JEMS

Michael Hochman

A ratio ergodic theorem for multiparameter non-singular actions

Abstract. We prove a ratio ergodic theorem for non-singular free $\mathbb{Z}^d$ and $\mathbb{R}^d$ actions, along balls in an arbitrary norm. Using a Chacon–Ornstein type lemma the proof is reduced to a statement about the amount of mass of a probability measure that can concentrate on (thickened) boundaries of balls in $\mathbb{R}^d$. The proof relies on geometric properties of norms, including the Besicovitch covering lemma and the fact that boundaries of balls have lower dimension than the ambient space. We also show that for general group actions, the Besicovitch covering property not only implies the maximal inequality, but is equivalent to it, implying that further generalization may require new methods.

Keywords. Group actions, measure preserving transformations, commuting transformations, non-singular actions, ergodic theorem, maximal inequality

1. Introduction

Consider a non-singular action of a group G on a standard σ-finite measure space $(\Omega, \mathcal{B}, \mu)$, which we denote $\omega \mapsto T^g \omega$; we shall assume that the action is free and ergodic. From the action on Ω there is induced an isometric linear action on L^∞, also denoted T^g, given by $T^g f = f \circ T^{g^{-1}}$; and this in turn induces an isometric linear action on the Banach dual of L^∞, whose restriction to L^1 is given by $\widehat{T}^g f = (f \circ T^{g^{-1}}) \cdot \frac{d(\mu \circ T^g)}{d\mu}$ (in the measure preserving case the Radon–Nikodym derivative is identically 1 and $\widehat{T}$ reduces to the usual Koopman operator).

For $\mathbb{Z}$-actions, there is in this setting an analogue to Birkhoff's ergodic theorem which is due to Hopf [Hop], later generalized by Hurewicz [Hur44] to a "measureless" statement, and by Chacon–Ornstein to the operator setting [CO60]. Hopf's ratio ergodic theorem states that, for an ergodic $\mathbb{Z}$-action generated by a transformation $T : \Omega \to \Omega$, for any $f, g \in L^1$ with $\int g \, d\mu \neq 0$, the following ratios converge almost surely:

$$R_n(f, g) = \frac{\sum_{k=0}^n \widehat{T}^k f}{\sum_{k=0}^n \widehat{T}^k g}.$$

M. Hochman: Department of Mathematics, Princeton University, Fine Hall, Washington Rd., Princeton, NJ 08544, USA; e-mail: hochman@princeton.math.edu

If in addition T is conservative (i.e. has no non-trivial wandering sets), or if the one-sided sum is replaced with the symmetric sum from $-n$ to n, then the limit is the constant function $\int f / \int g$. Note that for probability-preserving actions this is equivalent to the usual ergodic theorem: R_n becomes an ergodic average by setting $g \equiv 1$. For general actions this equivalence is false: for example, for measure-preserving actions of an infinite measure the ergodic averages converge to 0, and not to the mean.

While the ergodic theorem for measure-preserving actions on probability spaces has been broadly generalized to the group setting [OW87], the ratio theorem has not seen similar extensions, even to $\mathbb{Z}^d$-actions. For a time it was thought no such extension was possible. The natural thing to try in $\mathbb{Z}^d$ is to sum over the cubes $Q_n = [0; n]^d$, but there is a counter-example, due to Brunel and Krengel, showing that these ratios may diverge for $d > 1$ [Kre85]. However, recently J. Feldman [Fel07] proved a partial result for $\mathbb{Z}^d$, showing that if the generators of the action act conservatively then the ratio theorem holds for sums over the symmetric cubes $[-n; n]^d$. However, the conservativity requirement is essential to the argument, and is more restrictive than one would like, since there are certainly actions that are conservative but whose generators are not (consider for example the $\mathbb{Z}^2$ action generated by translation by $\sqrt{2}$ and $\sqrt{3}$ on $\mathbb{R}$; the action is conservative but no cyclic subgroup acts conservatively).

Our main result is an unconditional ratio theorem for multiparameter actions:

Theorem 1.1. *Let $\{T^u\}_{u \in \mathbb{Z}^d}$ be a free, non-singular ergodic action on a standard σ-finite measure space. Let $\|\cdot\|$ be a norm on $\mathbb{R}^d$ and let $B_n = \{u \in \mathbb{Z}^d : \|u\| \leq n\}$. Then for every $f, g \in L^1$ with $\int g \neq 0$, we have*

$$R_n(f, g) = \frac{\sum_{u \in B_n} \widehat{T}^u f}{\sum_{u \in B_n} \widehat{T}^u g} \xrightarrow[n \to \infty]{} \frac{\int f}{\int g}$$

almost everywhere.

A similar result holds for $\mathbb{R}^d$-actions.

The method of proof follows a two-step argument that is by now standard and goes back to Hopf. With g fixed, one first proves that $R_n(f, g)$ converges for f in some dense subset $\mathcal{F} \subseteq L^1$. Then one applies a maximal inequality to go from $\mathcal{F}$ to its closure (we shall discuss maximal inequalities in more detail below). In Feldman's proof the conservativity assumption is used to construct a special family of functions which is dense and for which the ratios converge. We shall instead work with the larger subspace generated by g and bounded co-boundaries:

$$\mathcal{F} = \text{span}\{g, h - \widehat{T}^u h : u \in \mathbb{Z}^d, h \in L^1 \cap L^\infty\}.$$

A standard argument shows that $\mathcal{F}$ is dense in L^1 (see e.g. [Fel07, Aar97]). Since $R_n(g, g) \equiv 1$, convergence of $R_n(f, g)$ for all $f \in \mathcal{F}$ will follow once it is established for co-boundaries $f = h - \widehat{T}^u h$. For such f some cancellation occurs in the sum $\sum_{v \in B_n} \widehat{T}^v f$, and some algebra (given in Section 5, or see [Bec83]) reduces the problem to the following variant of the Chacon–Ornstein lemma, into the proof of which goes most of the hard work:

Theorem 1.2. *Under the hypotheses of Theorem* 1.1, *for any* $h \in L^\infty \cap L^1$, *and for any* $t > 0$,

$$\frac{\sum_{u \in B_{n+t} \setminus B_{n-t}} \widehat{T}^u h}{\sum_{u \in B_n} \widehat{T}^u h} \to 0$$

almost surely.

Note that for $d = 1$ the numerator contains only two terms and it suffices to show that the denominator tends to ∞; this follows easily for conservative actions, while the non-conservative case can be proved directly. On the other hand, for $d > 1$ the number of terms in the numerator is on the order of n^{d-1}, and when the measure is infinite the denominator satisfies $n^{-d} \sum_{u \in B_n} \widehat{T}^u h \to 0$. Thus a more sophisticated argument is necessary.

Our proof of Theorem 1.2 applies the transference principle to reduce theorem 1.2 to a geometric statement about the amount of mass which can concentrate on boundaries of balls for finite measures in $\mathbb{R}^d$. In the proof we use two facts related to the finite dimension of $\mathbb{R}^d$ (with combinatorial analogs in $\mathbb{Z}^d$). One is the Besicovitch covering lemma, about which we shall have more to say below. The other is (a variant of) the fact that the boundaries of balls in $\mathbb{R}^d$ are manifolds of lower dimension, which is closely related to finite topological dimension of $\mathbb{R}^d$. This property has apparently not been exploited before in this context.

It is also worth noting that our method does not require us to distinguish between the conservative and non-conservative case.

As we have mentioned already, in order to derive the ratio theorem from Theorem 1.2 one uses the maximal inequality, which is the second subject of this paper. We shall denote by B_n an increasing sequence of finite subsets of G which satisfy $e \in B_0$ (here e is the identity element in G).

Definition 1.3. *An ergodic action of G admits a* ratio maximal inequality (*with respect to* (B_n)) *if, for every* $0 \le g \in L^1$ *there is a constant M such that, for every* $0 \le f \in L^1$ *and* $\varepsilon > 0$,

$$\mu_g\{\omega \in \Omega : \sup_n R_n(f, g) > \varepsilon\} \le \frac{M}{\varepsilon} \int f \, d\mu$$

where $d\mu_g = g \cdot d\mu$. *We say that* there is a maximal inequality *for G* (*with respect to* (B_n)) *if every action admits a maximal inequality.*

Notice that when $\mu(\Omega) = 1$, we can take $g \equiv 1$. Then $\mu_g = \mu$ and $R_n(f, g)$ are the ergodic averages of f, so the ratio maximal inequality reduces to the usual maximal inequality. Note also that we allow the constant M to depend on g, since this is what is used in the proof of the ergodic theorem.

