_n}$ converge to P weakly, hence P is a probability measure and, for every open set $U \subseteq \mathbf{R}^d$,

$$\int \nu U\, dP(\nu) \le \liminf_{n\to\infty} (|\log\delta_n|)^{-1} \int_{\delta_n}^1 \frac{\mu_{x,r}(U)}{r^\alpha}\,\frac{dr}{r}. \tag{4}$$

If μ has finite upper α-dimensional density at x, then there is a compact subset of $\mathcal{M}(\mathbf{R}^d)$ containing the supports of all measures P^x_δ, $0 < \delta \le 1$. In particular, P is a probability measure and, for every compact set $K \subseteq \mathbf{R}^d$,

$$\int \nu K\, dP(\nu) \ge \limsup_{n\to\infty} (|\log\delta_n|)^{-1} \int_{\delta_n}^1 \frac{\mu_{x,r}(K)}{r^\alpha}\,\frac{dr}{r}\,, \tag{5}$$

and, for every $R > 0$,

$$\int \nu U(0,R)\, dP(\nu) = \lim_{n\to\infty} (|\log\delta_n|)^{-1} \int_{\delta_n}^1 \frac{\mu U(x,rR)}{r^\alpha}\,\frac{dr}{r}\,. \tag{6}$$

Proof. As the sequence $P^x_{\delta_n}$ converges weakly if a subsequence converges weakly, we may assume that the limit in (3) exists. For any $m = 1, 2, \dots$ we infer from

$$\begin{aligned} \int \nu U(0,m)\, dP^x_{\delta_n}(\nu) &= (|\log\delta_n|)^{-1} \int_{\delta_n}^1 \frac{\mu U(x,mr)}{r^\alpha}\,\frac{dr}{r} \\ &= m^\alpha (|\log\delta_n|)^{-1} \int_{m\delta_n}^m \frac{\mu U(x,r)}{r^\alpha}\,\frac{dr}{r} \\ &\le m^\alpha (|\log\delta_n|)^{-1} \int_{\delta_n}^1 \frac{\mu U(x,r)}{r^\alpha}\,\frac{dr}{r} \\ &\quad + m^\alpha (|\log\delta_n|)^{-1} \int_1^m \frac{\mu U(x,r)}{r^\alpha}\,\frac{dr}{r} \end{aligned}$$

that the sequence $\int \nu U(0,m)\, dP^x_{\delta_n}(\nu)$, $n = 1, 2, \dots$ is bounded by, say, C_m. Thus, for any $\varepsilon > 0$, the measures $P^x_{\delta_n}$ are concentrated, up to measure ε, on the compact set

$$\bigcap_{m=1}^\infty \{\nu \in \mathcal{M}(\mathbf{R}^d) : \nu U(0,m) \le 2^m C_m/\varepsilon\}\,,$$

which shows that the vaguely convergent sequence $P^x_{\delta_n}$ converges weakly. The inequality (4) follows, since the function $\nu \mapsto \nu U$ is lower semi-continuous.

If μ has finite upper α-dimensional density at x, then the set $\{\mu_{x,r}/r^\alpha\ : 0 < r \le 1\}$ is relatively compact in $\mathcal{M}(\mathbf{R}^d)$; its closure has the required property. Thus the function $\nu \mapsto \nu K$ is upper semi-continuous and bounded on a compact set containing supports of all $P^x_{\delta_n}$. Hence

$$\int \nu K \, dP(\nu) \geq \limsup_{n\to\infty} (|\log \delta_n|)^{-1} \int_{\delta_n}^{1} \frac{\mu_{x,r}K}{r^\alpha} \, \frac{dr}{r} \, ,$$

and

$$\begin{aligned}
\int \nu\overline{U(0,R)} \, dP(\nu) \quad & \geq \quad \limsup_{n\to\infty} (|\log \delta_n|)^{-1} \int_{\delta_n}^{1} \frac{\mu\overline{U(x,Rr)}}{r^\alpha} \, \frac{dr}{r} \\
& \geq \quad \liminf_{n\to\infty} (|\log \delta_n|)^{-1} \int_{\delta_n}^{1} \frac{\mu U(x,Rr)}{r^\alpha} \, \frac{dr}{r} \\
& \geq \quad \int \nu U(0,R) \, dP(\nu) = \int \nu\overline{U(0,R)} \, dP(\nu) \, ,
\end{aligned}$$

using Proposition 4 in the last step. This proves Proposition 8. □

We may use the formula (2) of Theorem 1 with $G(\nu, u) = h(u)$ to infer that, at μ-almost all points x, every tangent measure distribution $P \in \mathcal{P}^\alpha(\mu, x)$ satisfies

$$\int\!\!\int h(u) \, d\nu(u) \, dP(\nu) = \int\!\!\int h(-u) \, d\nu(u) \, dP(\nu) \, , \tag{7}$$

for every Borel function $h : \mathbf{R}^d \longrightarrow [0, \infty)$. Choosing for h characteristic functions only, we get

$$\int \nu(U) \, dP(\nu) = \int \nu(-U) \, dP(\nu) \, , \tag{8}$$

for every Borel subset U of $\mathbf{R}^d$. If μ has finite upper α-dimensional density at x, the expectations $\int \nu \, dP$ of the tangent measure distributions agree with the average tangent measures by Proposition 8 and (8) shows that all average tangent measures are symmetric about the origin. A number of unexpected statements concerning average densities, which have no analogy for ordinary densities, also follow from this equality.

Proposition 9. *If μ has finite upper α-dimensional density almost everywhere, then at μ-almost all points x the following holds: Whenever $U \subseteq V \subseteq \mathbf{R}^d$ with U closed and V open and $\delta_n \downarrow 0$, we have*

$$\liminf_{n\to\infty} (|\log \delta_n|)^{-1} \int_{\delta_n}^{1} \frac{\mu(x + rU)}{r^\alpha} \, \frac{dr}{r} \leq \liminf_{n\to\infty} (|\log \delta_n|)^{-1} \int_{\delta_n}^{1} \frac{\mu(x - rV)}{r^\alpha} \, \frac{dr}{r} \, .$$

Proof. We may assume that $P^x_{\delta_n}$ converges to a measure distribution P. Employing the compactness result of Proposition 8 and (8), we infer that

$$\begin{aligned}
& \liminf_{n\to\infty} (|\log \delta_n|)^{-1} \int_{\delta_n}^{1} \frac{\mu(x + rU)}{r^\alpha} \, \frac{dr}{r} \\
\leq \quad & \int \nu(U) \, dP(\nu) = \int \nu(-U) \, dP(\nu) \\
\leq \quad & \int \nu(-V) \, dP(\nu) \\
\leq \quad & \liminf_{n\to\infty} (|\log \delta_n|)^{-1} \int_{\delta_n}^{1} \frac{\mu(x - rV)}{r^\alpha} \, \frac{dr}{r} \, .
\end{aligned}$$

□

Whenever C is a Borel cone in $\mathbf{R}^d$, one defines the lower average α-dimensional density of a measure μ at a point x with respect to C as

$$\liminf_{\delta\downarrow 0}\, (|\log\delta|)^{-1}\int_\delta^1 \frac{\mu((x+C)\cap U(x,r))}{r^\alpha}\,\frac{dr}{r}\;.$$

If $C=\mathbf{R}^d$, this is the lower average density of the measure μ at a point x, if $d=1$ and $C=(0,\infty)$ we speak about right lower average density of the measure μ at a point x, etc.

Proposition 10. *Suppose $\alpha>0$, $\mu\in\mathcal{M}(\mathbf{R}^d)$ has finite upper α-dimensional density almost everywhere and C is a closed cone which (except for the origin) is contained in the open cone K. Then, at μ-almost all points x, the lower average α-dimensional density of the measure μ with respect to C does not exceed its lower average α-dimensional density with respect to the cone $-K$.*

Proof. For every tangent measure distribution P the scaling invariance property implies that $\int \nu(\overline{U(0,1)}\cap C)\,dP(\nu)\le\int\nu(U(0,1)\cap K)\,dP(\nu)$ and we can employ the same argument as in Proposition 9. $\square$

As an immediate corollary of these propositions we obtain the following result, which was obtained only under the additional assumption of positive lower α-dimensional density in [Mö2] and, in a weaker form, in [MöPr].

Theorem 4. *Let $\alpha>0$ and $\mu\in\mathcal{M}(\mathbf{R})$ be a measure on the line with finite upper α-dimensional densities μ-almost everywhere. Then, at μ-almost all points x, the right and left lower average α-dimensional densities coincide and each of them is half of the lower average α-dimensional density of μ at x.*

Of course, analogous results can be found for upper average densities.
As our final application we show in Theorem 5 how the results of [FaSp] and [Ma2] follow naturally by using the higher regularity of tangent measure distributions. Our approach removes their unnatural assumption of finiteness of the upper densities for results about lower densities.

Lemma 3. *Suppose that a probability measure P on $\mathcal{M}(\mathbf{R}^d)$ defines an α-self similar random measure and that, with some $c>0$, P-almost every ν satisfies $\nu U(0,1)=c$. Then α is an integer.*

Proof. From the invariance of P under the rescaling group $(S_\lambda^\alpha)_{\lambda>0}$ we infer that, for every $r>0$ and P-almost every ν, we have $\nu U(0,r)=cr^\alpha$. Thus, using the Palm formula in the last step,

$$\begin{aligned}1 &= P(\{\mu\,:\,\mu U(0,r)=cr^\alpha\})=\frac{1}{cs^\alpha}\int\!\!\int_{U(0,s)} 1_{\{\mu\,:\,\mu U(0,r)=cr^\alpha\}}(\nu)\,d\nu(u)\,dP(\nu)\\ &= \frac{1}{cs^\alpha}\int\!\!\int_{U(0,s)} 1_{\{\mu\,:\,\mu U(u,r)=cr^\alpha\}}(\nu)\,d\nu(u)\,dP(\nu).\end{aligned}$$

As r and s are arbitrary, for P-almost every ν we infer that $\nu U(u,r)=cr^\alpha$ for ν-almost every u and every rational $r>0$. By continuity, P-almost every ν

satisfies that $\nu U(u,r) = cr^\alpha$ for every u in the support of ν and every $r > 0$. Now it suffices to use Marstrand's Theorem ([Ma1], see also [Mat, 14.10]) to infer that α is an integer. □

Theorem 5. *Suppose that $\mu \in \mathcal{M}(\mathbf{R}^d)$ and that $\alpha > 0$ is not an integer. Then μ-almost all points x satisfy:*

(a) The lower average α-dimensional density and the lower α-dimensional density of μ at x are either both zero, or both infinite, or the former is strictly bigger than the latter.
(b) The upper average α-dimensional density and the upper α-dimensional density of μ at x are either both zero, or both infinite, or the former is strictly smaller than the latter.

Proof. If $d(x)$ and $D(x)$ denote the (ordinary) lower and upper α-dimensional densities of μ, respectively, then, for μ-almost every x, each $\nu \in \mathrm{Tan}_\alpha(\mu, x)$ satisfies $d(x) \le \nu U(0,1) \le D(x)$. To use Lemma 3, it suffices to observe that if the average and ordinary lower, respectively upper, densities of μ at x coincide and are finite, then there is a probability measure $P \in \mathcal{P}^\alpha(\mu, x)$ such that P-almost every ν satisfies that $\nu U(0,1) = d(x)$, respectively $\nu U(0,1) = D(x)$. But if $d(x)$, respectively $D(x)$, is finite, then we can find $\delta_n \downarrow 0$ such that

$$\lim_{n\to\infty} (|\log \delta_n|)^{-1} \int_{\delta_n}^1 \frac{\mu U(x,r)}{r^\alpha} \frac{dr}{r} = d(x) \text{ , respectively } D(x).$$

We can choose P as the limit of a subsequence of $P^x_{\delta_n}$. Then, by Proposition 2(a), $d(x) \le \nu U(0,1) \le D(x)$ for P-almost every ν, and we may use the inequality (4) of Proposition 8 in case of lower densities and the equality (5) of the same proposition in case of upper densities to infer that P-almost every ν satisfies $\nu U(0,1) = d(x)$, respectively $\nu U(0,1) = D(x)$. □

3. The main lemma and its proof

In this section we formulate and prove the main lemma for Theorem 1. The proof of the theorem will then be completed in Section 4.

