MAGNITUDE, DIVERSITY, CAPACITIES, AND DIMENSIONS OF METRIC SPACES

MARK W. MECKES

Abstract. Magnitude is a numerical invariant of metric spaces introduced by Leinster, motivated by considerations from category theory. This paper extends the original definition for finite spaces to compact spaces, in an equivalent but more natural and direct manner than in previous works by Leinster, Willerton, and the author. The new definition uncovers a previously unknown relationship between magnitude and capacities of sets. Exploiting this relationship, it is shown that for a compact subset of Euclidean space, the magnitude dimension considered by Leinster and Willerton is equal to the Minkowski dimension.

1. Introduction

The magnitude of a metric space is a numerical isometric invariant introduced by Leinster in [9]. From the perspective of geometry, its definition was motivated in a rather unusual way. In [6], Leinster had defined the Euler characteristic of a finite category, which generalizes the Euler characteristic of a topological space or of a poset. This notion of Euler characteristic can be naturally generalized from categories to to enriched categories, a family of algebraic structures which, as observed by Lawvere in [5], includes metric spaces; in this context the generalization of Euler characteristic is named "magnitude". Specialized then to metric spaces, one obtains Leinster's definition of the magnitude of a finite metric space, stated in Definition 2.1 below. Magnitude was extended to compact metric spaces in multiple ways in [9, 11, 19, 20], which were shown by the author in [13] to agree with each other for many spaces (specifically, for so-called positive definite spaces, which include all compact subsets of Euclidean space).

Given this exotic provenance, it may come as a surprise that magnitude turns out to be closely related to classical invariants of integral geometry; see [9, 11, 20] for a number of results along these lines. Conjectures in [9, 11], which are supported by partial results in those papers and by heuristics and numerical computations in [19], suggest that the relationship between magnitude and integral geometry runs deeply enough that all the intrinsic volumes of convex bodies can be recovered from magnitude. A second surprise, in a completely different direction, is that the magnitude of a finite metric space has been introduced in the literature before, in connection with quantifying biodiversity in [16]. Although the theory of magnitude was not developed at all in [16], the relationship between magnitude and diversity has been investigated more fully in [7].

The present work grew out of the author's search for a more satisfactory definition of magnitude for compact metric spaces. The approach of [13] was to introduce yet another definition, a measure-theoretic generalization of a variational formula for the magnitude of a finite positive definite space derived in [7, 9], and to prove that this new definition agrees with all the earlier ones. Here we instead take a more functional-analytic approach to generalize directly the original definition of magnitude for finite metric spaces. Besides the æsthetic appeal of a more direct approach, the resulting definition, which again agrees with the earlier ones, can be used to prove new properties of magnitude in Euclidean space. More interestingly, it uncovers previously unknown connections between magnitude and potential theory. In fact, another surprise is that the magnitude of a compact subset of Euclidean space has (almost) been introduced in an equivalent form in the literature before both [9] and [16], as a type of capacity. (It seems likely that magnitude is the unique notion to have arisen independently in potential theory, theoretical ecology, and category theory.)

The relationship between magnitude and capacity has important consequences. There is a notion of dimension associated to magnitude which was first investigated in [11], and which provided some of the first compelling evidence that magnitude encodes interesting geometric information. Using in part a deep result on relationships between different capacities, we will see that in Euclidean space, this magnitude dimension turns out to be the same as Minkowski dimension. In establishing this result, we find an apparently new formulation of Minkowski dimension in terms of capacity. In addition, the conjectures from [9, 11] mentioned earlier would, if true, indicate previously unknown connections between capacities and intrinsic volumes of convex bodies.

The rest of this paper is organized as follows. Section 2 presents Leinster's original definition of the magnitude of a finite metric space and some related definitions we will need. Section 3 develops the functional-analytic generalization of the definition of magnitude for compact spaces. Section 4 presents a dual perspective on the definitions of Section 3, and discusses a quantity closely related to magnitude, the maximum diversity of a metric space. Section 5 specializes the constructions of the previous sections to subsets of Euclidean space, and uses them to prove new results about the behavior of magnitude in Euclidean space. Section 6 discusses the connections between magnitude, maximum diversity, and capacities. Section 7 proves a new characterization of Minkowski dimension in terms of maximum diversity, and uses a result from potential theory recalled in Section 6 to deduce that magnitude dimension and Minkowski dimension are equal in Euclidean space. Finally, in Section 8, we briefly investigate another, closely related instance of the magnitude of an enriched category: the case of ultrametric spaces, whose theory turns out to be much simpler but nevertheless intriguingly similar to that of metric spaces.