The ordinary maximal inequality for probability-preserving actions of amenable groups is known to hold quite generally [Lin01, Wei03], but this is not so in the nonsingular case. Indeed, if the Krengel–Brunel counter-example is examined closely it is evident that there is a dense class of functions f for which the ratio theorem holds. The problem must be that the maximal inequality fails. This is closely related to the fact that the sum is over one-sided cubes $Q_n = [1; n]^d$, which fail to satisfy the Besicovitch covering property:

Definition 1.4. *A sequence (B_n) of subsets of G satisfies the* Besicovitch covering prop-
erty *with constant C if the following holds. If $E \subseteq G$ is finite, and for each $g \in E$ we are
given a translate $B_{n(g)}g$ of one of the B_n's, then there is a subfamily of these translates
which covers E in such a way that no point in G is covered more than C times.*

This geometric property has found many applications in analysis; an excellent source on
this is [dG75]. That it implies the ratio maximal inequality was first shown by M. Becker
[Bec83] for balls $B_n \subseteq \mathbb{R}^d$ in a given norm. A maximal inequality relying on the Besicov-
itch property was later also established by E. Lindenstrauss and D. Rudolph for a more
general class of non-singular group actions [Lin06]. A short proof of the general case can
be found in Feldman's paper [Fel07]. Other applications of the Besicovitch property to
ergodic theory appear in [Hoc06].

   It is thus known that the Besicovitch property implies the ratio maximal inequality. It
has apparently not been observed before that it is also necessary.

Theorem 1.5. *Let G be a countable group and $B_r \subseteq G$ an increasing sequence of sym-
metric sets with $\bigcap B_r = \{e\}$. Then there is a ratio maximal inequality for G if and only
if B_n satisfies the Besicovitch property.*

Actually, more is true: if B_n is not Besicovitch then the ratio maximal inequality fails
for every free action of the group. Contrast this with the usual maximal inequality, which
holds for any measure-preserving action of an amenable group on a probability space, as
long as the averages are taken over a tempered Følner sequence [Lin01, Wei03].

   One should note that the Besicovitch property is rather rare. It fails, for example, for
the Heisenberg group when B_n are balls with respect to several natural metrics [Rig04].

   It is not clear what all this says about the ratio ergodic theorem. For probability-
preserving actions of amenable groups the ratio theorem along tempered Følner sequences
follows from the ordinary ergodic theorem. At the same time, the ratio maximal inequality
fails, as we saw before. The ratio ergodic theorem does hold, for trivial reasons, for dissi-
pative actions (e.g. on atomic measure spaces). This leaves the hope that a ratio theorem
may persist in a more general setting, even without a maximal inequality.

   The rest of this paper is organized as follows. In the next section we discuss the
Besicovitch property and prove Theorem 1.5. In Section 3 we discuss some covering and
disjointification lemmas. In Section 4 we define coarse dimension and prove our main
tool about concentration of measures on ball boundaries. In Section 5 we complete the
proofs of Theorem 1.2 and the ratio theorem, 1.1.

2. The Besicovitch lemma and the maximal inequality

In this section we prove Theorem 1.5. We shall reformulate the Besicovitch covering
property for metric spaces and present it in several equivalent forms. An excellent source
on these matters is [dG75].

   Given a metric space (X, d) we denote by $B_r(x)$ the open ball of radius r centered
at x. We think of balls as carrying with them the information about their center and radius,
which are not in general determined by the ball as a set.

A finite family of balls $\mathcal{U} = \{B_{r(i)}(x_i) : 1 \leq i \leq N\}$ is called a *carpet* over $\{x_1, \ldots, x_N\}$. It is sometimes convenient to regard carpets as ordered sets. Note that the statement that $\mathcal{U}$ is a carpet over E is stronger than the statement that it covers E, since the former asserts that each $x \in E$ is the center of some ball in $\mathcal{U}$, whereas the latter only says that x belongs to some ball.

We say that a collection of sets has *multiplicity* $\leq m$ if every point is contained in at most m elements of the collection.

A metric space satisfies the *Besicovitch property* with constant C [Bes45, dG75] if for any carpet over E there exists a subcarpet which covers E and has multiplicity $\leq C$. The main example for this is $\mathbb{R}^d$ with a norm-induced metric; this was shown by Morse [Mor47]. A more accessible proof can be found in [dG75] or can be deduced from Proposition 4.2 below.

This definition of the Besicovitch property is consistent with the one in the introduction if $X = G$ is a group, d is a right-invariant metric on G, and B_n are the balls of radius n around the group's identity element. We shall allow ourselves to switch freely between these two formalisms, which are notationally identical.

We say that a sequence $B_{r(i)}(x_i)$ is *incremental* if $r(i)$ is non-increasing and $x_i \notin \bigcup_{j<i} B_{r(j)}(x_i)$.

The Besicovitch property has several equivalent forms which are useful in applications.

Proposition 2.1. *Let X be a metric space and $C \in \mathbb{N}$. The following are equivalent:*

(1) *X has the Besicovitch property with constant C.*
(2) *For any carpet $\mathcal{U}$ over E and $A, B \subseteq E$, if $t > 0$ and $|A \cap F|/|B \cap F| < t$ for every $F \in \mathcal{U}$, then $|A|/|B| < Ct$.*
(3) *For any carpet $\mathcal{U}$ over E and $A, B \subseteq E$, if $t > 0$ and $|A \cap F|/|B \cap F| > t$ for every $F \in \mathcal{U}$, then $|A|/|B| > (1/C)t$.*
(4) *Every incremental sequence has multiplicity $\leq C$.*
(5) *For any carpet $\mathcal{U}$ over E there is an incremental sequence of sets from $\mathcal{U}$ covering E and with multiplicity $\leq C$*

Proof. (1) implies (2): Using (1) we may pass to a subcollection $\{F_i\}_{i \in I} \subseteq \mathcal{U}$ with multiplicity $\leq C$, and which covers E, and hence covers A and B. It now follows that

$$|A| \leq \sum_i |A \cap F_i| < t \sum |B \cap F_i| \leq Ct \left| \bigcup_i (B \cap F_i) \right| = Ct|B|.$$

(2) and (3) are equivalent on reversing the roles of A and B.

(2) implies (4): Let $B_{r(1)}(x_1), \ldots, B_{r(N)}(x_N)$ be an incremental sequence. For each i, note that if $j < i$ then $x_i \notin B_{r(j)}(x_j)$ because the sequence is incremental, and also, since $r(j)$ do not increase, if $j > i$ then $x_i \notin B_{r(j)}(x_j)$. Thus each x_i is covered by exactly one of the sets $B_{r(j)}(x_j)$. Next, let $y \neq x_i$, $i = 1, \ldots, N$. Choose[1] $\varepsilon \geq 0$ so that $x_i \notin B_\varepsilon(y)$; setting $A = \{x_i : y \in B_{r(i)}(x_i)\}$ and $B = \{y\}$ we see that for any $t > 1$ the

---

[1] In the case of sets in a group rather than balls in a metric space, the existence of such an ε follows from the assumption that $B_0 = \{e\}$.

hypothesis of (2) is satisfied with respect to the carpet $\{B_\varepsilon(y)\} \cup \{B_{r(i)}(x_i) : x_i \in A\}$ over $\{x_1, \ldots, x_N, y\}$, implying that $N = |A|/|B| < Ct$ for every $t > 1$, and the claim follows.

(4) implies (5): Let $\mathcal{U} = \{B_{r(1)}(x_1), \ldots, B_{r(N)}(x_N)\}$ be a carpet over the set $E = \{x_1, \ldots, x_N\}$. Without loss of generality we may assume that $r(1) \geq \cdots \geq r(N)$. Iterate over i from 1 to N and at each stage discard the set $B_{r(i)}(x_i)$ if x_i belongs to the union of the sets selected previously and otherwise keep it. We obtain an incremental sequence which covers E, and by (4) has multiplicity $\leq C$.

The implication (5)$\Rightarrow$(1) is trivial. $\square$

We can now prove Theorem 1.5:

Theorem. *Let G be a countable group and $B_r \subseteq G$ an increasing sequence of symmetric sets with $\bigcap B_r = \{e\}$. Then G satisfies a ratio maximal inequality for sums over B_r if and only if X satisfies the Besicovitch property with respect to B_r.*

Proof. One direction is the ratio maximal inequality of Becker and of Lindenstrauss and Rudolph [Bec83, Lin06].

Conversely, suppose the Besicovitch property fails. Given an action of G and functions $f, h \in L^1$ let us write

$$C(f, h) = \frac{\mu_h\{\sup_n R_n(f, h) > \varepsilon\}}{\int f \, d\mu}$$

where $d\mu_h = h \cdot d\mu$. We are out to show that for some $h \in L^1$ this quantity is not bounded above as f varies.