The main lemma

We start by giving an intuitive picture of the rôle of the main lemma without caring too much about technicalities. We approximate the left hand side in the Palm formula (2) by

$$\int\int G(\nu,u)\, d\nu(u)\, P^x_\delta(\nu) = (|\log\delta|)^{-1} \int_\delta^1 \int G\Big(\frac{\mu_{x,r}}{r^\alpha}, \frac{y-x}{r}\Big) \frac{d\mu(y)}{r^\alpha} \frac{dr}{r}$$

and the right hand side by

$$\begin{aligned}
&\int\int G(T^u\nu, -u)\, d\nu(u)\, P_\delta^x(\nu) \\
&= (|\log\delta|)^{-1}\int_\delta^1 \int G\Big(T^{\frac{y-x}{r}}\Big(\frac{\mu_{x,r}}{r^\alpha}\Big), \frac{x-y}{r}\Big)\,\frac{d\mu(y)}{r^\alpha}\,\frac{dr}{r} \\
&= (|\log\delta|)^{-1}\int_\delta^1 \int G\Big(\frac{\mu_{y,r}}{r^\alpha}, \frac{x-y}{r}\Big)\,\frac{d\mu(y)}{r^\alpha}\,\frac{dr}{r}\,.
\end{aligned}$$

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

$$\lim_{\delta\to 0}(|\log\delta|)^{-1}\int_\delta^1 \int \frac{H(x,y,r)-H(y,x,r)}{r^\alpha}\, d\mu(y)\,\frac{dr}{r} = 0\,.$$

Our main lemma shows this for general mappings $H : \mathbf{R}^d\times\mathbf{R}^d\times(0,\infty)\to(0,\infty)$ that fulfill certain conditions. It does not refer explicitly to the notion of tangent measure distributions. Although our results do not depend on the choice of the norm of $\mathbf{R}^d$, it is convenient to equip $\mathbf{R}^d$ with the maximum norm defined, for $u=(u_1,\dots,u_d)$, by

$$\|u\| = \max\{|u_i| \,:\, i=1,\dots,d\}$$

and denote by $B(u,r)$ the closed balls with respect to this norm, i.e., $B(u,r) = \{x\in\mathbf{R}^d \,:\, \|u-x\|\le r\}$.

Lemma 4. *Whenever a measure $\mu\in\mathcal{M}(\mathbf{R}^d)$, a Borel measurable function $H : \mathbf{R}^d\times\mathbf{R}^d\times(0,\infty)\longrightarrow[0,1]$ and $0<\alpha<d$ satisfy, for some constants $R, A_1, A_2\in(0,\infty)$, the conditions*

(i) $H(x,y;r)=0$ *whenever* $\|x-y\|\ge Rr$,
(ii) $H(x,y;r)=0$ *whenever* $\mu B(x,5Rr)\ge A_1 r^\alpha$, *and*
(iii) for all $x,\tilde{x},y,\tilde{y}\in\mathbf{R}^d$ *and* $0<r<\infty$,

$$|H(x,y;r)-H(\tilde{x},\tilde{y};r)| \le A_2\,\frac{\|x-\tilde{x}\|+\|y-\tilde{y}\|}{r}\,,$$

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

$$\lim_{\delta\downarrow 0}\frac{1}{|\log\delta|}\int_\delta^1\int \frac{H(x,y;r)-H(y,x;r)}{r^\alpha}\, d\mu(y)\,\frac{dr}{r} = 0. \tag{9}$$

The rest of this section is devoted to the proof of this lemma. But before we give any details we outline the strategy behind our proof.

Outline of the proof

Writing $G(x,y,r) = (1/r^\alpha)(H(x,y;r)-H(y,x;r))$ we have to show that, for large integers n,

$$\frac{1}{n}\int_{2^{-n}}^1 \int G(x,y,r)\, d\mu(y)\,\frac{dr}{r} \tag{10}$$

is small. Our first step in the proof is to approximate this situation by discretizing it in a dyadic grid.

Dyadic decomposition of domains is a common tool in analysis, in particular in singular integral theory, see e.g. [St]. The tool is useful when combined with Lebesgue integration over the decomposed domain. We will use this idea here by replacing the integration with respect to dr/r in (10) by a sum over cubes with sidelengths comparable to their distances to x. Roughly speaking we approximate (10) by

$$\int \sum_{k=0}^{n-1} H_k(x,y)\, d\mu(y)\,, \tag{11}$$

where $H_k(x,y) \approx \sum_Q 1_Q(x) 1_Q(y) G(x,y,2^{-k})$ and the summation is with respect to all dyadic cubes of sidelength 2^{-k}. 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.

In order to prove that (11) is small for μ-almost all x we write

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

where Q_k is the unique dyadic cube of sidelength 2^{-k} containing x, and show that

$$\int \Big[\int \sum_{k=0}^{n-1} H_k(x,y)\, d\mu(y) \Big]^2 d\mu(x) = \int \Big(\varphi_x(\mathbf{R}^d) \Big)^2 d\mu(x) \tag{12}$$

is small. We can now find a suitable partition of $\mathbf{R}^d$ into sets $\{G(I)\}_{I\in\mathcal{I}}$, which allows us to write

$$\int \Big(\varphi_x(\mathbf{R}^d) \Big)^2 d\mu(x) = \sum_{I\in\mathcal{I}} \sum_{J\in\mathcal{I}} \int \varphi_x(G(I)) \varphi(G(J))\, d\mu(x)\,. \tag{13}$$

Every set $G(I)$ is a finite union of subcubes of a cube I and for all these cubes we have the necessary lower control of the μ-mass. This will be accomplished by means of a stopping time argument. The proof finishes by giving careful individual estimates for the summands on the right hand side of (13).

We now start with the detailed proof. Our plan is first to formulate exactly the dyadic model for the main lemma, then show how this approximation may be used to prove its statement, and only then to embark upon the series of technical estimates which will eventually result in the proof of the dyadic version of our main lemma.

The dyadic model

For the exact formulation of the lemma we define, for $k = 0, 1, \ldots$, the k-th grid as the family of dyadic cubes

$$\mathcal{Q}_k \quad = \quad \Big\{ \prod_{i=1}^{d} \Big[\frac{k_i}{2^k}, \frac{k_i+1}{2^k} \Big) \,:\, (k_1, \ldots, k_d) \in \mathbf{Z}^d \Big\}.$$

Any $x \in \mathbf{R}^d$ is contained in a unique cube from $\mathcal{Q}_k$; this cube will be denoted by $Q_k(x)$. The dyadic approximation to our main lemma is given in the following statement.

Lemma 5. *Given $0 < \alpha < d$ and $\sigma > 0$ and constants $C_1, C_2 \in (0,\infty)$, we can find a constant $\mathcal{C}_0 = \mathcal{C}_0(\sigma, \alpha, d, C_1, C_2) \in (0,\infty)$ having the following property: Whenever a finite measure $\mu \in \mathcal{M}(\mathbf{R}^d)$ and Borel measurable functions $H_k : \mathbf{R}^d \times \mathbf{R}^d \longrightarrow \mathbf{R}$ satisfy the conditions*

(i) $H_k(x,y) = 0$ *whenever* $y \notin Q_k(x)$,
(ii) $H_k(x,y) = 0$ *whenever* $\mu(Q_k(x)) > C_1 2^{-k\alpha}$,
(iii) $|H_k(x,u) - H_k(y,u)| \le C_2 2^{k\alpha+k-i}$ *if* $i \ge k$, $x, y \in I \in \mathcal{Q}_i$ *and* $u \in Q_k(x)$,
(iv) $H_k(y,x) = -H_k(x,y)$ *for all* $k = 0, 1, \dots$ *and all* $x, y \in \mathbf{R}^d$,

then for any $n = 1, 2, \dots$ and any non-negative function $g \in L_2(\mu)$,

$$\int g(x)\Big(\int \sum_{k=0}^{n-1} H_k(x,u)\, d\mu(u) - \sigma n\Big) d\mu(x) \quad \le \quad \mathcal{C}_0 n^{3/4} \mu(\mathbf{R}^d)^{1/2} \|g\|_{L_2(\mu)}.$$

Although we shall not need it, it is interesting to note that our proof allows one to describe the dependence of $\mathcal{C}_0$ on all parameters explicitly; this shows, in particular, that $\mathcal{C}_0$ may be chosen independently of α as long as α is kept away from d.

Proof of Lemma 4 using Lemma 5

We need only prove Lemma 4 under the additional assumption that μ is finite, as the general case immediately follows by using the special one for the restrictions of μ to $B(0,k)$, $k = 1, 2, \dots$.

Lemma 6. *Suppose that $0 < \alpha < d$, $\mu \in \mathcal{M}(\mathbf{R}^d)$ is a finite measure and $G : \mathbf{R}^d \times \mathbf{R}^d \times (0,\infty) \longrightarrow \mathbf{R}$ is a Borel measurable function, and $\hat{A}_1, \hat{A}_2 \in (0,\infty)$ are constants such that*

(i) $|G(x,y;r)| \le r^{-\alpha}$ *for all* $x, y \in \mathbf{R}^d$ *and* $r > 0$,
(ii) $G(x,y;r) = 0$ *whenever* $\|x - y\| \ge r$,
(iii) $G(x,y;r) = 0$ *whenever* $\mu B(x, 4r) \ge \hat{A}_1 r^\alpha$,
(iv) *for all* $x, y, u \in \mathbf{R}^d$ *and* $0 < r < \infty$,

$$|G(x,u;r) - G(y,u;r)| \le \hat{A}_2 r^{-\alpha-1} \|x - y\| \,,$$

(v) $G(y,x;r) = -G(x,y;r)$ *for all* $x, y \in \mathbf{R}^d$ *and* $r > 0$.

Then for every $\sigma > 0$ there is $\mathcal{C}_1 < \infty$ such that for any $n = 1, 2, \dots$ and any non-negative $g \in L_2(\mu)$,

$$\int g(x)\Big[\int_{2^{-n-1}}^{1/2} \int G(x,y;r)\, d\mu(y) \frac{dr}{r} - \sigma n\Big]\, d\mu(x) \le \mathcal{C}_1 n^{3/4} \|g\|_{L_2(\mu)}. \tag{14}$$

Consequently, for μ-almost all $x \in \mathbf{R}^d$, the equality

$$\lim_{\delta \downarrow 0} \frac{1}{|\log \delta|} \int_{\delta R}^{R} \int G(x,y;r)\, d\mu(y) \frac{dr}{r} = 0 \tag{15}$$

holds for every $R > 0$.

Proof. The function

$$\omega(z) = \prod_{i=1}^{d} \big(1 - \min\{|z_i|, 1/2\}\big)^{-1}$$

verifies $1 \le \omega(z) \le 2^d$ and is Lipschitz on $\mathbf{R}^d$; let C_0 be its Lipschitz constant. We define $C_1 = \hat{A}_1$ and $C_2 = 4^\alpha C_0 + 2^{d+2\alpha+2}\hat{A}_2$, and we prove that, for any $a \in \mathbf{R}^d$ and any $1/4 \le s \le 1/2$, the functions

$$H_k(x,y;a,s) \;=\; \sum_{Q \in \mathcal{Q}_k} 1_Q(x) 1_Q(y)\, G(a+x, a+y; 2^{-k}s)\, \omega\big(2^k(y-x)\big)$$

satisfy the assumptions (i)–(iv) of Lemma 5 for the measure $T^a\mu$. Clearly, 5(i) follows directly from the definition of $H_k(x,y;a,s)$ and 5(iv) follows directly from 6(v). If $T^a\mu(Q_k(x)) > C_1 2^{-k\alpha}$, then

$$\begin{aligned} \mu(B(a+x, 2^{-k+2}s)) &\ge \mu(B(a+x, 2^{-k})) \ge \mu(a + Q_k(x)) = T^a\mu(Q_k(x)) \\ &> C_1 2^{-k\alpha} \ge \hat{A}_1 (2^{-k}s)^\alpha, \end{aligned}$$

which, by 6(iii), shows that $G(a+x, a+y; 2^{-k}s) = 0$ and hence the requirement of 5(ii) that $H_k(x,y;a,s) = 0$. Finally, the condition 5(iii) is also easy, since $i \ge k$, $x,y \in I \in \mathcal{Q}_i$ and $u \in Q_k(x)$ imply, because of 6(ii) and 6(iv), that

$$\begin{aligned} &|H_k(x,u;a,s) - H_k(y,u;a,s)| \\ &\quad= \big|G(x+a, u+a; 2^{-k}s)\, \omega\big(2^k(u-x)\big) - G(y+a, u+a; 2^{-k}s)\, \omega\big(2^k(u-y)\big)\big| \\ &\quad\le \big|\omega\big(2^k(u-x)\big) - \omega\big(2^k(u-y)\big)\big|\, \big|G(x+a, u+a; 2^{-k}s)\big| \\ &\qquad + \big|\omega\big(2^k(u-y)\big)\big|\, \big|G(x+a, u+a; 2^{-k}s) - G(y+a, u+a; 2^{-k}s)\big| \\ &\quad\le (2^{-k}s)^{-\alpha} C_0 2^k \|y - x\| \; + 2^d \hat{A}_2 (2^{-k}s)^{-\alpha-1} \|x - y\| \\ &\quad\le C_2 2^{k\alpha + k - i}\ . \end{aligned}$$