Acknowledgements. The author thanks Pierre Albin, Shiri Artstein-Avidan, Harsh Mathur, and especially Tom Leinster for helpful discussions. Particular thanks are due to Charles Clum for convincing the author of the usefulness of the potential function in analyzing magnitude. This work was supported in part by NSF grant DMS-0902203.

2. Finite metric spaces

We now recall the definition of the magnitude of a finite metric space.

Definition 2.1. *Given a finite metric space (A, d), define the matrix $\zeta \in \mathbb{R}^{A\times A}$ by $\zeta(a, b) := e^{-d(a,b)}$. A vector $w \in \mathbb{R}^A$ is a* **weighting** *for A if for each $a \in A$,*

$$(\zeta w)(a) = \sum_{b\in A} e^{-d(a,b)} w(b) = 1. \tag{2.1}$$

If A possesses a weighting w, then the **magnitude** *of A is*

$$|A| := \sum_{a\in A} w(a). \tag{2.2}$$

An arbitrary metric space need not possess a weighting (see [9, Example 2.2.7]), nor must a weighting be unique. However, it is easy to check that if a weighting exists, then the value of the sum in (2.2) is independent of the weighting. Thus magnitude is in general only partially defined on the class of finite metric spaces, but well-defined when it is defined. The magnitude of A is always defined if the matrix ζ is invertible, which is the case whenever A is a subset of Euclidean space [9, Theorem 2.5.3].

There is an arbitrary choice of scale implicit in Definition 2.1: instead of the metric space $A = (A, d)$, one may equally well consider any of the metric spaces $tA = (A, td)$ for $t > 0$. We thus define the **magnitude function** of A to be the function $t \mapsto |tA|$ for $t > 0$. In general the magnitude function may be only partially defined, but it is always defined for all but finitely many values of t [9, Proposition 2.2.6].

Observe that if $A \subseteq \mathbb{R}^n$, then the metric space tA as defined above is isometric to the space $\{ta \mid a \in A\} \subseteq \mathbb{R}^n$; as is standard, we will thus use tA to denote this latter set without fear of ambiguity. In particular, the magnitude function of a finite subset of Euclidean space is defined everywhere.

If we write ζ_t for the matrix associated to tA by Definition 2.1, so that $\zeta_t(a, b) = e^{-td(a,b)}$, then ζ_t tends toward the identity matrix indexed by A when $t \to \infty$. From this it follows that for sufficiently large t, tA has a unique weighting which tends to the vector whose entries are all 1, and the magnitude of tA tends to the cardinality of A. (See [9, Proposition 2.2.6] for more details.) This suggests the intuitive interpretation of magnitude as the "effective number of points" of A when A is viewed at a particular scale. (Indeed, in [16], the magnitude of A is called the "effective number of species" in an ecosystem whose species are represented by the points of A, when $e^{-d(a,b)}$ is interpreted as the similarity between species a and species b.)

3. Weightings and magnitude

In this and the next section, it will make no difference whether we choose to work with real or complex scalars. However, in section 5, Fourier-analytic tools will be brought to bear and it will be more natural to work with complex scalars. For now we write $\mathbb{F}$ for the scalar field, which may be taken as either $\mathbb{R}$ or $\mathbb{C}$; of course all complex conjugates may be ignored in the case $\mathbb{F} = \mathbb{R}$.

It is useful to introduce an ambient, possibly noncompact metric space (X, d) containing a compact space A whose magnitude we wish to consider, particularly in order to apply Fourier analysis when A is a subset of Euclidean space. Denote by $FM(X)$ the space of finitely supported, finite signed (if $\mathbb{F} = \mathbb{R}$) or complex (if $\mathbb{F} = \mathbb{C}$) measures on X. Define a symmetric bilinear (if $\mathbb{F} = \mathbb{R}$) or Hermitian sesquilinear (if $\mathbb{F} = \mathbb{C}$) form $\langle \cdot, \cdot \rangle_W$ on $FM = FM(X)$ by

$$\langle \mu, \nu \rangle_W := \int \int e^{-d(a,b)} \; d\mu(a) \; d\overline{\nu}(b),$$

where $\overline{\nu}$ denotes the complex conjugate of a complex measure ν. The metric space (X, d) is said to be **positive definite** if $\langle \cdot, \cdot \rangle_W$ is a positive definite inner product. Equivalently, X is positive definite if for each nonempty finite subset $A \subseteq X$, the associated symmetric matrix ζ from Definition 2.1 is positive definite. The phrase "positive definite metric space" will be abbreviated as **PDMS**. In particular, $\mathbb{R}^n$ is positive definite [9, Theorem 2.5.3].