We start with the action of G on itself by left translation, $T^g x = gx$, and let μ be Haar (counting) measure, which is clearly preserved. By Proposition 2.1, for every $M > 0$ there is a $t > 0$ and finite sets $U, V \subseteq G$, and $n(g) \in \mathbb{N}$ for $g \in U \cup V$, such that

$$|U \cap B_{n(g)}g|/|V \cap B_{n(g)}g| > t$$

for $g \in U \cup V$, but $|U|/|V| < t/M$. Let $f = 1_U$ and $h = 1_V$. Then

$$\left\{ x \in V : \sup_n \frac{\#\{g \in B_n : T^g x \in U\}}{\#\{g \in B_n : T^g x \in V\}} > t \right\} = V$$

so

$$\mu_h\left(x \in G : \sup_n \frac{\sum_{g \in B_n} f(T^g x)}{\sum_{g \in B_n} h(T^g x)} > t \right) = \mu_h(V) = |V| > \frac{M}{t}|U| = \frac{M}{t} \int f \, d\mu.$$

We have found that for each $M > 0$ there are $f, h \in L^1(G)$ with

$$C(f, h) > M.$$

This is already enough to conclude that the ratio maximal inequality cannot hold with a constant which is independent of h.

We next want to show that for every action of G on a measure space $(\Omega, \mathcal{F}, \mu)$, there is a fixed h with $\sup_{f \in L^1} C(f, h) = \infty$. We prove this for the case that the measure space

is non-atomic and the action measure preserving. The proof for the atomic case is simpler, so we omit it.

We construct by induction functions which will establish our claim. Suppose we have $0 \le f_k, h \in L^1(\Omega)$ for $1 \le k \le n$ such that $C(f_k, h) > k$. Using what we know about the action of G on itself, we can find $0 \le f', h' \in L^1(G)$ so that $C(f', h') > n+1$. Below we show how to merge h, h' into a function $h'' \in L^1(\Omega)$, and to construct a function $f_{n+1} \in L^1(\Omega)$ derived from f' so that $C(f_k, h'') > k$ for $1 \le k \le n+1$ and $\|h - h''\|_1 < \varepsilon$, where ε is a parameter which can be chosen arbitrarily small. Once this is done, we can iterate the process and pass to a limit function $h_* \in L^1(\Omega)$ which, for the sequence f_k constructed, satisfies $C(f_k, h_*) \to \infty$, completing the proof of the theorem.

Fix $\varepsilon > 0$. We may assume that the functions f', h' that we have found on G are supported inside B_{N_0} for some N_0 and that if $R_i(f', h')(g) > n+1$ for some i and $g \in G$ then $g \in B_{N_0}$. If $\omega \in \Omega$ let $i_k(\omega)$ be the first index such that $R_{i_k(\omega)}(f_k, h)(\omega) > k$. Since $i_k(\cdot)$ is measurable for $k = 1, \ldots, n$, there is some N_1 such that $\mu_h(i_k > N_1) < \varepsilon$ for each k. Set $N = \max\{N_0, N_1\}$.

Using the fact that the action is non-atomic and free, we can find a set $A \subseteq \Omega$ with positive μ-measure, so that $gA \cap g'A = \emptyset$ whenever $g, g' \in B_N^{-1} B_N$, and so that

$$\widetilde{A} = \cup_{g \in B_N^{-1} B_N} gA$$

has measure less than ε, both with respect to μ and with respect to μ_h [Wei03]. By the choice of N, if h'' is a function that differs from h only on $\bigcup_{g \in B_N} gA$ then $C(f, h'') > C(f_k, h'') - 2\varepsilon$, because for $\omega \in \Omega \setminus \widetilde{A}$, we have $\widehat{T}^g h(\omega) = \widehat{T}^g h''(\omega)$ as long as $g \in B_N$, implying $R_{i_k(\omega)}(f, h')(\omega) = R_{i_k(\omega)}(f, h)(\omega)$ outside of $\widetilde{A}$ and outside the set where $i_k > N$.

Define $h''(\omega) = h'(g)$ for $\omega \in gA$ and $g \in B_{N_0}$, and $h'' = h$ otherwise. By the above, $C(f_k, h'') > C(f_k, h) - \varepsilon$ for $k = 1, \ldots, n$ and $\|h - h''\|_1 = \int_{\widetilde{A}} |h - h''| \, d\mu$, which can be made $> k$ and $< 2^{-n}$ respectively by choosing ε small enough.

Finally, define the function $f_{n+1}(\omega) = f'(g)$ for $\omega \in gA$ and $g \in B_{N_0}$, and 0 otherwise. Since on $B_N A$ we have $R_i(f_{n+1}, h'') = R_i(f', h')$ for $i \le N_0$, it follows that $C(f_{n+1}, h'') > n+1$ (notice that $C(\cdot, \cdot)$ is invariant under scaling of μ, which explains why constructing f'', h'' on a part of the measure space which is small with respect to μ does not ruin the property $C(f', h') > n+1$). $\qquad\square$

3. The doubling property and disjointification

Another property of metric spaces which is related (but not equivalent) to the Besicovitch property is the doubling condition. Let (X, d) be a metric space, and suppose we are given a measure on X which we denote by $|\cdot|$; in our setting it will be Haar measure, and for $\mathbb{Z}^d$ will denote the usual counting measure. We say that (X, d) satisfies the *doubling condition* with constant D if for every ball $B_r(x)$ we have $|B_{2r}(x)| \le D|B_r(x)|$. This is satisfied for the groups $\mathbb{Z}^d, \mathbb{R}^d$ for any norm; for finitely generated groups with word metric this condition is equivalent to polynomial growth.

In this section we derive some covering lemmas based on the doubling and Besicovitch properties. For this we require some more notation. Write rad B for the radius of a ball B, and if $\mathcal{U}$ is a collection of balls we write $\operatorname{rmax}\mathcal{U}$ and $\operatorname{rmin}\mathcal{U}$ for the maximal and minimal radii of balls in $\mathcal{U}$, respectively.

We say that $\mathcal{U}$ is *well-separated* if any two balls in $\mathcal{U}$ are at distance at least $\operatorname{rmin}\mathcal{U}$ from each other (the distance between sets A and B is $\inf\{d(a,b) : a \in A, \ b \in B\}$).

The doubling condition together with the Besicovitch property implies the following standard covering result which can be found e.g. in [dG75].

Lemma 3.1. *Let X be a metric space with a measure, and suppose it satisfies the Besicovitch property with constant C and the doubling condition with constant D. Then for every finite $E \subseteq X$ and every carpet $\mathcal{U}$ over E there is a subcollection $\mathcal{V} \subseteq \mathcal{U}$ which covers E and which can be partitioned into $\chi = CD^2 + 1$ subcollections, each of which is well-separated.*

Proof. We begin with a few observations. Let $x \in X$ an let $\mathcal{W}$ be a collection of n balls of radius R centered inside $B_{3R}(x)$, and suppose $\mathcal{W}$ has multiplicity $\leq C$. Then $\bigcup \mathcal{W} \subseteq B_{4R}(x)$, so

$$n|B_R(x)| \leq C \cdot |B_{4R}(x)| \leq D^2 \cdot C \cdot |B_R(x)|,$$

hence $n \leq CD^2 = \chi - 1$.

Next, suppose $\mathcal{W}$ consists of balls of radius $\geq R$ which intersect $B_{2R}(x)$, and suppose $\mathcal{W}$ has multiplicity $\leq C$. By replacing each ball $B \in \mathcal{W}$ with a ball of radius R contained in B and centered within $B_{3R}(x)$, we conclude again that $|\mathcal{W}| \leq \chi - 1$.

We now prove the lemma. By (5) of Proposition 2.1, choose an incremental sequence $U_1, \ldots, U_n \in \mathcal{U}$ covering E, and assign colors $1, 2, \ldots, \chi$ to the U_i as follows. Color U_1 arbitrarily. Assuming we have colored $U_1, \ldots, U_k$ consider U_{k+1}. By the above, U_{k+1} cannot be within distance rad U_k of more than $\chi - 1$ of the balls we have already colored, so there is a color which we can assign to it without violating the coloring condition. When all the balls are colored, set $\mathcal{V}_k = $ the balls colored k. Clearly each collection is well-separated. $\qquad\square$

For a metric space X, we denote by $\chi(X)$ the smallest constant χ for which X satisfies the conclusion of the proposition. Clearly, if $Y \subseteq X$ then $\chi(Y) \leq \chi(X)$. If X satisfies the hypotheses of the proposition then $\chi(X) \leq CD^2 + 1$, so this bound holds for any $Y \subseteq X$, even though Y may no longer satisfy the doubling condition.