Let $\mathcal{C} = \mathcal{C}_0(\sigma/\log 2, \alpha, d, C_1, C_2)\mu(\mathbf{R}^d)^{1/2}$, where $\mathcal{C}_0$ is the constant introduced in Lemma 5 and let $g \in L_2(\mu)$ be non-negative. The statement of Lemma 5 then implies that, for every $a \in \mathbf{R}^d$ and every $1/4 \le s \le 1/2$,

$$\int g(x+a) \Big(\int \sum_{k=0}^{n-1} H_k(x,u;a,s)\ dT^a\mu(u) - \sigma n/\log 2 \Big) dT^a\mu(x)$$

$$\le \mathcal{C} n^{3/4} \|g\|_{L_2(\mu)}\ .$$

Hence, denoting

$$\begin{aligned} G_k(x,y;a,s) &= H_k(x-a,y-a;a,s) \\ &= \sum_{Q\in\mathcal{Q}_k} 1_{a+Q}(x)1_{a+Q}(y)G(x,y;2^{-k}s)\,\omega\big(2^k(y-x)\big), \end{aligned}$$

we have that

$$\int g(x)\Big(\int\sum_{k=0}^{n-1} G_k(x,u;a,s)\,d\mu(u)-\sigma n/\log 2\Big)d\mu(x)\le \mathcal{C}n^{3/4}\|g\|_{L_2(\mu)}. \quad (16)$$

If $\|y-x\|\le 2^{-k}s$, we use that $s\le 1/2$ to infer that $\omega\big(2^k(y-x)\big)=\prod_{i=1}^d\big(1-2^k|y_i-x_i|\big)^{-1}$. Thus

$$\begin{aligned} &\int_{[0,1]^d}\sum_{Q\in\mathcal{Q}_k} 1_{a+Q}(x)1_{a+Q}(y)\,d\mathcal{L}^d(a) \\ &\quad= \mathcal{L}^d\{a\in[0,1]^d:\sum_{Q\in\mathcal{Q}_k}1_{a+Q}(x)1_{a+Q}(y)=1\} \\ &\quad= 1/\omega\big(2^k(y-x)\big) \end{aligned}$$

and we conclude that

$$G(x,y;2^{-k}s)=\int_{[0,1]^d} G_k(x,y;a,s)\,d\mathcal{L}^d(a). \quad (17)$$

If $\|y-x\|>2^{-k}s$, then both sides of (17) are zero according to (ii); hence (17) may be used for all x,y. Thus, integrating (16) over $[0,1]^d$ with respect to the Lebesgue measure and over $1/4\le s\le 1/2$ with respect to the measure $\frac{ds}{s}$ and using that

$$\int_{1/4}^{1/2} G(x,u;2^{-k}s)\frac{ds}{s}=\int_{2^{-k-2}}^{2^{-k-1}} G(x,u;r)\frac{dr}{r}\ ,$$

we get

$$\int g(x)\Big[\int_{2^{-n-1}}^{1/2}\int G(x,y;r)\,d\mu(y)\frac{dr}{r}-\sigma n\Big]\,d\mu(x)\le \mathcal{C}n^{3/4}\|g\|_{L_2(\mu)}\log 2,$$

which shows that (14) holds with $\mathcal{C}_1=\mathcal{C}\log 2$.
To prove (15), we denote

$$\bar G(x;r)=\int G(x,y;r)d\mu(y)$$

and note that, by 6(ii),6(v) and 6(iv),

$$\begin{aligned} |G(x,y;r)| &= 1_{B(x,r)}(y)|G(x,y;r)-G(y,y;r)|\le \hat A_2 r^{-\alpha-1}\|x-y\|1_{B(x,r)}(y) \\ &\le \hat A_2 r^{-\alpha}1_{B(x,r)}(y). \end{aligned}$$

Since, by 6(iii), the left side is zero unless $\mu(B(x,r))\le \hat A_1 r^{\alpha}$, we obtain by integrating with respect to $d\mu(y)$ that for all x and r,

$$|\bar{G}(x;r)| \le \hat{A}_2 r^{-\alpha}\mu(B(x,r)) \le \hat{A}_1\hat{A}_2. \tag{18}$$

Next we use (14) for the characteristic function g of the set

$$E_n \quad = \quad \Big\{x : \int_{2^{-n-1}}^{1/2} \bar{G}(x;r)\,\frac{dr}{r} > 2\sigma n\Big\},$$

to get that

$$\sigma n \mu(E_n) \quad \le \quad \mathcal{C}_1 n^{3/4}\mu(E_n)^{1/2},$$

hence $\mu(E_n) \le \mathcal{C}_1^2 n^{-1/2}/\sigma^2$, which shows that $\sum_{n=1}^{\infty}\mu(E_{n^3})$ converges. By the Borel-Cantelli Lemma, for μ-almost every $x \in \mathbf{R}^d$ there is a positive integer m such that $x \notin \bigcup_{n=m}^{\infty} E_{n^3}$. Consider any such x and the corresponding m; we may clearly also assume that $m \ge 7\hat{A}_1\hat{A}_2/\sigma$. Whenever $0 < \delta < 2^{-m^3}$ we find $n \ge m$ such that $2^{-(n+1)^3} \le \delta < 2^{-n^3}$ and we note that $x \notin E_{n^3}$ implies that

$$\frac{1}{|\log\delta|}\int_{2^{-n^3-1}}^{1/2} \bar{G}(x;r)\,\frac{dr}{r} = \frac{n^3}{|\log\delta|}\frac{1}{n^3}\int_{2^{-n^3-1}}^{1/2} \bar{G}(x;r)\,\frac{dr}{r} \le 2\sigma/\log 2\,.$$

Since (18) gives

$$\frac{1}{|\log\delta|}\int_{1/2}^{R} \bar{G}(x;r)\,\frac{dr}{r} \le \frac{\hat{A}_1\hat{A}_2(|\log R| + \log 2)}{|\log\delta|}$$

and

$$\begin{aligned}\frac{1}{|\log\delta|}\int_{\delta R}^{2^{-n^3-1}} \bar{G}(x;r)\,\frac{dr}{r} \quad &\le \quad \frac{\hat{A}_1\hat{A}_2(|\log R| + ((n+1)^3 - n^3)\log 2)}{|\log\delta|} \\ &\le \quad \frac{\hat{A}_1\hat{A}_2|\log R|}{|\log\delta|} + \frac{7\hat{A}_1\hat{A}_2}{n},\end{aligned}$$

we infer that

$$\limsup_{\delta\downarrow 0}\frac{1}{|\log\delta|}\int_{\delta R}^{R}\int G(x,y;r)\,d\mu(y)\frac{dr}{r} \le 2\sigma/\log 2$$

for every $R > 0$. Since $\sigma > 0$ may be arbitrary, it follows that for μ-almost every x the inequality

$$\limsup_{\delta\downarrow 0}\frac{1}{|\log\delta|}\int_{\delta R}^{R}\int G(x,y;r)\,d\mu(y)\frac{dr}{r} \le 0$$

holds for every $R > 0$. The statement (15) follows by using this also with G replaced by $-G$. $\square$

Proof of Lemma 4. If the assumptions of Lemma 4 hold, we let $\hat{A}_1 = A_1 R^{-\alpha}$, $\hat{A}_2 = 2A_2 R$, and

$$G(x,y;r) = \frac{H(x,y;r/R) - H(y,x;r/R)}{r^\alpha} \ ,$$

and note that the assumptions (i), (ii), (iv) and (v) of Lemma 6 are clearly satisfied. To establish 6(iii), we observe that it follows directly from 4(i) if $\|x - y\| \geq r$. If $\|x-y\| < r$ then $B(x,5r) \cap B(y,5r) \supseteq B(x,4r)$, hence $\mu B(x,4r) \geq \hat{A}_1 r^\alpha$ implies that $\mu B(x,5r) \geq A_1(r/R)^\alpha$ and $\mu B(y,5r) \geq A_1(r/R)^\alpha$, hence $H(x,y;r/R) = H(y,x;r/R) = 0$ according to 4(ii) .
Hence, using (15) of Lemma 6, we have

$$\begin{aligned}
&\lim_{\delta\downarrow 0} \frac{1}{|\log\delta|} \int_\delta^1 \int \frac{H(x,y;r) - H(y,x;r)}{r^\alpha} \, d\mu(y) \, \frac{dr}{r} \\
= \quad & R^\alpha \lim_{\delta\downarrow 0} \frac{1}{|\log\delta|} \int_{\delta R}^R \int G(x,y;r) \, d\mu(y) \frac{dr}{r} = 0.
\end{aligned}$$

□

Proof of Lemma 5

For the remainder of this section we shall assume that $0 < \alpha < d$ and $\sigma, C_1, C_2 > 0$, the measure μ, the functions H_k and an integer n are fixed. Define the cut-off function

$$\omega_k(x) = \begin{cases} 1 & \text{if } C_2^{-1}\sigma 2^{-k\alpha} \leq \mu(Q_k(x)) \leq C_1 2^{-k\alpha}, \\ 0 & \text{otherwise,} \end{cases}$$

and let, for $k \in \{0,1,\ldots,n-1\}$, $x \in \mathbf{R}^d$ and Borel sets $E \subseteq \mathbf{R}^d$,

$$H_k(x) \quad = \quad \omega_k(x) \int_{Q_k(x)} H_k(x,u) \, d\mu(u)$$

and

$$\varphi_x(E) \quad = \quad (C_1C_2)^{-1} \sum_{k=0}^{n-1} \frac{\mathcal{L}^d\big(E \cap (Q_k(x) \setminus Q_{k+1}(x))\big)}{\mathcal{L}^d\big(Q_k(x) \setminus Q_{k+1}(x)\big)} H_k(x).$$

These notions are justified by the following observation.

Lemma 7. *For every positive function $g \in L_2(\mu)$,*

$$\begin{aligned}
&\int g(x)\Big(\int \sum_{k=0}^{n-1} H_k(x,u) \, d\mu(u) - \sigma n\Big) d\mu(x) \\
\leq \quad & C_1 C_2 \int g(x) \varphi_x(\mathbf{R}^d) d\mu(x) \\
\leq \quad & C_1 C_2 \|g\|_{L_2(\mu)} \left(\int \big(\varphi_x(\mathbf{R}^d)\big)^2 d\mu(x) \right)^{1/2}.
\end{aligned}$$

Proof. By (i) and by the case $i = k$ of (iii), $H_k(x,u) \le C_2 2^{k\alpha} 1_{Q_k(x)}(u)$. Thus, if $\omega_k(x) = 0$, then either $\mu(Q_k(x)) < C_2^{-1}\sigma 2^{-k\alpha}$ and $\int H_k(x,u)\,d\mu(u) \le C_2 2^{k\alpha}\mu(Q_k(x)) < \sigma$, or $\mu(Q_k(x)) > C_1 2^{-k\alpha}$ in which case we infer from (ii) that $\int H_k(x,u)\,d\mu(u) = 0 < \sigma$, and hence

$$\int \sum_{k=0}^{n-1} H_k(x,u)\,d\mu(u) - \sigma n \le \int \sum_{k=0}^{n-1} \omega_k(x) H_k(x,u)\,d\mu(u) = C_1 C_2 \varphi_x(\mathbf{R}^d)\ .$$

The statement follows by multiplying this inequality by $g(x)$, integrating with respect to $d\mu(x)$, and using Hölder's inequality. $\square$

We therefore devote the rest of this section to the estimate of $\int \big(\varphi_x(\mathbf{R}^d)\big)^2 d\mu(x)$. Since working with the signed measures φ_x may sometimes be slightly awkward, we majorize them by Radon measures ψ_x $(x \in \mathbf{R}^d)$ defined by

$$\psi_x(E) \quad = \quad \sum_{k=0}^{n-1} \frac{\mathcal{L}^d\big(E \cap (Q_k(x) \setminus Q_{k+1}(x))\big)}{\mathcal{L}^d\big(Q_k(x) \setminus Q_{k+1}(x)\big)} \omega_k(x).$$

To see the majorization, we first note that $\omega_k(x) \neq 0$ implies that $\omega_k(x) = 1$ and $\mu(Q_k(x)) \le C_1 2^{-k\alpha}$, which proves that

Lemma 8. *For every $x \in \mathbf{R}^d$ and every $k \in \{0,1,\ldots,n-1\}$,*

$$\omega_k(x) 2^{k\alpha} \mu(Q_k(x)) \le C_1\ .$$

Consequently, $|H_k(x)| \le \omega_k(x) \int_{Q_k(x)} |H_k(x,u)|\,d\mu(u) \le \omega_k(x) 2^{k\alpha} C_2 \mu(Q_k(x)) \le \omega_k(x) C_1 C_2$, and we infer that

Lemma 9. *For every $x \in \mathbf{R}^d$ and every Borel set $E \subseteq \mathbf{R}^d$,*

$$|\varphi_x(E)| \le \psi_x(E)\ .$$

In addition to the above auxiliary functions and measures, we need to decompose the family of dyadic cubes into several subfamilies. We denote by

$$\mathcal{I} \quad = \quad \mathcal{Q}_0 \cup \bigcup_{k=1}^{n-1} \{Q \in \mathcal{Q}_k : \mu(Q) \ge C_2^{-1}\sigma 2^{-k\alpha}\}$$

the family of those dyadic cubes which have a non-negligible μ measure (or belong to the starting grid).