In order to talk sensibly of the magnitude of tA for each $t > 0$, we will sometimes need the additional assumption that tA is a PDMS for each t. By [13, Theorem 3.3], this property of A is equivalent to the classical property of **negative type** (whose original definition will not be needed here). Of course if $A \subseteq \mathbb{R}^n$, then tA is also (isometric to) a subset of $\mathbb{R}^n$, and thus A is of negative type. Many other spaces of negative type which are of interest are collected in [13, Theorem 3.6].

For the rest of this paper (X, d) is assumed to be a PDMS. Let $\mathcal{W}$ denote the completion of FM with respect to the inner product $\langle \cdot, \cdot \rangle_{\mathcal{W}}$; we call $\mathcal{W}$ the **weighting space** of X. For a compact subset $A \subseteq X$, denote by $\mathcal{W}_A$ the closure in $\mathcal{W}$ of $FM(A)$. Note that $\mathcal{W}_A$ is simply the weighting space of A, and is independent of the ambient space X. (Another characterization of $\mathcal{W}$, which may feel more concrete to some readers, will be given in Section 4 below.)

Denote by $C^{1/2}$ the space of Hölder continuous functions $f : X \to \mathbb{F}$ with exponent $1/2$, equipped with the norm

$$\|f\|_{C^{1/2}} := \max \left\{ \|f\|_\infty , \sup_{x,y \in X,\ x \neq y} \frac{|f(x) - f(y)|}{\sqrt{2d(x,y)}} \right\}.$$

We adopt this slightly unusual version of the $C^{1/2}$ norm purely for convenience; any equivalent norm would work equally well for our purposes.

For $\mu \in FM$, define $Z\mu : X \to \mathbb{F}$ by

$$Z\mu(x) := \int e^{-d(x,y)} \, d\mu(y), \tag{3.1}$$

so that

$$\langle \mu, \nu \rangle_{\mathcal{W}} = \int (Z\mu) \, d\overline{\nu} = \langle Z\mu, \overline{\nu} \rangle \tag{3.2}$$

for each $\mu, \nu \in FM$, where $\langle \cdot, \cdot \rangle$ denotes the standard bilinear pairing between functions and measures.

Lemma 3.1. *If $\mu \in FM$, then $Z\mu \in C^{1/2}$ and $\|Z\mu\|_{C^{1/2}} \leq \|\mu\|_{\mathcal{W}}$.*

Proof. By the Cauchy–Schwarz inequality, for each $x, y \in X$,

$$|Z\mu(x)| = |\langle \mu, \delta_x \rangle_{\mathcal{W}}| \leq \|\mu\|_{\mathcal{W}} \|\delta_x\|_{\mathcal{W}} = \|\mu\|_{\mathcal{W}}$$

and

$$|Z\mu(x) - Z\mu(y)| = \left| \langle \mu, \delta_x - \delta_y \rangle_{\mathcal{W}} \right| \leq \|\mu\|_{\mathcal{W}} \|\delta_x - \delta_y\|_{\mathcal{W}} .$$

Now

$$\begin{aligned} \|\delta_x - \delta_y\|_{\mathcal{W}}^2 &= \|\delta_x\|_{\mathcal{W}}^2 - \langle \delta_x, \delta_y \rangle_{\mathcal{W}} - \langle \delta_y, \delta_x \rangle_{\mathcal{W}} + \|\delta_y\|^2 \\ &= 2\big(1 - e^{-d(x,y)}\big) \leq 2d(x,y), \end{aligned}$$

so $\|Z\mu\|_{C^{1/2}} \leq \|\mu\|_{\mathcal{W}}$. ∎

Proposition 3.2. *The map $Z : FM \to C^{1/2}$ defined by* (3.1) *extends uniquely to an injective linear operator $Z : \mathcal{W} \to C^{1/2}$ with $\|Z\| = 1$. Furthermore, for each $w \in \mathcal{W}$ and $\mu \in FM$,*

$$\langle w, \mu \rangle_{\mathcal{W}} = \int (Zw) \, d\overline{\mu}. \tag{3.3}$$

Proof. Lemma 3.1 imply that Z extends uniquely to a linear operator $Z : \mathcal{W} \to C^{1/2}$ with norm at most 1, and (3.2) then implies (3.3). For each $x \in X$, $\|Z\delta_x\|_\infty = 1 = \|\delta_x\|_\mathcal{W}$, so $\|Z\| = 1$.