Corollary 3.2. *In the notation of the previous lemma, assume there is given a finite measure μ supported in a set E. Then there is a well-separated subset of $\mathcal{U}$ which covers a set of mass $\geq (1/\chi)\mu(E)$.*

Proof. Color the balls as in the previous lemma. Each monochromatic collection of balls is well-separated and since there are χ colors, and the union covers E, one color class covers a $1/\chi$-fraction of the mass. $\qquad\square$

Our next objective is a lemma like the above except that, instead of capturing mass in a well-separated collection of balls, we do so with spheres, or more precisely thick spheres.

This is too much to hope for in general, but we can do it under the hypothesis that the balls we begin with contain some fraction of their mass on their boundaries.

For a metric space X, the *t-boundary* of a ball $B_r(x)$ is defined for $t \le r$ by

$$\partial_t B_r(x) = B_{r+t}(x) \setminus B_{r-t}(x);$$

this is called a *thick sphere*; we say that its radius is r and thickness t, and agree that the sphere carries this information with it (r, t are not determined from $\partial_t B_r(x)$ in general). Also write $\partial B_r(x)$ for the usual topological boundary of $B_r(x)$, which is a *sphere*. We apply these operations to collections elementwise, i.e. if $\mathcal{U}$ is a collection of balls we write $\partial \mathcal{U} = \{\partial B : B \in \partial U\}$, etc.

If $\mathcal{U}$ is a collection of spheres we define $\mathrm{rmin}\,\mathcal{U}$, $\mathrm{rmax}\,\mathcal{U}$ in the same way as for balls. For $R > 0$, we say the collection is *R-separated* if any two members are at distance at least R from each other. If this is true for $R = \mathrm{rmin}\,\mathcal{U}$ we say the collection is *well-separated*. Thus the t-boundaries of an R-separated collection of balls are $(R - 2t)$-separated. Note that an R-separated family of spheres may be nested: although the spheres are disjoint, the corresponding balls may be contained in each other.

A sequence $\mathcal{U}_1, \ldots, \mathcal{U}_p$ of carpets over E is called a *stack*, and p is its *height*.

Given a measure μ, a set F and a collection $\mathcal{U}$ of sets, we say that $\mathcal{U}$ *covers an ε-fraction of F* if $\mu(F \cap \bigcup \mathcal{U}) \ge \varepsilon \mu(F)$.

Lemma 3.3. *Let X be a metric space satisfying the Besicovitch and doubling properties and let $\chi = \chi(X)$. For $0 < \varepsilon, \delta < 1$ and $d \in \mathbb{N}$ let $p = \left\lceil \frac{2\chi}{\varepsilon\delta} \right\rceil$, and suppose that*

(1) *μ is a finite measure on X.*
(2) *$F \subseteq X$ is finite and $\mu(F) > \delta\mu(X)$.*
(3) *$\mathcal{U}_1, \ldots, \mathcal{U}_p$ is a stack over F with $\mathrm{rmin}\,\mathcal{U}_i \ge \mathrm{rmax}\,\mathcal{U}_{i-1}$.*
(4) *$\mu(\partial_1 B) > \varepsilon\mu(B)$ for each $B \in \bigcup_i \mathcal{U}_i$.*

Then there is an integer $k \ge 1$ and a subcollection $\mathcal{V} \subseteq \bigcup_{i \ge k} \mathcal{U}_i$ of spheres such that:

(a) *$\partial\mathcal{V}$ is well-separated.*
(b) *For $r = \mathrm{rmax}\,\mathcal{U}_{k-1}$, the set $\bigcup_{B \in \mathcal{V}} \partial_{2r} B$ contains more than $1/2$ of (the μ-mass of) F.*

Remark. If we assume that $\mathrm{rmin}\,\mathcal{U}_k > 4\,\mathrm{rmax}\,\mathcal{U}_{k-1}$ then we can conclude that the collection $\partial_{2r} B$, $B \in \mathcal{V}$, is pairwise disjoint.

Proof. For convenience we assume $\mu(X) = 1$. The proof follows the usual Vitali-like exhaustion scheme. We describe a recursive procedure for constructing $\mathcal{V}$, and show that it will eventually terminate with a suitable collection. Our induction hypothesis is that at the k-th stage we have constructed a collection $\mathcal{V} \subseteq \bigcup_{i > p-k} \mathcal{U}_i$ with $\partial\mathcal{V}$ well-separated, and with $\mu(\bigcup_{B \in \mathcal{V}} \partial_{2r} B) \ge \frac{\varepsilon\delta}{2\chi} \cdot k$ for $r = \mathrm{rmax}\,\mathcal{U}_{k-1}$.

We begin for $k = 0$ with $\mathcal{V} = \emptyset$, which satisfy this trivially. Assuming we have completed the k-th stage, let $r = \mathrm{rmax}\,\mathcal{U}_{k-1}$. Distinguish two cases.

If $\mu(F \cap \bigcup_{B \in \mathcal{V}} \partial_{2r} B) > \frac{1}{2}\mu(F)$, then $\mathcal{V}$ is the desired collection and we are done.

Otherwise let $G = F \setminus \bigcup_{B \in \mathcal{V}} \partial_{2r} B$, so $\mu(G) > \frac{1}{2} \mu(F) \geq \frac{1}{2} \delta$. By Corollary 3.2 we may choose a well-separated subcollection of balls $\mathcal{U}' \subseteq \mathcal{U}_{k-1}$ with $\mu(\bigcup \mathcal{U}') > \frac{\delta}{2\chi}$, so by assumption $\mu(\bigcup_{B \in \mathcal{U}'} \partial B) > \frac{\varepsilon \delta}{2\chi}$. Since the centers of $B \in \mathcal{U}'$ are at distance at least $2r \geq 2 \operatorname{rmax} \mathcal{U}'$ from each $S \in \partial \mathcal{V}$, the collection $\partial \mathcal{V} \cup \partial \mathcal{U}'$ is well-separated, and we have

$$\mu\Big(\bigcup_{B \in \mathcal{V} \cup \mathcal{U}'} \partial B\Big) = \mu\Big(\bigcup_{B \in \mathcal{V}} \partial B\Big) = \mu\Big(\bigcup_{B \in \mathcal{U}'} \partial B\Big) \geq \frac{\varepsilon \delta}{2\chi} k + \frac{\varepsilon \delta}{2\chi} = \frac{\varepsilon \delta}{2\chi}(k+1),$$

so we complete the recursive step by adding $\mathcal{U}'$ to $\mathcal{V}$.

It only remains to show that this cannot continue for p steps; and indeed, if it did we would have $\mu(\bigcup_{B \in \mathcal{V}} \partial B) > 1$, which is impossible. $\qquad \square$

4. Coarse dimension and non-concentration of mass on boundaries

For a metric space with a measure, let us say that a ball is ε-*thick* if an ε-fraction of its mass is concentrated on its boundary. In this section we derive a theorem which says, roughly, that given a finite measure on $\mathbb{R}^d$, only a relatively small mass of points can have the property that they lie at the center of many ε-thick balls. This result depends on a metric property that is closely related to topological dimension, which we call *coarse dimension*. Informally, we wish to express the fact that the boundary of balls is of a lower dimension than the ambient space. This is not quite what we need, since we are using thick boundaries in place of topological boundaries. In general it is not true that $\partial_1 B$ has lower dimension than X; in $\mathbb{R}^d$, for example, $\partial_1 B_r(x)$ has non-empty interior so it has full dimension. However, from the point of view of balls with radius $\gg 1$, $\partial_1 B_r(x)$ looks more or less like the lower-dimensional subset $\partial B_r(x)$ (and for balls whose radius is $\gg r$, $\partial_1 B_r(x)$ looks like a point). For this reason we introduce a parameter R_0 which specifies how big balls must be in order to pick up the "large scale" geometry. We make the following provisional definition, which is neither general nor particularly elegant, but is convenient for the induction which is to follow.