Noting that $Q_0(x) \in \mathcal{I}$ for every $x \in \mathbf{R}^d$, we define

$$k(x) \quad = \quad \max\{k = 0,1,\ldots,n-1 : Q_k(x) \in \mathcal{I}\}$$

and, for $I \in \mathcal{Q}_m$,

$$G(I) \quad = \quad \{x \in I : k(x) = m\}.$$

It will also be convenient to extend the notation $Q_k(x)$ in the following way: If $I \in \mathcal{Q}_m$ and $0 \le k \le m$, then $Q_k(I)$ will denote the unique cube from $\mathcal{Q}_k$ containing I.

Finally, we choose the smallest integer q such that $q^2 \geq n$, define, for $i \in \{0, 1, \ldots, q-1\}$,

$$\mathcal{I}_i \quad = \quad \mathcal{I} \cap \bigcup_{k=iq}^{(i+1)q-1} \mathcal{Q}_k,$$

and denote by $\mathcal{K}_i$ the set of maximal cubes from the set $\mathcal{I} \cap \bigcup_{k=iq}^{n-1} \mathcal{Q}_k = \bigcup_{j=i}^{q-1} \mathcal{I}_j$, i.e.,

$$\mathcal{K}_i = \Big\{ K \in \mathcal{I} \cap \bigcup_{k=iq}^{n-1} \mathcal{Q}_k : K \subseteq I \in \bigcup_{k=iq}^{n-1} \mathcal{Q}_k,\, K \neq I \text{ implies } I \notin \mathcal{I} \Big\} .$$

Lemma 10. *If $K \in \mathcal{I}$, the sets $G(I)$ $(I \in \mathcal{I},\, I \subseteq K)$ form a (Borel measurable, finite) partition of K. Consequently, the union of the family $\mathcal{K}_i$ is partitioned into the cubes from $\mathcal{K}_i$ as well as into the sets $G(I)$ $(I \in \bigcup_{k=i}^{q-1} \mathcal{I}_k)$. In particular, the sets $G(I)$ $(I \in \mathcal{I})$ form a (Borel measurable, locally finite) partition of $\mathbf{R}^d$.*

Proof. Suppose that $K \in \mathcal{I} \cap \mathcal{Q}_m$. If $x \in K$, then $k(x) \geq m$, so $I = Q_{k(x)}(x) \subseteq K$, $I \in \mathcal{I}$ and $x \in G(I)$. Hence $K \subseteq \bigcup_{I \in \mathcal{I}, I \subseteq K} G(I)$ which, since $G(I) \subseteq I$, implies that $K = \bigcup_{I \in \mathcal{I}, I \subseteq K} G(I)$. Finally, $x \in G(I) \cap G(J)$ implies that $I = Q_{k(x)}(x) = J$, so $G(I) \cap G(J) = \emptyset$ for $I \neq J$. The bracketed statements about Borel measurability and finiteness are obvious since each $G(I)$ is a finite union of cubes from $\mathcal{Q}_{n-1}$ and the number of those $I \in \mathcal{Q}$ for which $I \subseteq K$ is finite.
Since the cubes from the family $\mathcal{K}_i$ are clearly disjoint, their union is partitioned by them as well as by the sets $G(I)$, where I runs through cubes from $\mathcal{I}$ that are subsets of cubes from $\mathcal{K}_i$. But each such I belongs to $\bigcup_{k=i}^{q-1} \mathcal{I}_k$ and, conversely, each $I \in \bigcup_{k=i}^{q-1} \mathcal{I}_k$ belongs to $\mathcal{I}$, hence, by definition, it is a subset of some cube from $\mathcal{K}_i$. This proves the second statement and, since the union of $\mathcal{K}_0 = \mathcal{Q}_0$ is $\mathbf{R}^d$, also the last one. $\square$

We are now ready to start our main estimates. Note that, according to Lemma 10,

$$\varphi_x(\mathbf{R}^d) \quad = \quad \sum_{I \in \mathcal{I}} \varphi_x(G(I))$$

and

$$\sum_{j=i+1}^{q-1} \sum_{J \in \mathcal{I}_j} \varphi_x(G(J)) \quad = \quad \sum_{K \in \mathcal{K}_{i+1}} \varphi_x(K).$$

Hence, using that $|\varphi_x(E)| \leq \psi_x(E)$ by Lemma 9, we may estimate

$$\begin{aligned} &\int \big(\varphi_x(\mathbf{R}^d)\big)^2 \, d\mu(x) \\ &= \sum_{I,J \in \mathcal{I}} \int \varphi_x(G(I))\, \varphi_x(G(J))\, d\mu(x) \\ &= \sum_{i=0}^{q-1} \sum_{I,J \in \mathcal{I}_i} \int \varphi_x(G(I))\, \varphi_x(G(J))\, d\mu(x) \end{aligned}$$

$$+\,2\sum_{i=0}^{q-2}\sum_{j=i+1}^{q-1}\sum_{I\in\mathcal{I}_i}\sum_{J\in\mathcal{I}_j}\int\varphi_x(G(I))\,\varphi_x(G(J))\,d\mu(x)$$

$$\le\ \sum_{i=0}^{q-1}\sum_{I,J\in\mathcal{I}_i}\int\psi_x(G(I))\,\psi_x(G(J))\,d\mu(x) \tag{19}$$

$$+\,2\sum_{i=0}^{q-2}\sum_{I\in\mathcal{I}_i}\sum_{K\in\mathcal{K}_{i+1}}\int\varphi_x(G(I))\,\varphi_x(K)\,d\mu(x). \tag{20}$$

For $i\ge 1$ we divide the integral $\int\psi_x(G(I))\,\psi_x(G(J))\,d\mu(x)$ in (19) into integrals over $Q_{(i-1)q}(I)\cap Q_{(i-1)q}(J)$ and $\mathbf{R}^d\setminus(Q_{(i-1)q}(I)\cap Q_{(i-1)q}(J))$ and, using that ψ_x is a non-negative measure, estimate the second integral by twice the integral over $\mathbf{R}^d\setminus Q_{(i-1)q}(I)$. Similarly, we divide the integral $\int\varphi_x(G(I))\,\varphi_x(K)\,d\mu(x)$ in (20) into integrals over $\mathbf{R}^d\setminus K$ and K and replace φ_x by ψ_x in the first, but then estimate the second integral in a slightly more sophisticated way by

$$\begin{aligned}
&\int_K\varphi_x(G(I))\,\varphi_x(K)\,d\mu(x)\\
&\quad=\ \int_K\big(\varphi_x(G(I))-\varphi_{\zeta(K)}(G(I))\big)\,\varphi_x(K)\,d\mu(x)\\
&\qquad+\int_K\varphi_{\zeta(K)}(G(I))\,\varphi_x(K)\,d\mu(x)\\
&\quad\le\ \sup_{z\in K}\big|\varphi_z(G(I))-\varphi_{\zeta(K)}(G(I))\big|\int_K\psi_x(K)\,d\mu(x)\\
&\qquad+\varphi_{\zeta(K)}(G(I))\int_K\varphi_x(K)\,d\mu(x),
\end{aligned}$$

where $\zeta(K)$ denotes the centre of K. Thus

$$\int\big(\varphi_x(\mathbf{R}^d)\big)^2\,d\mu(x)$$

$$\le\ \sum_{I,J\in\mathcal{I}_0}\int\psi_x(G(I))\,\psi_x(G(J))\,d\mu(x) \tag{21}$$

$$+\sum_{i=1}^{q-1}\sum_{I,J\in\mathcal{I}_i}\int_{Q_{(i-1)q}(I)\cap Q_{(i-1)q}(J)}\psi_x(G(I))\,\psi_x(G(J))\,d\mu(x) \tag{22}$$

$$+\,2\sum_{i=1}^{q-1}\sum_{I,J\in\mathcal{I}_i}\int_{\mathbf{R}^d\setminus Q_{(i-1)q}(I)}\psi_x(G(I))\,\psi_x(G(J))\,d\mu(x) \tag{23}$$

$$+\,2\sum_{i=0}^{q-2}\sum_{I\in\mathcal{I}_i}\sum_{K\in\mathcal{K}_{i+1}}\int_{\mathbf{R}^d\setminus K}\psi_x(G(I))\,\psi_x(K)\,d\mu(x) \tag{24}$$

$$+\,2\sum_{i=0}^{q-2}\sum_{I\in\mathcal{I}_i}\sum_{K\in\mathcal{K}_{i+1}}\sup_{z\in K}\big|\varphi_z(G(I))-\varphi_{\zeta(K)}(G(I))\big|\int_K\psi_x(K)d\mu(x) \tag{25}$$

$$+ 2\sum_{i=0}^{q-2}\sum_{I\in\mathcal{I}_i}\sum_{K\in\mathcal{K}_{i+1}} \varphi_{\zeta(K)}(G(I)) \int_K \varphi_x(K)\,d\mu(x). \tag{26}$$

Taking into account Lemma 7, we conclude that to finish the proof of Lemma 5 it suffices to estimate each of the terms (21-26) by $\mathcal{C}n^{3/2}\mu(\mathbf{R}^d)$, or by $\mathcal{C}q^3\mu(\mathbf{R}^d)$, since $n \le q^2 \le (\sqrt{n}+1)^2 \le 4n$ implies $n^{3/2} \le q^3 \le 8n^{3/2}$. This will be done, term by term, in a series of lemmas. The first of them follows immediately from the definitions.

Lemma 11. *If* $0 \le i < j \le n$, *then*

$$\psi_x(Q_i(x)\setminus Q_j(x)) = \sum_{k=i}^{j-1}\omega_k(x) \le j-i.$$

In particular,

$$\psi_x(\mathbf{R}^d) = \psi_x(Q_0(x)\setminus Q_n(x)) \le n \le q^2\ .$$

Lemma 12. *Whenever* $0 \le l < m \le n$ *and* $x \in I \in \mathcal{Q}_l$, *then*

$$\psi_x(Q_m(x)\cap G(I)) = 0\ .$$

Proof. Since $\psi_x(Q_m(x)\cap G(I)) \le \sum_{k=m}^{n-1}\omega_k(x)$, the statement is obvious if $m = n$ or $\omega_k(x) = 0$ for all $m \le k \le n-1$. On the other hand, if $p \ge m$ is the least index for which $\omega_p(x) \ne 0$, then $Q_p(z) = Q_p(x) \in \mathcal{I}$ for all $z \in Q_p(x)$, so $k(z) \ge p$ for all such z and we infer that $Q_p(x) \cap G(I) = \emptyset$. Hence $\psi_x(Q_m(x)\cap G(I)) \le \sum_{k=m}^{p-1}\omega_k(x) = 0$, since $\omega_k(x) = 0$ for all $m \le k < p$. □

Lemma 13. *If* $0 \le l \le i < j \le n$ *and* $\mathcal{J} \subseteq \mathcal{I} \cap \bigcup_{k=i}^{j-1}\mathcal{Q}_k$, *then*

$$\sum_{I\in\mathcal{J},x\in Q_l(I)} \psi_x(G(I)) \le 2(j-l)\ .$$

Consequently,

$$\sum_{I,J\in\mathcal{J}} \int_{Q_l(I)\cap Q_l(J)} \psi_x(G(I))\,\psi_x(G(J))\,d\mu(x) \le 4(j-l)^2\mu(\mathbf{R}^d)\ .$$