To prove injectivity, suppose that $Zw = 0$ for some $w \in \mathcal{W}$. Then for each $x \in X$, by (3.3),

$$0 = Zw(x) = \langle w, \delta_x \rangle_\mathcal{W} ,$$

from which it follows by linearity that $\langle w, \mu \rangle_\mathcal{W} = 0$ for each $\mu \in FM$. Since FM is dense in $\mathcal{W}$, this implies that $w = 0$. □

Definition 3.3. *Let $A \subseteq X$ be compact. A* **weighting** *for A is an element $w \in \mathcal{W}_A$ such that for each $a \in A$, $Zw(a) = 1$. If A possesses a weighting w, then the* **magnitude** *of A is*

$$|A| := \|w\|_\mathcal{W}^2 . \tag{3.4}$$

If A does not possess a weighting, then $|A| := \infty$.

Since Z is injective, if A possesses a weighting, then the weighting is unique. If A possesses a complex weighting w, then the real part of w (defined by extending the "real part" map on FM to $\mathcal{W}$) is also a weighting for A, and hence equal to w; thus the existence of a weighting for A and the magnitude of A are independent of the scalar field. It is an open question whether there exists a compact PDMS whose magnitude is infinite.

If A is a compact metric space of negative type, then **magnitude function** of A is defined as before to be the function $t \mapsto |tA|$ for $t > 0$.

We next compare Definition 3.3 to the original Definition 2.1 of magnitude for a finite metric space A. The definitions of weighting are clearly equivalent when $w \in \mathbb{R}^A$ is identified with $\sum_{a\in A} w(a)\delta_a \in FM(A)$. As noted earlier, an arbitrary finite metric space may not possess a weighting, but if A is a finite PDMS then the matrix ζ is positive definite, hence invertible, so A possesses a unique weighting $w \in \mathbb{R}^A$.

The equivalence of (2.2) and (3.4) is less immediately obvious, so it is desirable to motivate(3.4). If one thinks of measures as dual to functions, then the right hand side of (2.2) is interpreted not as $w(A)$ but as $\langle 1, \overline{w} \rangle$, where 1 denotes the function with the constant value 1. However, interesting function spaces on noncompact domains generally do not contain constant functions, which suggests that this interpretation is also unsatisfactory; one should replace the constant 1 with some canonical function which is equal to 1 everywhere on A. Fortunately, (2.1) provides such a function, namely ζw. The definition (2.2) can thus be rewritten as

$$|A| := \sum_{a,b\in A} w(a)\zeta(a,b)\overline{w(b)}.$$

and then reinterpreted as

$$|A| := \int (\zeta w)\ d\overline{w} = \langle \zeta w, \overline{w} \rangle ,$$

which is clearly equivalent to (3.4).

The next result shows that Definition 3.3 agrees with the definition of magnitude adopted in [13] (cf. Theorem 2.4 in [13]), and motivates the definition of $|A| = \infty$ when A does not possess a weighting.

Theorem 3.4. *If $A \subseteq X$ is compact then*

$$|A| = \sup \left\{ \frac{|\mu(A)|^2}{\|\mu\|_{\mathcal{W}}^2} \;\middle|\; \mu \in FM(A),\ \mu \neq 0 \right\}. \tag{3.5}$$

Proof. Let κ denote the supremum in (3.5). We need to show that A possesses a weighting w if and only if $\kappa < \infty$, and that in that case $\|w\|_{\mathcal{W}}^2 = \kappa$.

Suppose that A possesses a weighting w. Then for each $\mu \in FM(A)$, by (3.3),

$$\langle w, \mu \rangle_{\mathcal{W}} = \langle Zw, \overline{\mu} \rangle = \overline{\mu(A)},$$

and so by the Cauchy–Schwarz inequality,

$$|\mu(A)|^2 = |\langle w, \mu \rangle_{\mathcal{W}}|^2 \leq \|\mu\|_{\mathcal{W}}^2 \|w\|_{\mathcal{W}}^2,$$

which implies that $\kappa < \infty$.