Definition 4.1. *For metric spaces X and $R_0 > 1$, the relation $\operatorname{cdim}_{R_0} X = k$ (read: X has coarse dimension k at scales $\geq R_0$) is defined by recursion on k:*

- $\operatorname{cdim}_{R_0} X = -1$ *for $X = \emptyset$ and any R_0,*
- $\operatorname{cdim}_{R_0} X = k$ *if $\operatorname{cdim}_{R_0} X \neq k - 1$ and, for every $t \geq 1$, every $r \geq t R_0$ and every $x \in X$, the subspace $Y = \partial_t B_r(x)$ satisfies $\operatorname{cdim}_{t R_0} Y = m$ for some $m \leq k - 1$.*

In showing that $\mathbb{R}^d$ has finite coarse dimension we use a property which is closely related to (and implies) the Besicovitch property, though the two are apparently not equivalent for general metric spaces.

Proposition 4.2. *Let $\| \cdot \|$ be a norm on $\mathbb{R}^d$. Then there are $R_0 > 1$ and $k \in \mathbb{N}$ with the following property. Suppose that $r(1) \geq \cdots \geq r(k) \geq R_0$ and $x_1, \ldots, x_k \in \mathbb{R}^d$ are such that $x_i \in \mathbb{R}^d \setminus \bigcup_{j<i} B_{r(j)-1}(x_j)$. Then $\bigcap_{i=1}^{k} \partial_1 B_{r(i)}(x_i) = \emptyset$.*

Proof. Let $x \in \mathbb{R}^d$ and $r > 1$. If $y \in \partial B_r(x)$, then the point $y' = 2x - y$ antipodally opposite to y on $\partial B_r(x)$ is at distance at least r from any ball $B_s(y)$ with $1 \leq s \leq r$. It follows that there is an $\varepsilon > 0$ such that if $z \in \partial B_r(x) \cap \partial B_s(y)$, then the angle $\angle xyz$ is greater than ε. By compactness of $\partial B_r(x)$ we can choose ε uniform in y. By continuity of the map $(u, v, w) \mapsto \angle uvw$, compactness and the assumption $s \geq 1$, we find that for some $\delta > 0$, the same remains true if we perturb y, z by δ.

Since the metric is translation invariant, we have shown the following: there is a $0 < \delta < 1$ such that, for any $1 \leq s \leq r$ and any three points $x \in \mathbb{R}^d$, $y \in \partial_{\delta/2} B_r(x)$ and $z \in \partial_{\delta/2} B_s(y)$, the angle $\angle xyz$ is at least ε. Rescaling and setting $R_0 = 2/\delta$, we find that if $R_0 \leq s \leq r$, $y \in \partial_1 B_r(x)$ and $z \in \partial_1 B_s(y)$ then $\angle xyz > \varepsilon$.

Returning to the situation in the formulation of the lemma, if $x \in \bigcap_{i=1}^k \partial_1 B_{r(i)}(x_i)$, then $\angle x_i x x_j > \varepsilon$ for all $1 \leq i < j \leq k$, and by compactness of the unit sphere this cannot happen for k arbitrarily large. The lemma follows. $\qquad\square$

Corollary 4.3. $\mathbb{R}^d$ *has finite coarse dimension with respect to any norm-induced metric.*

Proof. Let $\| \cdot \|$ be a fixed norm and let k', R_0 be the constants as in Proposition 4.2. Let k'' be the size of the maximal $(1 - 1/R_0)$-separated set of points in $B_2(0)$. Let $k = k'k''$; we claim that $\mathrm{cdim}_{R_0} \mathbb{R}^d \leq k$. Unraveling the definition of coarse dimension, it is apparent that in order to prove this it suffices to show that if we are given

(1) a sequence $t(1), \ldots, t(k) \geq 1$,
(2) a sequence $r(1), \ldots, r(k)$ such that $r(i) \geq t(1) \cdot \ldots \cdot t(i) R_0$,
(3) points $x_1, \ldots, x_k \in \mathbb{R}^d$ such that $x_i \in \partial_{t(j)} B_{r(j)}(x_j)$ for $j < i$,

then $\bigcap_{i=1}^k \partial_{t(i)} B_{r(i)}(x_i) = \emptyset$.

First, we claim that we may assume that each of the sequences is of length k', but that the radii are non-increasing. This will follow if we show that $r(j) \leq r(1)$ for some $2 \leq j \leq k'' + 1$, because we can then repeat this with $r(j)$ instead of $r(1)$, and so on k' times. To show that such a j exists, consider the points $x_2, \ldots, x_{k''+1}$ and suppose that $r(j) \geq r(1)$ for $2 \leq j \leq k'' + 1$. Observe that by (3) the x_j are all located within the ball $B_{r(1)+t(1)}(x_1) \subseteq B_{2r(1)}(x_1)$, because by (2), $t(1) \leq r(1)/R_0 \leq r(1)$. Also, by (2) and (3), if $i > j$ then

$$d(x_i, x_j) \geq r(j) - t(j) \geq r(j)(1 - 1/R_0)$$

so that if $r(j) \geq r(1)$ we have $d(x_i, x_j) \geq r(1)(1 - 1/R_0)$. Thus $x_2, \ldots, x_{k''+1}$ is an $r(1)(1-1/R_0)$-separated set in the ball $B_{2r(1)}(x_1)$, and rescaling we obtain a contradiction to the definition of k''.

Now we assume the sequences have length k' and the radii are non-increasing. Let $t = \max t_i$ and replace $\| \cdot \|$ with $\| \cdot \|^* = (1/t)\| \cdot \|$. After this rescaling, we wish to show that $\bigcap_{i=1}^{k'} \partial_{t_i/t}^* B_{r(i)/t}^*(x_i) = \emptyset$, where the $*$'s indicate operations with respect to $\| \cdot \|^*$. For this it is enough to show that $\bigcap_{i=1}^{k'} \partial_1^* B_{r(i)/t}^*(x_i) = \emptyset$, and this will follow once we verify the hypothesis of the previous proposition for the norm $\| \cdot \|^*$; and indeed, clearly $r(i)/t$ is still decreasing; $x_i \in \partial_{t(j)} B_{r(j)}(x_j)$ implies $x_i \in \partial_1^* B_{r(j)/t}^*(x_j)$; and the inequalities $r(i) \geq r(k) \geq t(1) \cdot \ldots \cdot t(k') R_0$ and $t_i \geq 1$ imply $r(i)/t \geq R_0$, as required there. $\qquad\square$

We can now state and prove the main result of this section, which is the main tool in the proof of Theorem 1.2. Although it is $\mathbb{R}^d$ that we have in mind, the formulation is for general metric spaces in order to facilitate the inductive proof.

Theorem 4.4. *Fix $k, \chi \in \mathbb{N}$ and $0 < \varepsilon, \delta < 1$ and set $q = \left(\frac{200\chi^2}{\varepsilon^2\delta^3}\right)^k \cdot 1000^{k^2}$. Suppose that*

(1) *X is a metric space with $\chi(X) \leq \chi$ and $\mathrm{cdim}_{R_0} X = k$ for some $R_0 > 2$,*
(2) *μ is a finite measure on X,*
(3) *$F \subseteq X$ is finite,*
(4) *$\mathcal{U}_1, \ldots, \mathcal{U}_q$ is a stack over F with*
    (a) *$\mathrm{rmin}\,\mathcal{U}_i \geq (\mathrm{rmax}\,\mathcal{U}_{i-1})^2$,*
    (b) *$\mathrm{rmin}\,\mathcal{U}_1 \geq \max\{2, R_0\}$,*
(5) *$\mu(\partial_1 B) \geq \varepsilon\mu(B)$ for each $B \in \bigcup_i \mathcal{U}_i$.*

Then $\mu(F) \leq \delta\mu(X)$.

Remark. No attempt has been made to optimize the conditions. A slower rate of growth in (4a) would probably suffice.

Proof. Define integers $Q(k, \chi, \varepsilon, \delta)$ recursively by

$$Q(0, \chi, \varepsilon, \delta) = 1,$$
$$Q(k, \chi, \varepsilon, \delta) = \left\lceil \frac{2\chi}{\varepsilon\delta} \right\rceil \cdot \left(1 + \left\lceil \frac{64\chi}{\varepsilon\delta^2} \right\rceil\right) \cdot \left(1 + Q\left(k - 1, \chi, \frac{\varepsilon}{2}, \frac{\delta}{8}\right)\right).$$

One may verify that $q \geq Q(k, \chi, \varepsilon, \delta)$, so it suffices to prove the claim for $q = Q(k, \chi, \varepsilon, \delta)$; this we do by induction on k.

For $k = 0$ the claim is trivial, since then $\partial_1 B = \emptyset$ for any ball B. We can therefore have $\mu(\partial_1 B) \geq \varepsilon\mu(B)$ only when $\mu(B) = 0$, implying that each point in F has mass 0, so $\mu(F) = 0$.