Proof. If $I \in \mathcal{J}$ and $x \in Q_l(I)\setminus I$, then $G(I) \subseteq I \subseteq Q_l(x)\setminus Q_j(x)$. Using also the disjointness of $G(I)$ $(I \in \mathcal{I})$ (Lemma 10) and Lemma 11, we have

$$\sum_{I\in\mathcal{J},x\in Q_l(I)\setminus I} \psi_x(G(I)) \le \psi_x(Q_l(x)\setminus Q_j(x)) \le j-l.$$

Similarly, but using first Lemma 12, we get

$$\sum_{I\in\mathcal{J},x\in I} \psi_x(G(I)) = \sum_{I\in\mathcal{J},x\in I} \psi_x(G(I)\setminus Q_j(x)) \le \psi_x(Q_l(x)\setminus Q_j(x)) \le j-l.$$

Adding these two inequalities gives the first statement. The second statement follows, since

$$\sum_{I,J\in\mathcal{J}}\int_{Q_l(I)\cap Q_l(J)}\psi_x(G(I))\,\psi_x(G(J))\,d\mu(x)$$
$$=\int\Big(\sum_{I\in\mathcal{J},x\in Q_l(I)}\psi_x(G(I))\Big)^2 d\mu(x)\le 4(j-l)^2\mu(\mathbf{R}^d).$$

□

Estimate of (21). From Lemma 13 with $\mathcal{J}=\mathcal{I}_0$ we infer

$$\sum_{I,J\in\mathcal{I}_0}\int\psi_x(G(I))\,\psi_x(G(J))\,d\mu(x)\le 4q^3\mu(\mathbf{R}^d)\ .$$

Estimate of (22). Using Lemma 13, we see that

$$\sum_{i=1}^{q-1}\sum_{I,J\in\mathcal{I}_i}\int_{Q_{(i-1)q}(I)\cap Q_{(i-1)q}(J)}\psi_x(G(I))\psi_x(G(J))\,d\mu(x)$$
$$\le\sum_{i=0}^{q-1}4(2q)^2\mu(\mathbf{R}^d)=16q^3\mu(\mathbf{R}^d)\ .$$

Lemma 14. *If $0\le l<k\le n$ and $I\in\mathcal{I}\cap\mathcal{Q}_k$, then*

$$\int_{Q_l(I)\setminus Q_{l+1}(I)}\psi_x(I)\,d\mu(x)\le 2C_1C_2 2^{(l-k)(d-\alpha)}\mu(I)/\sigma\ .$$

Proof. We first note that $\mu(I)\ge 2^{-k\alpha}C_2^{-1}\sigma$ since $k>0$ and $I\in\mathcal{I}\cap\mathcal{Q}_k$. If $x\in Q_l(I)\setminus Q_{l+1}(I)$ then $I\cap(Q_i(x)\setminus Q_{i+1}(x))=\emptyset$ if $i\ne l$ and $I\cap(Q_l(x)\setminus Q_{l+1}(x))=I$. Hence

$$\begin{aligned}\psi_x(I)&=\mathcal{L}^d(I)(2^{-ld}-2^{-(l+1)d})^{-1}\omega_l(x)\le 2^{-kd}2^{ld+1}\omega_l(x)\\&\le 2^{-k(d-\alpha)}2^{ld+1}C_2\mu(I)\omega_l(x)/\sigma.\end{aligned}$$

Since, according to Lemma 8,

$$\int_{Q_l(I)\setminus Q_{l+1}(I)}\omega_l(x)\,d\mu(x)\le C_1 2^{-l\alpha}\ ,$$

we infer that

$$\begin{aligned}&\int_{Q_l(I)\setminus Q_{l+1}(I)}\psi_x(I)\,d\mu(x)\\&\le 2^{-k(d-\alpha)}2^{ld+1}C_2\mu(I)\int_{Q_l(I)\setminus Q_{l+1}(I)}\omega_l(x)\,d\mu(x)/\sigma\\&\le 2C_1C_2 2^{(l-k)(d-\alpha)}\mu(I)/\sigma.\end{aligned}$$

□

Lemma 15. *If $0 \le m \le k \le n$ and $I \in \mathcal{I} \cap \mathcal{Q}_k$, then*

$$\int_{\mathbf{R}^d \setminus Q_m(I)} \psi_x(I)\, d\mu(x) \le 2C_1 C_2 2^{(m-k)(d-\alpha)} \mu(I)/(\sigma(2^{d-\alpha}-1)) \,.$$

In particular,

$$\int_{\mathbf{R}^d \setminus I} \psi_x(I)\, d\mu(x) \le 2C_1 C_2 \mu(I)/(\sigma(2^{d-\alpha}-1)) \,.$$

Proof. Since $\psi_x(\mathbf{R}^d \setminus Q_0(x)) = 0$, the first inequality follows by adding the inequalities from Lemma 14 for $l = 0, 1, \dots, m-1$. The second inequality is the case $m = k$ of the first one. □

Estimate of (23). Using that $\sum_{J\in\mathcal{I}_i} \psi_x(G(J)) \le n \le q^2$ according to Lemma 11 and that according to Lemma 15, for any $I \in \mathcal{I}_i$,

$$\int_{\mathbf{R}^d \setminus Q_{(i-1)q}(I)} \psi_x(G(I))\, d\mu(x) \le \mathcal{C}_1 2^{-q(d-\alpha)} \mu(I),$$

where $\mathcal{C}_1 = 2C_1C_2/(\sigma(2^{d-\alpha}-1))$, we estimate that

$$\begin{aligned}
&\sum_{i=1}^{q-1} \sum_{I,J\in\mathcal{I}_i} \int_{\mathbf{R}^d \setminus Q_{(i-1)q}(I)} \psi_x(G(I))\, \psi_x(G(J))\, d\mu(x) \\
&\quad\le\ q^2 \sum_{i=1}^{q-1} \sum_{I\in\mathcal{I}_i} \int_{\mathbf{R}^d \setminus Q_{(i-1)q}(I)} \psi_x(G(I))\, d\mu(x) \\
&\quad\le\ \mathcal{C}_1 2^{-q(d-\alpha)} q^2 \sum_{i=1}^{q-1} \sum_{I\in\mathcal{I}_i} \mu(I) \\
&\quad\le\ \mathcal{C}_1 2^{-q(d-\alpha)} q^4 \mu(\mathbf{R}^d) \\
&\quad\le\ \mathcal{C}_2 q^3 \mu(\mathbf{R}^d),
\end{aligned}$$

where $\mathcal{C}_2 = \mathcal{C}_1 \sup_{m\ge 1} 2^{-m(d-\alpha)} m$.

Estimate of (24). We first note that the disjointness of the cubes from $\mathcal{K}_{i+1}$ and the second inequality of Lemma 15 imply that

$$\sum_{K\in\mathcal{K}_{i+1}} \int_{\mathbf{R}^d \setminus K} \psi_x(K)\, d\mu(x) \le \mathcal{C}_1 \sum_{K\in\mathcal{K}_{i+1}} \mu(K) \le \mathcal{C}_1 \mu(\mathbf{R}^d) \,.$$

Thus, using first that $\sum_{I\in\mathcal{I}_i} \psi_x(G(I)) \le q^2$, we have

$$\begin{aligned}
&\sum_{i=0}^{q-2} \sum_{I\in\mathcal{I}_i} \sum_{K\in\mathcal{K}_{i+1}} \int_{\mathbf{R}^d \setminus K} \psi_x(G(I))\, \psi_x(K)\, d\mu(x) \\
&\quad\le\ q^2 \sum_{i=0}^{q-2} \sum_{K\in\mathcal{K}_{i+1}} \int_{\mathbf{R}^d \setminus K} \psi_x(K)\, d\mu(x)
\end{aligned}$$

$$\begin{aligned} &\leq \quad q^2 \sum_{i=0}^{q-2} \mathcal{C}_1 \mu(\mathbf{R}^d) \\ &\leq \quad \mathcal{C}_1 q^3 \mu(\mathbf{R}^d)\,. \end{aligned}$$

Lemma 16. *For any dyadic cube K,*

$$\sum_{I \in \mathcal{I}} \sup_{x,y \in K} \big|\varphi_x(G(I) \setminus K) - \varphi_y(G(I) \setminus K)\big| \leq 1.$$

Proof. If $x, y \in K \in \mathcal{Q}_m$ and $0 \leq k < m$, then $Q_k(x) = Q_k(y)$, $Q_{k+1}(x) = Q_{k+1}(y)$ and $\omega_k(x) = \omega_k(y)$. Using also that $\omega_k(x) 2^{k\alpha} \mu(Q_k(x)) \leq C_1$ by Lemma 8 and that, for all $u \in Q_k(x)$, $|H_k(y,u) - H_k(x,u)| \leq C_2 2^{k\alpha+k-m}$ by assumption, we infer that

$$\begin{aligned} |H_k(y) - H_k(x)| \quad &\leq \quad \omega_k(x) \int_{Q_k(x)} |H_k(y,u) - H_k(x,u)|\, d\mu(u) \\ &\leq \quad C_2 \omega_k(x) 2^{k\alpha} \mu(Q_k(x)) 2^{k-m} \\ &\leq \quad C_1 C_2 2^{k-m}. \end{aligned}$$

Hence

$$\begin{aligned} &|\varphi_x(G(I) \setminus K) - \varphi_y(G(I) \setminus K)| \\ &\leq \quad (C_1 C_2)^{-1} \sum_{k=0}^{m-1} \frac{\mathcal{L}^d\big(G(I) \cap (Q_k(x) \setminus Q_{k+1}(x))\big)}{\mathcal{L}^d\big(Q_k(x) \setminus Q_{k+1}(x)\big)} C_1 C_2 2^{k-m} \end{aligned}$$

and, taking supremum over $x, y \in K$, we have

$$\begin{aligned} &\sup_{x,y \in K} |\varphi_x(G(I) \setminus K) - \varphi_y(G(I) \setminus K)| \\ &\leq \quad \sum_{k=0}^{m-1} \frac{\mathcal{L}^d\big(G(I) \cap (Q_k(\zeta(K)) \setminus Q_{k+1}(\zeta(K)))\big)}{\mathcal{L}^d\big(Q_k(\zeta(K)) \setminus Q_{k+1}(\zeta(K))\big)} 2^{k-m}. \end{aligned}$$

Thus, summing over $I \in \mathcal{I}$ and using the disjointness of the sets $G(I)$, we get

$$\sum_{I \in \mathcal{I}} \sup_{x,y \in K} |\varphi_x(G(I) \setminus K) - \varphi_y(G(I) \setminus K)| \leq \sum_{k=0}^{m-1} 2^{k-m} \leq 1.$$

$\square$

Estimate of (25). To use Lemma 16, we observe that, whenever $u \in K \in \mathcal{K}_{i+1}$ and $I \in \bigcup_{j=0}^{i} \mathcal{I}_j$, then $k(u) > i$ and hence $u \notin G(I)$. Consequently, $\varphi_x(G(I)) = \varphi_x(G(I) \setminus K)$ for every $x \in K$ and, using first that $\int_K \psi_x(K)\, d\mu(x) \leq q^2 \mu(K)$, since $\psi_x(K) \leq \psi_x(\mathbf{R}^d) \leq q^2$ by Lemma 11, then Lemma 16 and finally the disjointness of $\mathcal{K}_{i+1}$, we infer that

$$\begin{aligned}
&\sum_{i=0}^{q-2}\sum_{I\in\mathcal{I}_i}\sum_{K\in\mathcal{K}_{i+1}}\sup_{z\in K}\left|\varphi_z(G(I))-\varphi_{\zeta(K)}(G(I))\right|\int_K\psi_x(K)\,d\mu(x)\\
&\quad\le\; q^2\sum_{i=0}^{q-2}\sum_{K\in\mathcal{K}_{i+1}}\mu(K)\sum_{j=0}^{i}\sum_{I\in\mathcal{I}_j}\sup_{z\in K}\left|\varphi_z(G(I))-\varphi_{\zeta(K)}(G(I))\right|\\
&\quad=\; q^2\sum_{i=0}^{q-2}\sum_{K\in\mathcal{K}_{i+1}}\mu(K)\sum_{j=0}^{i}\sum_{I\in\mathcal{I}_j}\sup_{z\in K}\left|\varphi_z(G(I)\setminus K)-\varphi_{\zeta(K)}(G(I)\setminus K)\right|\\
&\quad\le\; q^2\sum_{i=0}^{q-2}\sum_{K\in\mathcal{K}_{i+1}}\mu(K)\\
&\quad\le\; q^3\mu(\mathbf{R}^d).
\end{aligned}$$

Lemma 17. *For every dyadic cube K,*

$$\int_K\varphi_x(K)\,d\mu(x)=0\ .$$

Proof. If $Q\in\mathcal{Q}_k$, the anti-symmetry of H_k and the fact that $\omega_k(x)=\omega_k(\zeta(Q))$ for every $x\in Q$ imply that

$$\int_Q H_k(x)\,d\mu(x)\;=\;\omega_k(\zeta(Q))\int_Q\int_Q H_k(x,u)\,d\mu(x)\,d\mu(u)=0.$$

Hence, if $K\in\mathcal{Q}_m$,

$$\begin{aligned}
\int_K\varphi_x(K)\,d\mu(x)&=(C_1C_2)^{-1}\sum_{k=m}^{n-1}\int_K H_k(x)\,d\mu(x)\\
&=(C_1C_2)^{-1}\sum_{k=m}^{n-1}\sum_{Q\in\mathcal{Q}_k,Q\subseteq K}\int_Q H_k(x)\,d\mu(x)\\
&=0.
\end{aligned}$$

□

Estimate of (26). By Lemma 17,

$$\sum_{i=0}^{q-2}\sum_{I\in\mathcal{I}_i}\sum_{K\in\mathcal{K}_{i+1}}\varphi_{\zeta(K)}(G(I))\int_K\varphi_x(K)\,d\mu(x)=0\ .$$

4. Proof of the Palm formula

Noting that trivial cases have been handled in Proposition 5, we assume, for the rest of this section, that $0<\alpha<d$ and a measure $\mu\in\mathcal{M}(\mathbf{R}^d)$ are fixed.

We will need to make some, not very difficult but careful, estimates of measures of sets. This will be facilitated by defining the usual distances $F_k(\nu,\tilde\nu)$ of measures ν

and $\tilde\nu$ as the supremum of $|\int f(w)\,d\nu(w) - \int f(w)\,d\tilde\nu(w)|$, where f runs through functions $f : \mathbf{R}^d \to \mathbf{R}$ having support in $B(0,k)$ and Lipschitz constant at most one. It is not difficult to see that $(F_k)_{k=1}^{\infty}$ is a sequence of pseudo-metrics inducing the vague topology of $\mathcal{M}(\mathbf{R}^d)$ (see e.g. [Mat, 14.13]). In addition, we will only need the following simple estimate of the distance of enlargements of measures.

Lemma 18. *Whenever $x, \tilde x \in \mathbf{R}^d$ and $r > 0$, then*

$$F_k\Big(\frac{\mu_{x,r}}{r^\alpha}, \frac{\mu_{\tilde x,r}}{r^\alpha}\Big) \le 2\frac{\mu(B(x, \|x-\tilde x\| + kr))}{r^\alpha}\,\frac{\|x-\tilde x\|}{r}.$$

Proof. If f is as in the definition of F_k then

$$\begin{aligned}
&\Big|\int f(w)\,d\mu_{x,r}(w) - \int f(w)\,d\mu_{\tilde x,r}(w)\Big| \\
&= \int \Big| f\Big(\frac{w-x}{r}\Big) - f\Big(\frac{w-\tilde x}{r}\Big)\Big|\,d\mu(w) \\
&\le \int_{B(x,kr)} \frac{\|x-\tilde x\|}{r}\,d\mu(w) + \int_{B(\tilde x,kr)} \frac{\|x-\tilde x\|}{r}\,d\mu(w) \\
&\le 2\mu(B(x,\|x-\tilde x\| + kr))\frac{\|x-\tilde x\|}{r},
\end{aligned}$$

which, divided by r^α, proves the statement. □

Before coming to our last lemma, we need to introduce the distances also in the space $\mathcal{M}(\mathbf{R}^d) \times \mathbf{R}^d$. We define them by $\varrho_p((\nu,u),(\tilde\nu,\tilde u)) = F_p(\nu,\tilde\nu) + \|u-\tilde u\|$; they form a non-decreasing sequence of pseudo-metrics generating the topology of $\mathcal{M}(\mathbf{R}^d) \times \mathbf{R}^d$. We will also use the usual notation $\varrho_p((\nu,u),M) = \inf_{(\tilde\nu,\tilde u)\in M} \varrho_p((\nu,u),(\tilde\nu,\tilde u))$ for the ϱ_p distance of $(\nu,u) \in \mathcal{M}(\mathbf{R}^d) \times \mathbf{R}^d$ from a set $M \subseteq \mathcal{M}(\mathbf{R}^d) \times \mathbf{R}^d$ and recall that the ϱ_p distance of sets $M, N \subseteq \mathcal{M}(\mathbf{R}^d) \times \mathbf{R}^d$ is defined as the infimum of $\varrho_p((\nu,u),(\tilde\nu,\tilde u))$, where $(\nu,u) \in M$ and $(\tilde\nu,\tilde u) \in N$.

Lemma 19. *Suppose that $R > 1$ and that $U \subseteq V \subseteq \mathcal{M}(\mathbf{R}^d) \times B(0,R-1)$ are open subsets of $\mathcal{M}(\mathbf{R}^d) \times \mathbf{R}$ such that, for some positive integer p, the ϱ_p distance of U from the complement of V is positive and the values of $\nu(B(0,5R+p+1))$ for $(\nu,u) \in U$ are bounded. Then, for μ-almost every $x \in \mathbf{R}^d$,*

$$\liminf_{\delta\downarrow 0} \iint \big[1_V(\nu,y) - 1_U(T^y\nu,-y)\big]\,d\nu(y)\,dP_\delta^x(\nu) \ge 0.$$

Proof. We choose $\xi \in (0,1)$ such that $\varrho_p((\nu,u),(\tilde\nu,\tilde u)) > \xi$ whenever $(\nu,u) \in U$ and $(\tilde\nu,\tilde u) \notin V$, and we let

$$A_1 = 1 + \sup_{(\nu,u)\in U} \nu(B(0,5R+p+1)), \qquad A_2 = (2A_1+1)/\xi\ ,$$

and observe that our assumptions imply that A_1 is finite. (It may be pointed out that the names of the constants are chosen with the future application of Lemma 4 in mind. Also, to avoid trivialities, note that in the case $U = \emptyset$ the statement of the lemma is obvious.)

Define $S : \mathbf{R}^d \times \mathbf{R}^d \times (0,\infty) \to \mathcal{M}(\mathbf{R}^d) \times \mathbf{R}^d$ by

$$S(x,y;r) = \Big(\frac{\mu_{x,r}}{r^\alpha}, \frac{y-x}{r}\Big) .$$

Since S is continuous, the set $S^{-1}(U)$ is open, which implies that the function

$$H(x,y;r) \quad = \quad \max\Big\{0, 1 - A_2 \inf_{(\tilde{x},\tilde{y};r)\in S^{-1}(U)} \frac{\|x-\tilde{x}\| + \|y-\tilde{y}\|}{r}\Big\}.$$

is upper semi-continuous, hence Borel measurable.

We prove that

$$H(x,y;r) \le 1_V\Big(\frac{\mu_{x,r}}{r^\alpha}, \frac{y-x}{r}\Big) \tag{27}$$

for all $(x,y;r) \in \mathbf{R}^d \times \mathbf{R}^d \times (0,\infty)$. Indeed, otherwise there would exist $(x,y;r)$ and $(\tilde{x},\tilde{y};r)$ such that $S(x,y;r) \notin V$, $S(\tilde{x},\tilde{y};r) \in U$, and $\|x-\tilde{x}\|+\|y-\tilde{y}\| < r/A_2$, and we would infer from $S(\tilde{x},\tilde{y};r) \in U$ that

$$\begin{aligned} \mu(B(\tilde{x}, \|\tilde{x}-x\| + pr)) &\le \mu(B(\tilde{x},(p+1)r)) = \mu_{\tilde{x},r}(B(0,(p+1))) \\ &\le \mu_{\tilde{x},r}(B(0, 5R+p+1)) < A_1 r^\alpha. \end{aligned}$$

But then Lemma 18 would imply that

$$\begin{aligned} \xi &< \varrho_p\Big(\Big(\frac{\mu_{x,r}}{r^\alpha}, \frac{y-x}{r}\Big), \Big(\frac{\mu_{\tilde{x},r}}{r^\alpha}, \frac{\tilde{y}-\tilde{x}}{r}\Big)\Big) \\ &\le 2\frac{\mu(B(\tilde{x}, \|\tilde{x}-x\| + pr))}{r^\alpha} \frac{\|x-\tilde{x}\|}{r} + \frac{\|x-\tilde{x}\| + \|y-\tilde{y}\|}{r} \\ &\le 2A_1/A_2 + 1/A_2 \\ &= \xi. \end{aligned}$$

Noting that, for any G,

$$\begin{aligned} \iint G(\nu,y)\, d\nu(y)\, dP_\delta^x(\nu) &= \frac{1}{|\log\delta|}\int_\delta^1 \int G\Big(\frac{\mu_{x,r}}{r^\alpha}, y\Big)\, \frac{d\mu_{x,r}(y)}{r^\alpha}\, \frac{dr}{r} \\ &= \frac{1}{|\log\delta|}\int_\delta^1 \int G\Big(\frac{\mu_{x,r}}{r^\alpha}, \frac{y-x}{r}\Big)\, d\mu(y)\, \frac{dr}{r^{1+\alpha}}, \end{aligned}$$

we integrate (27) first with respect to $d\mu(y)$ and then with respect to $dr/r^{1+\alpha}$ to find out that

$$\frac{1}{|\log\delta|}\int_\delta^1 \int H(x,y;r)\, d\mu(y)\, \frac{dr}{r^{1+\alpha}} \le \iint 1_V(\nu,y) d\nu(y)\, dP_\delta^x(\nu) \tag{28}$$

for every $x \in \mathbf{R}^d$ and every $\delta > 0$. In a similar way we find an estimate of the integral involving the set U: Since $H(y,x;r) = 1$ if $S(y,x;r) \in U$, we have

$$1_U\Big(\frac{\mu_{y,r}}{r^\alpha}, \frac{x-y}{r}\Big) \le H(y,x;r). \tag{29}$$

Hence, noting that, for any G,

$$\begin{aligned}\iint G(T^y\nu,-y)\,d\nu(y)\,dP^x_\delta(\nu) &= \frac{1}{|\log\delta|}\int_\delta^1\int G\Big(\frac{T^y\mu_{x,r}}{r^\alpha},-y\Big)\,\frac{d\mu_{x,r}(y)}{r^\alpha}\,\frac{dr}{r}\\ &= \frac{1}{|\log\delta|}\int_\delta^1\int G\Big(\frac{\mu_{y,r}}{r^\alpha},\frac{x-y}{r}\Big)\,d\mu(y)\,\frac{dr}{r^{1+\alpha}},\end{aligned}$$

we obtain, integrating (29) first with respect to $d\mu(y)$ and then with respect to $dr/r^{1+\alpha}$, that

$$\iint 1_U(T^y\nu,-y)\,d\nu(y)\,dP^x_\delta(\nu) \le \frac{1}{|\log\delta|}\int_\delta^1\int H(y,x;r)\,d\mu(y)\,\frac{dr}{r^{1+\alpha}} \tag{30}$$

for every $x\in\mathbf{R}^d$ and every $\delta>0$.

We prove that $H(x,y;r)$ satisfies the assumptions of Lemma 4. We have already noted that H is Borel measurable, and it is clear that its range is contained in $[0,1]$. If $H(x,y;r)\neq 0$, then we find $(\tilde x,\tilde y;r)\in S^{-1}(U)$ such that $\|x-\tilde x\|+\|y-\tilde y\|\le r/A_2\le r$ and infer from $S(\tilde x,\tilde y;r)\in U$ that $\|\tilde x-\tilde y\|<(R-1)r$ and $\mu_{\tilde x,r}(B(0,5R+1))<A_1r^\alpha$. Hence

$$\|x-y\|\le\|x-\tilde x\|+\|y-\tilde y\|+\|\tilde x-\tilde y\|<r+(R-1)r=Rr\ ,$$

which proves 4(i) , and

$$\mu(B(x,5Rr))\le\mu(B(\tilde x,(5R+1)r))=\mu_{\tilde x,r}(B(0,5R+1))<A_1r^\alpha\ ,$$

which proves 4(ii) . Finally, the requirement 4(iii) follows directly from the definition of $H(x,y;r)$.