Assume that the claim holds for $k - 1$. We suppose that $X, R_0, \chi, \mu, F, \mathcal{U}_1, \ldots, \mathcal{U}_q$ satisfy the hypotheses of the theorem but $\mu(F) > \delta\mu(X)$, and proceed to derive a contradiction.

Preliminary disjointification. We first pass to a subsequence of the given carpets and extract a disjoint family of balls from them. Let

$$N = q / \left\lceil \frac{2\chi}{\varepsilon\delta} \right\rceil.$$

Since $q/N = \left\lceil \frac{2\chi}{\varepsilon\delta} \right\rceil$, we may apply Corollary 3.2 to the stack $\{\mathcal{U}_{iN}\}_{1 \leq i \leq q/N}$ obtained by choosing each N-th element of the original stack. We get an $n_0 \geq 1$ and a collection

$$\mathcal{V} \subseteq \bigcup_{i \geq N(n_0+1)} \mathcal{U}_i$$

such that $\partial\mathcal{V}$ is well-separated, and such that, setting $r = \mathrm{rmax}\,\mathcal{U}_{n_0}$ we have

$$\mu\left(F \cap \bigcup_{B \in \mathcal{V}} \partial_{2r} B\right) \geq \frac{1}{2}\mu(F) > \frac{\delta}{2}\mu(X).$$

We denote the union on the left hand side by

$$Y = \bigcup_{B \in \mathcal{V}} \partial_{2r} B.$$

From now on we can forget about the carpets $\mathcal{U}_i$ for $i < n_0 N$ and $i > (n_0 + 1)N$; we work only with $\mathcal{V}$ and $\mathcal{U}_i$, $n_0 N \leq i \leq (n_0 + 1)N$.

Outline of the argument. Roughly, our argument proceeds as follows. The set Y is made up of a union of thick spheres, and each is of lower coarse dimension than X. Let S be one of these spheres, and suppose that some non-trivial fraction of its mass comes from F. Consider the stack obtained by fixing a large p (but still much smaller than N) and selecting from the stack $\mathcal{U}_{n_0 N}, \mathcal{U}_{n_0 N+1}, \ldots, \mathcal{U}_{n_0 N+p}$ those balls centered in $F \cap S$. The induction hypothesis can be applied to show that for a non-trivial fraction of points $x \in F \cap S$ there is a ball in this stack whose 1-boundary *with respect to S* contains only an $\varepsilon/2$-fraction of the ball's mass. However, with respect to X these 1-boundaries contain an ε-fraction of the mass. Therefore, the difference—an $\varepsilon/2$ of the ball's mass—lies outside S. Passing to a disjoint subcollection of these balls centered in S, we obtain a set of mass equal to some non-trivial fraction of $F \cap S$, located outside of S but nearby. Now, since the mass of $F \cap Y$ is large, the situation described can be repeated for spheres S containing a non-negligible fraction of Y, and the masses obtained outside each sphere will be disjoint from each other. We conclude that, in the near vicinity of Y but disjoint from Y there is a set with mass a small but constant fraction of $\mu(X)$. Next, we repeat this argument, replacing Y with a small neighborhood of Y, and using the next p carpets $\mathcal{U}_{n_0 N+p+1}, \ldots, \mathcal{U}_{n_0 N+2p}$, and get another mass increment. After doing this sufficiently many times we will have accumulated more mass than there is in X altogether, a contradiction.

Partitioning into further substacks. Let us denote

$$r_i^+ = \operatorname{rmax} \mathcal{U}_i, \qquad r_i^- = \operatorname{rmin} \mathcal{U}_i,$$

and set

$$p = Q(k-1, \chi, \varepsilon/2, \delta/8).$$

We partition the carpets $\{\mathcal{U}_i\}_{n_0 N+1 \leq i \leq (n_0+1)N-1}$ into substacks of height $p+1$. More precisely, let $M = N/(p+1)$, and for $0 \leq j \leq M-1$ define

$$m(j) = n_0 N + 1 + (p+1) \cdot j$$

so for each such j we get the stack $\{\mathcal{U}_{m(j)+i}\}_{1 \leq i \leq p}$. The first thing to note is that all the balls in these stacks are from carpets $\mathcal{U}_i$ below $\mathcal{U}_{(n_0+1)N}$, which is the level where $\mathcal{V}$ begins. Consequently, the radii of all these balls are much smaller than the radii of balls in $\mathcal{V}$; indeed, the largest possible radius in our substacks is

$$r_{m(M-1)+p}^+ = r_{(n_0+1)N-1}^+ < \frac{1}{r_{(n_0+1)N-1}^+} r_{(n_0+1)N}^+ < \frac{1}{2} \operatorname{rmin} \mathcal{V}$$

by (4a) and (4b).

Thickening the set Y. For $0 \le j \le M - 1$ it will be convenient to denote

$$\Delta_j B = \partial_{r^+_{m(j)}} B$$

and to thicken the set Y by thickening each sphere in Y, obtaining

$$Y_j = \bigcup_{B \in \mathcal{V}} \Delta_j B.$$

Let us note several properties of the Y_j. First, clearly $Y_1 \subseteq \cdots \subseteq Y_M$, and $Y \subseteq Y_j$, implying

$$\mu(F \cap Y_j) \ge \mu(F \cap Y) \ge \frac{\delta}{2}\mu(X),$$

To see this it is enough to show that $Y \subseteq Y_0$, and indeed by (4a) and (4b),

$$r^+_{m(0)} = r^+_{n_0 N + 1} > r^2 \ge 2r,$$

so, since Y_0, Y are obtained, respectively, as the $r^+_{m(0)}$-thickening and r-thickening of the same spheres, the claim follows.

Second, each Y_j is the disjoint union of the thick spheres $\Delta_j B$, $B \in \mathcal{V}$. This follows from the inequality $r^+_m(j) < \frac{1}{2} \operatorname{rmin} \mathcal{V}$, noted above, and the fact that $\mathcal{V}$ is well-spaced.

Third, let $0 \le j \le M - 2$, and $x \in \partial_{r^+_{m(j)}} B$ for some $B \in \mathcal{V}$. Suppose that $B' \in \mathcal{U}_{m(j)+i}$ for some $1 \le i \le p$ is centered at x; then $\partial_1 B' \subseteq \Delta_{j+1} B$. To see this, suppose that $y \in B$ and $z \in \partial_1 B'$. Then

$$d(y, z) \le d(y, x) + d(x, z) \le r^+_{m(j)} + r^+_{m(j)+i} + 1 \le 2r^+_{m(j)+p} < r^+_{m(j+1)},$$

which proves the claim.

Spheres in Y_j containing a large proportion of F. For each $0 \le j \le M - 1$, define

$$\mathcal{W}_j = \left\{ B \in \mathcal{V} : \mu(F \cap \Delta_j B) > \frac{\delta}{4}\mu(\Delta_j B) \right\}.$$

A Markov-type argument now shows that the spheres in $\mathcal{W}_j$ contain a large fraction of X:

$$\frac{\delta}{2}\mu(X) < \mu(F \cap Y_j)$$

$$= \mu\left(F \cap \bigcup_{B \in \mathcal{V}} \Delta_j B\right) = \mu\left(F \bigcup_{B \in \mathcal{W}_j} \Delta_j B\right) + \mu\left(F \cap \bigcup_{B \in \mathcal{V} \setminus \mathcal{W}_j} \Delta_j B\right)$$

$$\le \mu\left(F \cap \bigcup_{B \in \mathcal{W}_j} \Delta_j B\right) + \frac{\delta}{4}\mu\left(\bigcup_{B \in \mathcal{V} \setminus \mathcal{W}_j} \Delta_j B\right)$$

$$\le \mu\left(F \cap \bigcup_{B \in \mathcal{W}_j} \Delta_j B\right) + \frac{\delta}{4}\mu(X)$$

and, rearranging, we get

$$\mu\left(F \cap \bigcup_{B \in \mathcal{W}_j} \Delta_j B\right) > \frac{\delta}{4}\mu(X).$$

Applying the induction hypothesis to fat spheres. Fix $0 \le j \le M-1$ and let $S = \Delta_j B$ for some $B \in \mathcal{W}_j$. Put $\mu_S = \mu|_S$, i.e. $\mu_S(A) = \mu(A \cap S)$, and $F_S = F \cap S$.

Consider the stack $\{\mathcal{U}_t'\}_{1 \le t \le p}$ over F_S obtained by selecting from $\{\mathcal{U}_{m(j)+t}\}_{1 \le t \le p}$ those balls with centers in F_S. This is a stack in X, but from it we get a stack in S by intersecting each ball with S.