This finishes the proof of the Lemma, since (30) and (28) and Lemma 4 imply that

$$\begin{aligned}&\liminf_{\delta\downarrow 0}\iint\big[1_V(\nu,y)-1_U(T^y\nu,-y)\big]\,d\nu(y)\,dP^x_\delta(\nu)\\ \ge\ &\liminf_{\delta\to 0}\frac{1}{|\log\delta|}\int_\delta^1\int\frac{H(x,y;r)-H(y,x;r)}{r^\alpha}\,d\mu(y)\,\frac{dr}{r}\\ =\ &0.\end{aligned}$$

□

Proof of Theorem 1

Using that $\mathcal{M}(\mathbf{R}^d)\times\mathbf{R}^d$ is a separable metric space, we find a countable basis $\mathcal{U}$ of open sets which is closed under finite unions. Let $\tilde{\mathcal{U}}$ be the set of pairs (U,V) of sets from $\mathcal{U}$ for which there are $R>1$ and a positive integer p such that $U\subseteq V\subseteq\mathcal{M}(\mathbf{R}^d)\times B(0,R-1)$, the ϱ_p distance of U from the complement of V is positive and the values of $\nu(B(0,5R+p+1))$ for $(\nu,u)\in U$ are bounded. Since $\tilde{\mathcal{U}}$ is countable, we infer from Lemma 19 that there is a μ null set $N\subseteq\mathbf{R}^d$ such that for every $x\in\mathbf{R}^d\setminus N$,

$$\liminf_{\delta\downarrow 0}\iint\big[1_V(\nu,y)-1_U(T^y\nu,-y)\big]\,d\nu(y)\,dP^x_\delta(\nu)\ge 0 \tag{31}$$

for every pair $(U,V)\in\tilde{\mathcal{U}}$. We now need one more technical lemma from measure theory.

Lemma 20. *Suppose that X is a locally compact Polish space and that P_k are Radon measures on $\mathcal{M}(X)$ vaguely converging to P_0. Let Λ_k be Borel measures on $\mathcal{M}(X) \times X$ defined by*

$$\Lambda_k(E) = \int\!\!\int 1_E(\nu, y)\, d\nu(y)\, dP_k(\nu)\ .$$

Then Λ_k are Radon measures converging vaguely to Λ_0.

Proof. If $K \subseteq \mathcal{M}(X) \times X$ is compact, let $K_1 \subseteq \mathcal{M}(X)$ and $K_2 \subseteq X$ be compact sets such that $K \subseteq K_1 \times K_2$ and let $L_2 \supseteq K_2$ be an open set contained in a compact set $K_3 \subseteq X$. Then there is $0 < c < \infty$ such that $\nu(K_3) < c$ for all $\nu \in K_1$. Given any $\eta > 0$, we use Lemma A3 to find an open set $L_1 \supseteq K_1$ such that

$$\limsup_{k\to\infty} \left| \int \varphi(\nu)\, dP_k(\nu) - \int \varphi(\nu)\, dP_0(\nu) \right| < \eta/c \tag{32}$$

for every continuous function $\varphi : \mathcal{M}(X) \to [0,1]$ such that $\varphi(\nu) = 0$ for $\nu \notin L_1$; intersecting L_1 with the open set $\{\nu : \nu(K_3) < c\}$, if necessary, we may also assume that $\nu(L_2) < c$ for all $\nu \in L_1$. As a byproduct we note that $K_1 \times L_2$ contains K and $\Lambda_k(K_1 \times L_2) < \infty$, so that Λ_k are Radon measures. Denote $L = L_1 \times L_2$. If $\psi : \mathcal{M}(X) \times X \to [0,1]$ is a continuous function such that $\psi(\nu, y) = 0$ if $(\nu, y) \notin L$, we note that the function $\varphi(\nu) = \frac{1}{c} \int \psi(\nu, y)\, d\nu(y)$ is non-negative, continuous, bounded by one, and satisfies $\varphi(\nu) = 0$ for $\nu \notin L_1$. Thus (32) implies that

$$\limsup_{k\to\infty} \left| \int \psi(\nu, y)\, d\Lambda_k(\nu, y) - \int \psi(\nu, y)\, d\Lambda_0(\nu, y) \right| < \eta\ ,$$

which, according to Lemma A3, implies that (Λ_k) vaguely converges to Λ_0. $\square$

Let $x \in \mathbf{R}^d \setminus N$ and $P \in \mathcal{P}^\alpha(\mu, x)$ be fixed. Let $\delta_k \downarrow 0$ be such that $P^x_{\delta_k}$ converge weakly to P. Then

$$\Lambda_k(E) \quad = \quad \int\!\!\int 1_E(\nu, y)\, d\nu(y) dP^x_{\delta_k}(\nu)$$

are, according to the lemma, Radon measures on $\mathcal{M}(\mathbf{R}^d) \times \mathbf{R}^d$ converging vaguely to the Radon measure

$$\Lambda(E) \quad = \quad \int\!\!\int 1_E(\nu, y)\, d\nu(y) dP(\nu).$$

Moreover, since the mapping $\tilde{T} : \mathcal{M}(\mathbf{R}^d) \times \mathbf{R}^d \to \mathcal{M}(\mathbf{R}^d) \times \mathbf{R}^d$ defined by $\tilde{T}(\nu, y) = (T^y \nu, -y)$ is a homeomorphism, the measures $\tilde{\Lambda}_k = \Lambda_k \circ \tilde{T}^{-1}$ converge vaguely to $\tilde{\Lambda} = \Lambda \circ \tilde{T}^{-1}$.

We intend to use Lemma A5 to show that $\tilde{\Lambda} \le \Lambda$. Whenever $K \subseteq L$ are subsets of $\mathcal{M}(\mathbf{R}^d) \times \mathbf{R}^d$ such that K is compact and L is open, we first choose $R > 2$ such that $K \subseteq \mathcal{M}(\mathbf{R}^d) \times B(0, R-2)$ and find $V \in \mathcal{U}$ such that $K \subseteq V \subseteq L \cap \mathcal{M}(\mathbf{R}^d) \times B(0, R-1)$. Then we choose a positive integer p such that the ϱ_p distance of K from the complement of V is positive. We let ξ denote this distance, define

$$A = 1 + \sup_{(\nu,u)\in K} \nu(B(0, 5R + p + 1))$$

and

$$\hat{U} = \{(\nu, u) \in V : \varrho_p((\nu, u), K) < \xi/2, \nu(B(0, 5R + p + 1)) < A\},$$

and choose $U \in \mathcal{U}$ such that $K \subseteq U \subseteq \hat{U}$. Then $(U, V) \in \tilde{\mathcal{U}}$, and we conclude from (31) that

$$\begin{aligned}
\liminf_{k\to\infty} \tilde{\Lambda}_k(U) &= \liminf_{k\to\infty} \int\int 1_U(T^y\nu, -y)\, d\nu(y)\, dP^x_{\delta_k}(\nu) \\
&\le \liminf_{k\to\infty} \int\int 1_V(\nu, y)\, d\nu(y)\, dP^x_{\delta_k}(\nu) \\
&= \liminf_{k\to\infty} \Lambda_k(V) \\
&\le \limsup_{k\to\infty} \Lambda_k(L).
\end{aligned}$$

By Lemma A5, this shows that $\tilde{\Lambda} \le \Lambda$. But then $\Lambda = \tilde{\Lambda} \circ \tilde{T}^{-1} \le \Lambda \circ \tilde{T}^{-1} = \tilde{\Lambda}$, and we conclude that $\tilde{\Lambda} = \Lambda$, which is the statement of Theorem 1.

Appendix: Measure theoretical definitions and preliminaries

In this appendix we recall some notions concerning Radon measures on Polish spaces. Although we use the theory only for the two cases of (possibly infinite) measures on a locally compact space and sub-probability measures on a (non-locally compact) space we think it is worth giving a unified treatment of the two situations.

If X is a separable and completely metrizable topological space, in short a *Polish space*, a non-negative Borel measure ν is termed a *Radon measure* if every compact subset K of X has finite measure and

$$\nu(B) = \sup\{\nu(K) : K \subseteq B,\ K \text{ compact}\} \text{ for every Borel set } B \subseteq X.$$

We shall often use the facts that Radon measures ν on X also satisfy (see e.g. [S, I,§I.2])

$$\nu(B) = \inf\{\nu(U) : B \subseteq U,\ U \text{ open}\} \text{ for every Borel set } B \subseteq X$$

and that, under the conditions imposed upon X, a Borel measure is Radon if and only if it is locally finite, i.e. every point has a neighbourhood of finite measure (see e.g.[S, I,§II.3]).

If ν_k are Radon measures on X, we say that the sequence (ν_k) *converges vaguely* to the Radon measure ν if the following two conditions hold

1. $\nu(K) \ge \limsup_{k\to\infty} \nu_k(K)$ for every compact set $K \subseteq X$, and
2. $\nu(L) \le \liminf_{k\to\infty} \nu_k(L)$ for every open set $L \subseteq X$.

In the following we recall some useful descriptions of the vague convergence of Radon measures. For this purpose the following lemma, which is proved for finite measures in [T, Lemma 7.7], will turn out to be useful.

Lemma A1. *Let (ν_k) be a sequence of Radon measures on a Polish space X. Then, for every $\eta > 0$ and every open set $G \subseteq X$ and compact set $K \subseteq X$ with $K \subseteq G$, there are a compact set C with $K \subseteq C \subseteq G$ and an open set H with $K \subseteq H \subseteq G$ such that*

(a) $\eta + \limsup_{k\to\infty} \nu_k(C) \geq \limsup_{k\to\infty} \nu_k(H)$ *and*
(b) $\eta + \liminf_{k\to\infty} \nu_k(C) \geq \liminf_{k\to\infty} \nu_k(H)$.

Proof. Let G_j be the intersection of G with the open $1/j$-neighbourhoods of K. Pick strictly increasing sequences (a_j) converging to $\inf_i \limsup \nu_k(G_i)$ and (b_j) increasing to $\inf_i \liminf \nu_k(G_i)$. We can find a sequence $k_1 < k_2 < \dots$ of indices such that $\nu_k(G_j) > b_j$ for all $k \geq k_j$ and $\nu_l(G_j) > a_j$ for some $k_j \leq l < k_{j+1}$. For every j we may pick a compact set $K \subseteq C_j \subseteq G_j$ such that $\nu_k(C_j) \geq b_j$ for all $k_j \leq k < k_{j+1}$ and $\nu_l(C_j) \geq a_j$ for some $k_j \leq l < k_{j+1}$. Then the set $C = \bigcup C_j$ is compact (see e.g. [T, Lemma 7.5]) and satisfies $\liminf \nu_k(C) \geq \liminf b_j$ and $\limsup \nu_k(C) \geq \limsup a_j$. Picking j so large that $H = G_j$ satisfies $\limsup \nu_k(H) \leq \inf_i \limsup \nu_k(G_i) + \eta$ and $\liminf \nu_k(H) \leq \inf_i \liminf \nu_k(G_i) + \eta$ finishes the proof. □

Lemma A2. *Let (ν_k) be a sequence of Radon measures on a Polish space X converging vaguely to the Radon measure ν.*

(a) For every compact $K \subseteq X$ and every $\eta > 0$, there is an open set $H \supseteq K$ such that $\limsup_{k\to\infty} \nu_k(H) < \nu(K) + \eta$.
(b) For every open set $G \subseteq X$ with $\nu(G) < \infty$ and every $\eta > 0$, there is a compact set $C \subseteq G$ such that $\liminf_{k\to\infty} \nu_k(C) > \nu(G) - \eta$.

Proof. To prove (a) we pick an open set $G \supseteq K$ such that $\nu(G \setminus K) < \eta/2$. By Lemma A1(a) we can find a compact set $C \subseteq G$ and an open set $H \supseteq K$ with $\limsup \nu_k(H) \leq \eta/2 + \limsup \nu_k(C)$ and, as $\limsup \nu_k(C) \leq \nu(C) < \nu(K) + \eta/2$, this proves (a). To prove (b) we pick a compact set $K \subseteq G$ with $\nu(G \setminus K) < \eta/2$ and use Lemma A1(b) to find a compact set $C \subseteq G$ and an open set $H \supseteq K$ with $\eta/2 + \liminf \nu_k(C) \geq \liminf \nu_k(H)$. As $\liminf \nu_k(H) \geq \nu(H) > \nu(G) - \eta/2$, (b) is proved. □

Lemma A3. *A sequence (ν_k) of Radon measures on a Polish space X converges vaguely to a Radon measure ν if and only if for every compact set $K \subseteq X$ and every $\eta > 0$ there is an open set L with $K \subseteq L \subseteq X$ such that*

$$\limsup_{k\to\infty} \left| \int \varphi(x)\, d\nu_k(x) - \int \varphi(x)\, d\nu(x) \right| < \eta$$

for every continuous function $\varphi : X \to [0,1]$ such that $\varphi(x) = 0$ for $x \notin L$.

Proof. An easy approximation shows that the given condition implies that (ν_k) converges vaguely to ν. On the other hand, if the convergence is assumed, $\eta > 0$ and $K \subseteq X$ is compact, we can use Lemma A2(a) and (b) to first find an open set $L \supseteq K$ such that $\limsup \nu_k(L) < \nu(K) + \eta/2$ and then a compact set C with $K \subseteq C \subseteq L$ with $\limsup \nu_k(L \setminus C) < \eta$. Observe that $\sup \nu_k(C) < \infty$. If φ is given as in the lemma, we calculate, using Fubini's Theorem and Fatou's Lemma,

$$
\begin{aligned}
\int \varphi\, d\nu &\geq \int_C \varphi\, d\nu = \int_0^1 \nu(\{x \in C : \varphi(x) \geq t\})\, dt \\
&\geq \int_0^1 \limsup_{k\to\infty} \nu_k(\{x \in C : \varphi(x) \geq t\})\, dt \\
&\geq \limsup_{k\to\infty} \int_0^1 \nu_k(\{x \in C : \varphi(x) \geq t\})\, dt \\
&= \limsup_{k\to\infty} \int_C \varphi\, d\nu_k \geq \limsup_{k\to\infty} \int \varphi\, d\nu_k - \eta\,,
\end{aligned}
$$

and, similarly, we get

$$
\begin{aligned}
\int \varphi\, d\nu &= \int_0^1 \nu(\{x \in L : \varphi(x) > t\})\, dt \\
&\leq \int_0^1 \liminf_{k\to\infty} \nu_k(\{x \in L : \varphi(x) > t\})\, dt \\
&\leq \liminf_{k\to\infty} \int_0^1 \nu_k(\{x \in L : \varphi(x) > t\})\, dt = \liminf_{k\to\infty} \int \varphi\, d\nu_k\,,
\end{aligned}
$$

which together implies the condition given in the lemma. □

As an important consequence of the lemma we note that the vague limit of a sequence (ν_k), if it exists, is uniquely determined. Although we have not defined a topology inducing the vague convergence (and, in fact, this may not be possible, see [T, Chap.7]), the notion of relative sequential compactness with respect to the vague convergence is naturally defined for a set of Radon measures.

Lemma A4. *A set M of Radon measures on a Polish space X is relatively sequentially compact with respect to the vague convergence if and only if* $\sup\{\nu(K) : \nu \in M\} < \infty$ *for every compact set $K \subseteq X$.*

Proof. If there is a sequence (ν_k) in M such that, for some compact K, the sequence $\nu_k(K)$ tends to ∞, this sequence clearly has no convergent subsequence and M cannot be relatively sequentially compact. On the other hand suppose (ν_k) is a sequence in M and $(\nu_k(K))$ is bounded for every compact K. We have to find a convergent subsequence. Using Lemma A1(a) we can find a sequence (G_j) of open sets increasing to X with $\sup_k \nu_k(G_j) < \infty$. As any family of finite measures with bounded mass has a vaguely convergent subsequence (see e.g.[T, Theorem 7.8]), we can pick a subsequence (μ_k) of (ν_k) such that, for every j, the sequence $1_{G_j}\mu_k$ converges, say to $\mu^{(j)}$. For every compact set K there is an index n such that $K \subseteq G_n$ and therefore $\mu^{(i)}|_K = \mu^{(j)}|_K$ for all $i, j \geq n$. Using this, it is easy to see that we can define a Radon measure μ by $\mu(B) = \sup \mu^{(n)}(K)$, where the supremum extends over all compact sets $K \subseteq B$ and n is chosen according to K. It is then straightforward to check that μ is indeed the limit of the sequence (μ_j). □

Lemma A5. *Suppose that μ_k and ν_k are Radon measures on a Polish space X such that (μ_k) converges vaguely to μ_0 and (ν_k) converges vaguely to ν_0. Suppose further that for every $\eta > 0$ and for every pair $K \subseteq L$ of subsets of X such that K is compact and L is open there is an open set U such that $K \subseteq U \subseteq L$ and*

$$\liminf_{k\to\infty} \mu_k(U) \le \eta + \limsup_{k\to\infty} \nu_k(L) .$$

Then $\mu_0 \le \nu_0$.

Proof. If $K \subseteq X$ is compact and $\eta > 0$, we use Lemma A3 to find an open set $L \supseteq K$ such that

$$\limsup_{k\to\infty} \left| \int \varphi(x)\, d\mu_k(x) - \int \varphi(x)\, d\mu_0(x) \right| \quad < \quad \eta$$

and

$$\limsup_{k\to\infty} \left| \int \varphi(x)\, d\nu_k(x) - \int \varphi(x)\, d\nu_0(x) \right| \quad < \quad \eta$$

for every continuous function $\varphi : X \to [0,1]$ such that $\varphi(x) = 0$ for $x \notin L$. Let $L \supseteq L_0 \supseteq K$ be an open set such that $\nu_0(L_0 \setminus K) < \eta$ and let $L_0 \supseteq L_1 \supseteq K$ be an open set whose closure is contained in L_0. Using the assumption of the lemma, we find an open set $L_1 \supseteq U \supseteq K$ such that

$$\liminf_{k\to\infty} \mu_k(U) \le \eta + \limsup_{k\to\infty} \nu_k(L_1) .$$

Let $\varphi : X \to [0,1]$ be a continuous function such that $\varphi(x) = 1$ for $x \in K$ and $\varphi(x) = 0$ for $x \notin U$ and let $\psi : X \to [0,1]$ be a continuous function such that $\psi(x) = 1$ for $x \in L_1$ and $\psi(x) = 0$ for $x \notin L_0$. Then

$$\begin{aligned} \mu_0(K) \quad &\le \quad \int \varphi(x)\, d\mu_0(x) \le \eta + \liminf_{k\to\infty} \int \varphi(x)\, d\mu_k(x) \le \eta + \liminf_{k\to\infty} \mu_k(U) \\ &\le \quad 2\eta + \limsup_{k\to\infty} \nu_k(L_1) \le 2\eta + \limsup_{k\to\infty} \int \psi(x)\, d\nu_k(x) \\ &\le \quad 3\eta + \int \psi(x)\, d\nu_0(x) \le 3\eta + \nu_0(L_0) \le 4\eta + \nu_0(K). \end{aligned}$$

Since $\eta > 0$ and the compact set K are arbitrary, this shows that $\mu_0 \le \nu_0$. □

A sequence (ν_k) of Radon measures with $\limsup_{k\to\infty} \nu_k(X) < \infty$ is said to *converge weakly* to the Radon measure ν if $\int \varphi\, d\nu_k \to \int \varphi\, d\nu$ for every bounded continuous function φ on X.

Lemma A6. *Let* ν_k *be Radon measures on a Polish space* X *and suppose that* $\limsup_{k\to\infty} \nu_k(X) < \infty$ *and* (ν_k) *converges vaguely to a Radon measure* ν. *Then the following statements are equivalent:*

(a) (ν_k) *converges weakly to* ν.

(b) $\nu(X) \ge \limsup_{k\to\infty} \nu_k(X)$.

(c) For every $\eta > 0$ *there is* $K \subseteq X$, *compact, such that* $\limsup \nu_k(X \setminus K) \le \eta$.

Proof. As the function 1_X is continuous and bounded, (a) implies (b). If (b) holds, then $\nu(A) \ge \limsup_{k\to\infty} \nu_k(A)$ for every closed set $A \subseteq X$ and we can conclude the weak convergence by means of Fubini's Theorem as exercised in the proof of Lemma A3. Finally, the equivalence of (a) and (c) follows from Prohorov's Theorem (see [S, App. §3,Th. 3 and 4]). □

Of course, if the Polish space X is compact, then Radon measures are finite and weak and vague convergence coincide. More generally, if X is a locally compact Polish space, then ν_k converges vaguely to ν if and only if $\int \varphi\, d\nu_k \to \int \varphi\, d\nu$ for every continuous function φ on X with compact support. Then the vague convergence of Radon measures is induced by the vague topology, the smallest topology on the set of Radon measures that makes the mappings $\mu \mapsto \int \varphi\, d\mu$ continuous for all continuous φ with compact support. In the case $X = \mathbf{R}^d$ the set $\mathcal{M}(\mathbf{R}^d)$ of all Radon measures on $\mathbf{R}^d$ equipped with the vague topology is a Polish space, see [Mat, 14.15], and we observe, as a consequence of Lemma A4, that a subset M of $\mathcal{M}(\mathbf{R}^d)$ is relatively compact if and only if for each $n = 1, 2, \ldots$ there is $c(n) \in (0, \infty)$ such that $\mu U(0, n) \le c(n)$ for each $\mu \in M$, where $U(x, r)$ denotes the open Euclidean ball of radius r centred in x.

References

[APa] M. Arbeiter and N. Patzschke. Random self similar multifractals. *Math. Nachr.*, 181:5–42, 1996.

[Ba] C. Bandt. The tangent distribution for self similar measures. *Lecture at the 5th Conference on Real Analysis and Measure Theory, Capri*, 1992.

[BeFi] T. Bedford and A.M. Fisher. Analogues of the Lebesgue density theorem for fractal sets of reals and integers. *Proc. London Math. Soc.(3)*, 64:95–124, 1992.

[Bo] A. Bonnet. On the regularity of edges in image segmentation. *Ann. Inst. Poincaré, Analyse non lineaire 13*, 4:485–528, 1996.

[FaSp] K.J. Falconer and O.B. Springer. Order–two density of sets and measures with non–integral dimension. *Mathematika*, 42:1–14, 1995.

[Fe] H. Federer. *Geometric Measure Theory.* Springer Verlag, Berlin, 1968.

[Fi] A.M. Fisher. Convex-invariant means and a pathwise central limit theorem. *Adv. in Math.*, 63:213–246, 1987.

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

[H] P. Huovinen. Singular integrals and rectifiability of measures in the plane. *Ann. Acad. Sci. Fennicae, Diss.* 109:1–63, 1997.

[K] O. Kallenberg. *Random Measures.* Akademie-Verlag, Berlin, 1983.

[KM] D. Krieg and P. Mörters. The tangent measure distributions of hyperbolic Cantor sets. *To appear in Mh. Math.*, 1998.

[Ma1] J.M. Marstrand. The (ϕ, s)–regular subsets of n–space. *Trans. Am. Math. Soc.*, 113:369–392, 1964.

[Ma2] J.M. Marstrand. Order–two density and the strong law of large numbers. *Mathematika*, 43:1–22, 1996.

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

[Me] J. Mecke. Stationäre zufällige Maße auf lokalkompakten abelschen Gruppen. *Zeitschr. Wahrsch. verw. Gebiete*, 9:36–58, 1967.

[Mö1] P. Mörters. *Tangent Measure Distributions and the Geometry of Measures.* PhD thesis, University College London, 1995.

[Mö2] P. Mörters. Symmetry properties of average densities and tangent measure distributions of measures on the line. *To appear in Adv. Appl. Math.*, 1998.

[Mö3] P. Mörters. The tangent measure distributions of the paths of Brownian motion. *In preparation*, 1998.

[MöPr] P. Mörters and D. Preiss. On one–sided average densities of fractal measures on the line. *Sem. d'initiation à l'analyse, Publ. Univ. Pierre et Marie Curie, Paris*, 117: Exposé 4, 1994/5.

[O1] T. C. O'Neil. *A Local Version of the Projection Theorem and Other Results in Geometric Measure Theory.* PhD Thesis, University College London, 1995.

[O2] T. C. O'Neil. A local version of the projection theorem. *Proc. London Math. Soc.(3)*, 73:68–104, 1996.

[Pr] D. Preiss. Geometry of measures in $\mathbf{R}^n$: Distribution, rectifiability and densities. *Ann. Math.*, 125:537–643, 1987.

[PaZU] N. Patzschke and U. Zähle. Self similar random measures IV. The recursive construction model of Falconer, Graf, Mauldin and Williams. *Math. Nachr.*, 149:285–302, 1990.

[PaZM] N. Patzschke and M. Zähle. Self similar random measures are locally scale invariant. *Prob. Theory Rel. Fields*, 97:559–574, 1993.

[S] L. Schwartz. *Radon measures on arbitrary topological spaces and cylindrical measures.* Oxford University Press, 1973.

[St] E.M. Stein. *Singular integrals and differentiability properties of functions.* Princeton University Press, 1970.

[T] F. Topsøe. *Topology and Measure.* Springer LNM 133, 1970.

[V] M. Valadier. Young measures. *In: Methods of Nonconvex Analysis, Springer LNM 1446,* pp.152–188, 1989.

[ZU] U. Zähle. Self similar random measures I. Notion, carrying Hausdorff dimension and hyperbolic distribution. *Prob. Theory Rel. Fields*, 80:79–100, 1988.

This article was processed by the author using the LaTeX style file *cljour1* from Springer-Verlag.