We claim that S, μ_S, F_S and this stack satisfy conditions (1) to (4b) of the theorem, with $k-1$ in place of k and $r_{m(j)}^+ R_0$ in place of R_0. Indeed, $S = \partial_{r_{m(j)}^+} B$, so it has coarse dimension $\le k-1$ at scales $\ge r_{m(j)}^+ R_0$, and $\chi(S) \le \chi(X)$ because $S \subseteq X$. The relative growth of radii in $\mathcal{U}_i'$ is inherited from $\mathcal{U}_i$. Finally, $\mathcal{U}_1' \subseteq \mathcal{U}_{m(j)+1}$, so by the growth assumption for the original stack,

$$\mathrm{rmin}\,\mathcal{U}_1' \ge r_{m(j)+1}^- \ge (r_{m(j)}^+)^2 \ge r_{m(j)}^+ R_0$$

and clearly also $\mathrm{rmin}\,\mathcal{U}_1' \ge 2$, which verifies (4b).

Let

$$F_S' = \left\{x \in F_S : \mu_S(\partial_1 B') \ge \frac{\varepsilon}{2}\mu_S(B') \text{ for every } B' \in \mathcal{U}_t', \, 1 \le t \le p\right\}$$

Applying the induction hypothesis to the stack obtained by restricting each $\mathcal{U}_t'$ to balls with center in F_S', and recalling the definition of p, we find that

$$\mu_S(F_S') \le \frac{\delta}{8}\mu_S(S).$$

Since $S = \Delta_j B$ for some $B \in \mathcal{W}_j$, we know that $\mu_S(F_S) > (\delta/4)\mu_S(S)$; so

$$\mu_S(F_S \setminus F_S') > \frac{\delta}{8}\mu_S(S) = \frac{\delta}{8}\mu(S).$$

Estimating the mass outside of a fat sphere. For each $x \in F_S \setminus F_S'$ there are some $1 \le t \le p$ and $B' \in \mathcal{U}_t'$, centered at x, with

$$\mu(\partial_1 B' \cap S) = \mu_S(\partial_1 B') \le \frac{\varepsilon}{2}\mu_S(B' \cap S).$$

But by (5) we have

$$\mu(\partial_1 B') \ge \varepsilon\mu(B') \ge \varepsilon\mu_S(B');$$

therefore,

$$\mu(\partial_1 B' \setminus S) \ge \frac{\varepsilon}{2}\mu(B' \cap S).$$

Estimating the mass between Y_j and Y_{j+1}. Applying Corollary 3.2 to each of the balls above as x runs over $F_S \setminus F_S'$, we choose a disjoint collection $\mathcal{C}$ of balls centered in $F_S \setminus F_S'$

satisfying the last inequality, and which cover a $\frac{1}{\chi}$-fraction of $F_S \setminus F_S'$ and so has mass $> \frac{\delta}{8\chi}\mu(S)$. The corresponding union of 1-spheres, since each contains an ε-fraction of the mass of the solid ball, has mass $> \frac{\varepsilon\delta}{8\chi}\mu(S)$; and since at least half this mass lies outside of S, we get

$$\mu\left(\left(\bigcup_{B' \in \mathcal{C}} \partial_1 B'\right) \setminus S\right) \geq \frac{\varepsilon\delta}{16\chi}\mu(S).$$

The set on the left hand side is in the complement of $S = \Delta_j B$, but certainly lies inside $\Delta_{j+1} B$, and these sets are disjoint for distinct $B \in \mathcal{V}$. Hence the contribution of mass near each S is disjoint from the contributions of other S's, so

$$\mu(Y_{j+1} \setminus Y_j) \geq \sum_{B \in \mathcal{W}_j} \mu(\Delta_{j+1} B \setminus \Delta_j B) \geq \sum_{B \in \mathcal{W}_j} \frac{\varepsilon\delta}{16\chi}\mu(\Delta_j B)$$

$$= \frac{\varepsilon\delta}{16\chi}\mu\left(\bigcup_{B \in \mathcal{W}_j} \Delta_j B\right) \geq \frac{\varepsilon\delta}{16\chi} \cdot \frac{\delta}{4}\mu(X)$$

because $\mu(\bigcup_{B \in \mathcal{W}_j} \Delta_j B) \geq \mu(Y_j) > \frac{\delta}{4}\mu(X)$.

The punchline. The number of Y_j's has been arranged to be

$$M = \frac{N}{p+1} = \frac{q}{\lceil 2\chi/\varepsilon\delta\rceil(p+1)} \geq \frac{64\chi}{\varepsilon\delta^2} + 1$$

so the relations $Y_j \subseteq Y_{j+1}$ and $\mu(Y_{j+1} \setminus Y_j) > \frac{\varepsilon\delta}{16\chi} \cdot \frac{\delta}{4}\mu(X)$ imply

$$\mu(Y_M) \geq M \cdot \frac{\varepsilon\delta}{16\chi} \cdot \frac{\delta}{4}\mu(X) > \mu(X),$$

which is the desired contradiction.   □

It is not hard to see that a similar result holds if μ is a Borel measure, F is a Borel set and the carpets are measurable (i.e. the function $r_n : F \to \mathbb{R}^+$ describing the radii of the balls in the n-th carpet is measurable). One way to see this is to discretize the data. For a fine partition $\mathcal{P} = \{P_i\}$ of X, choose a representative $x_i \in P_i$ in each atom, set $r_n'(x) = \int_{P_i} r_n$, and replace μ with the atomic measure supported on the x_i with $\mu'(\{x_i\}) = \mu(P_i)$. Applying the discrete lemma above to μ' and the new stack, with suitably modified parameters, we can deduce the result for the original measure.

5. Proof of the ratio theorem for $\mathbb{Z}^d$

Given what we have proven so far, Theorems 1.2 and 1.1 now follow by fairly standard arguments.

Proof of Theorem 1.2. This is a standard application of the transference together with Theorem 4.4. Let $\mathbb{Z}^d$ act on a σ-finite measure space $(\Omega, \mathcal{B}, \mu)$ by non-singular transformations. By passing to an equivalent measure, we may assume that $\mu(\Omega) = 1$. Let T

act by translation on L^∞, and let $\widehat{T}$ be the dual action of T on $L^1 \subseteq (L^\infty)^*$, which is a linear, order-preserving isometry defined by the condition $\int \widehat{T}^u f \cdot g \, d\mu = \int f \cdot T^u g \, d\mu$ for $f \in L^\infty$ and $g \in L^1$, and explicitly by

$$\widehat{T}^u f(\omega) = f(T^{-u}\omega) \cdot \frac{dT^u\mu}{d\mu}.$$

Fix a norm $\|\cdot\|$ on $\mathbb{R}^d$, and let $\mathrm{cdim}_{R_0} \mathbb{R}^d = k$ and $\chi(\mathbb{R}^d) = \chi$ for appropriate parameters R_0, k, χ.

Let $1 \le f \in L^\infty \subseteq L^1$. We are out to prove that

$$s_n(\omega) = \frac{\sum_{u \in \partial_1 B_n} \widehat{T}^u f(\omega)}{\sum_{u \in B_n} \widehat{T}^u f(\omega)} \to 0$$

for a.e. ω (from this the case of thick boundaries ∂_t follows by rescaling the norm). Set

$$A_\varepsilon = \{\omega \in \Omega : \limsup s_n(\omega) > \varepsilon\}$$

and suppose that $\mu(A_\varepsilon) > 0$ for some ε. We construct a sequence

$$2R_0 = r_0^- = r_0^+ \le r_1^- \le r_1^+ \le r_2^- \le r_2^+ \le \cdots$$

satisfying $r_i^- \ge (r_{i-1}^+)^2$ and $r_1^- \ge \max\{2, R_0\}$, and a set of points $A \subseteq A_\varepsilon$, so that for every $\omega \in A$ and $i \ge 1$ there is an $n_i = n_i(\omega) \in (r_i^-, r_i^+)$ with $s(n_i, \omega) > \varepsilon$, and $\mu(A) > \frac{1}{2}\mu(A_\varepsilon)$. We do this by recursion, so that going into the i-th stage we have defined $r_j^\pm$ for $j < i$ and sets $A_\varepsilon \supseteq C_0 \supseteq C_1 \supseteq \cdots \supseteq C_{i-1}$ satisfying the above and $\mu(C_j) \ge (\frac{1}{2} + \frac{1}{j+1})\mu(A_\varepsilon)$. In order to define $r_i^\pm$ and C_i, first set $r_i^- = (2 \vee r_{i-1}^+)^2$. Now, since $C_{i-1} \subseteq A_\varepsilon$, for every $\omega \in C_{i-1}$ there is an $n = n(\omega) \ge r_i^-$ with $s(n, \omega) > \varepsilon$; so we can choose r_i^+ so that $n(\omega) \le r_i^+$ on a subset of C_{i-1} of measure $> (\frac{1}{2} + \frac{1}{i+1})\mu(A_\varepsilon)$. This set will be C_i, and $A = \bigcap_{j=1}^\infty C_j$.

We are now ready to apply the transference principle. Fix δ and $n > r_q^+$, where $q = q(k, \chi, \varepsilon, \delta)$ is as in Theorem 4.4. Then

$$\mu(A) = \int 1_A \, d\mu = \frac{1}{|B_n|} \sum_{u \in B_n} \int \widehat{T}^u 1_A \, d\mu = \frac{1}{|B_n|} \int \sum_{u \in B_n} \widehat{T}^u 1_A \, d\mu.$$

Next, we bound the sum $\sum_{u \in B_n} \widehat{T}^u 1_A$. Fix $\omega \in \Omega$ and consider the measure $\nu = \nu_{\omega,n}$ on B_{2n} defined by $\nu(\{u\}) = \widehat{T}^u f(\omega)$. Let

$$U = U_{\omega,n} = \{u \in B_n : T^{-u}\omega \in A\}.$$

By the definition of A there is a stack of height q over U satisfying the hypothesis of Theorem 4.4, and all the balls in the stack are of radius $\le r_q^+ < n$, implying that they are contained in B_{2n}. Thus, by Theorem 4.4,

$$\nu(U) \le \delta\nu(B_{2n}) = \delta \sum_{u \in B_{2n}} \widehat{T}^u f(\omega) \le \delta\|f\|_\infty \sum_{u \in B_{2n}} \widehat{T}^u 1.$$

We have arranged things so that

$$\sum_{u \in B_n} \widehat{T}^u 1_A(\omega) = \nu_{\omega,n}(U_{\omega,n}),$$

therefore

$$\sum_{u \in B_n} \widehat{T}^u 1_A(\omega) \le \nu_{\omega,n}(U_{\omega,n}) \le \delta \|f\|_\infty \sum_{u \in B_{2n}} \widehat{T}^u 1.$$

Dividing by $|B_n|$ and integrating we get

$$\mu(A) \le \frac{1}{|B_n|} \int \delta \|f\|_\infty \sum_{u \in B_{2n}} \widehat{T}^u 1 \, d\mu \le \frac{\delta \|f\|_\infty}{|B_n|} |B_{2n}| \int 1 \, d\mu \le 2^d \|f\|_\infty \delta.$$

because $|B_n|/|B_{2n}| \le 2^d$. The right hand side can be made arbitrarily small, so $\mu(A) = 0$; hence also $\mu(A_\varepsilon) = 0$.

Finally, $s_n(\omega) \to 0$ if and only if $\omega \notin \bigcup_{m=1}^\infty A_{1/m}$, and the set on the right is seen to have measure 0. This completes the proof of Theorem 1.2. $\qquad\square$

Proof of Theorem 1.1. The proof is standard. By passing to an equivalent measure we may assume that μ is finite. We first prove the case $g \equiv 1$. Consider the space

$$\mathcal{F} = \mathrm{span}\{1, f - \widehat{T}^v f : v \in \mathbb{Z}^d \text{ and } f \in L^\infty\}.$$

One shows that $\mathcal{F}$ is dense in L^1; the proof follows the same lines as Riesz's proof of the mean ergodic theorem, using the duality relation $(L^1)^* = L^\infty$ instead of self-duality of L^2. See [Fel07, Aar97].

Next, one shows that the ratios $R_n(f, 1)$ converge for every member f of $\mathcal{F}$. Indeed, note that $R_n(1, 1) \equiv 1$; whereas if $f \in L^\infty$ then the ratios $R_n(f - T^v f, 1)$ satisfy

$$\left| \frac{\sum_{u \in B_n} \widehat{T}^u (f - \widehat{T}^v f)}{\sum_{u \in B_n} \widehat{T}^u 1} \right| \le \frac{\sum_{u \in \partial_{\|v\|} B_n} \widehat{T}^u |f|}{\sum_{u \in B_n} \widehat{T}^u 1} = \frac{\sum_{u \in \partial_{\|v\|} B_n} \widehat{T}^u |f|}{\sum_{u \in B_n} \widehat{T}^u |f|} \cdot \frac{\sum_{u \in B_n} \widehat{T}^u |f|}{\sum_{u \in B_n} \widehat{T}^u 1}$$

$$\le \frac{\sum_{u \in \partial_{\|v\|} B_n} \widehat{T}^u |f|}{\sum_{u \in B_n} \widehat{T}^u |f|} \cdot \|f\|_\infty$$

and the right hand side converges to 0 a.e. by Theorem 1.2. From this it follows that $R_n(f, 1) \to \int f$ for any $f \in \mathcal{F}$.

The case $g \equiv 1$ is concluded by applying the maximal inequality to get convergence on the closure of $\mathcal{F}$, which is all of L^1. This standard argument can be found in [Aar97]. It is also easy to check that the correct limit is obtained.

Finally, the case of general $g \in L^1$ is deduced from the equality $R_n(f, g) = R_n(f, 1)/R_n(g, 1)$. $\qquad\square$

References

[Aar97]   Aaronson, J.: An Introduction to Infinite Ergodic Theory. Math. Surveys Monogr. 50, Amer. Math. Soc., Providence, RI (1997)   Zbl 0882.28013   MR 1450400

[Bec83]   Becker, M. E.: A ratio ergodic theorem for groups of measure-preserving transformations. Illinois J. Math. **27**, 562–570 (1983)   Zbl 0523.28022   MR 0720092

[Bes45]   Besicovitch, A. S.: A general form of the covering principle and relative differentiation of additive functions. Proc. Cambridge Philos. Soc. **41**, 103–110 (1945)   Zbl 0063.00352   MR 0012325

[CO60]   Chacon, R. V., Ornstein, D. S.: A general ergodic theorem. Illinois J. Math. **4**, 153–160 (1960)   Zbl 0134.12102   MR 0110954

[dG75]   de Guzmán, M.: Differentiation of Integrals in R^n. Lecture Notes in Math. 481, Springer, Berlin (1975)   Zbl 0327.26010   MR 0457661

[Fel07]   Feldman, J.: A ratio ergodic theorem for commuting, conservative, invertible transformations with quasi-invariant measure summed over symmetric hypercubes. Ergodic Theory Dynam. Systems **27**, 1135–1142 (2007)   Zbl 1136.37005   MR 2342969

[Hoc06]   Hochman, M.: Return times, recurrence densities and entropy for actions of some discrete amenable groups. J. Anal. Math. **100**, 1–51 (2006)   Zbl 1173.37303   MR 2303303

[Hop]   Hopf, E.: Ergodentheorie. Ergeb. Math. Grenzgeb. 5, Heft 2, Springer (1937)   JFM 63.0786.07

[Hur44]   Hurewicz, W.: Ergodic theorem without invariant measure. Ann. of Math. (2) **45**, 192–206 (1944)   Zbl 0063.02944   MR 0009427

[Kre85]   Krengel, U.: Ergodic Theorems. de Gruyter Stud. Math. 6, de Gruyter, Berlin (1985)   Zbl 0575.28009   MR 0797411

[Lin01]   Lindenstrauss, E.: Pointwise theorems for amenable groups. Invent. Math. **146**, 259–295 (2001)   Zbl 1038.37004   MR 1865397

[Lin06]   Lindenstrauss, E: Invariant measures and arithmetic quantum unique ergodicity. Ann. of Math. (2) **163**, 165–219 (2006)   Zbl 1104.22015   MR 2195133

[Mor47]   Morse, A. P.: Perfect blankets. Trans. Amer. Math. Soc. **61**, 418–442 (1947)   Zbl 0031.38702   MR 0020618

[OW87]   Ornstein, D. S., Weiss, B.: Entropy and isomorphism theorems for actions of amenable groups. J. Anal. Math. **48**, 1–141 (1987)   Zbl 0637.28015   MR 0910005

[Rig04]   Rigot, S.: Counter example to the Besicovitch covering property for some Carnot groups equipped with their Carnot–Carathéodory metric. Math. Z. **248**, 827–848 (2004)   Zbl 1082.53030   MR 2103544

[Wei03]   Weiss, B.: Actions of amenable groups. In: Topics in Dynamics and Ergodic Theory, London Math. Soc. Lecture Note Ser. 310, Cambridge Univ. Press, Cambridge, 226–262 (2003)   Zbl 1079.37002   MR 2052281