Now suppose that $\kappa < \infty$. Then the linear functional $\mu \mapsto \mu(A)$ on $\big(FM(A), \|\cdot\|_{\mathcal{W}}\big)$ is bounded with norm $\sqrt{\kappa}$. It thus extends to a linear functional on $\mathcal{W}_A$ with norm $\sqrt{\kappa}$. By the Riesz representation theorem for bounded linear functionals on Hilbert spaces, there exists a $w \in \mathcal{W}_A$ such that $\langle \mu, w \rangle_{\mathcal{W}} = \mu(A)$ for each $\mu \in FM(A)$ and $\|w\|_{\mathcal{W}}^2 = \kappa$. In particular, for each $a \in A$, by (3.3),

$$Zw(a) = \langle w, \delta_a \rangle_{\mathcal{W}} = \overline{\delta_a(A)} = 1.$$

Thus w is a weighting for A, and $|A| = \|w\|_{\mathcal{W}}^2 = \kappa$. □

Besides the agreement of Definition 3.3 with the original definition of magnitude for finite PDMSs, the results of [13, Section 2] support Definition 3.3 as the "correct" notion of magnitude for a compact PDMS. As shown in [13, Theorem 2.6], magnitude as defined here or in [13] is lower semicontinuous on the class of compact PDMSs equipped with the Gromov–Hausdorff topology. In fact, since Theorem 3.4 implies that

$$|A| = \sup \{ |B| \mid B \subseteq A \text{ is finite} \},$$

magnitude as defined here is the minimal lower semicontinuous extension of magnitude from the class of finite PDMSs to the class of compact PDMSs.

4. The reproducing kernel and potential function of a PDMS

For technical reasons, it is fruitful to shift the emphasis from the weighting w of a compact subset $A \subseteq X$ to the function $Zw : X \to \mathbb{F}$, and thus from the weighting space $\mathcal{W}$ to the function space $Z(\mathcal{W}) \subseteq C^{1/2}$, which we consider next. The results of the next two sections should convince the skeptical reader that this additional layer of complexity is worthwhile.

Define $\mathcal{H} := Z(\mathcal{W}) \subseteq C^{1/2}$, and for a compact set $A \subseteq X$, define $\mathcal{H}_A := Z(\mathcal{W}_A)$. Equip $\mathcal{H}$ with the inner product

$$\langle g, h \rangle_{\mathcal{H}} := \left\langle Z^{-1} g, Z^{-1} h \right\rangle_{\mathcal{W}},$$

recalling that Z is injective by Proposition 3.2. Then $Z : \mathcal{W} \to \mathcal{H}$ is a surjective isometry, and $\|h\|_{C^{1/2}} \leq \|h\|_{\mathcal{H}}$ for each $h \in \mathcal{H}$. Furthermore, the Hilbert spaces $\mathcal{H}$ and $\mathcal{W}$ act as duals to each other via the bilinear pairing

$$\langle h, w \rangle = \left\langle Z^{-1} h, \overline{w} \right\rangle_{\mathcal{W}} = \left\langle h, \overline{Zw} \right\rangle_{\mathcal{H}}. \tag{4.1}$$

By (3.3), (4.1) extends the standard pairing between functions and measures. In particular, for each $x \in X$,

$$h(x) = \langle h, \delta_x \rangle = \left\langle h, e^{-d(x,\cdot)} \right\rangle_{\mathcal{H}}.$$

That is, $\mathcal{H}$ is the reproducing kernel Hilbert space (RKHS) on X with the reproducing kernel $e^{-d(x,y)}$. (Readers unfamiliar with RKHSs are referred to [2].) We have chosen to define first the weighting space $\mathcal{W}$, and then define $\mathcal{H}$ in terms of it, since this more closely parallels Leinster's original definition of the magnitude of a finite metric space. Alternatively, one could first define $\mathcal{H}$ to be this RKHS, and then define $\mathcal{W}$ as the dual space of $\mathcal{H}$. We will not make use of the theory of RKHSs in this paper, opting instead for more self-contained arguments.

If a compact set $A \subseteq X$ possesses a weighting w, then the function $h = Zw \in C^{1/2}$ is called the **potential function** of A. This name is motivated by the following physical model, which amounts to a less picturesque version of the penguin analogy discussed in [19], and which is not intended here as a realistic model of any physical phenomenon. (Actually, it is related to the Yukawa potential in one dimension, but not in any higher dimension; see [12, Section 6.23] for background from a mathematical perspective.)

We posit a type of charge such that a unit charge at $x \in X$ creates a potential at $y \in X$ of $e^{-d(x,y)}$. A finite set of fixed points $A \subseteq X$ represents a conductor which is connected to a ground to and from which charge may flow freely, but which otherwise does not interact with anything in the space X. The conductor is held at a uniform potential of 1. Suppose further that there is no other charge anywhere in the ambient space X. Then the charge at each point of A is given by the weighting w of A, and

$$h(x) = \sum_{a \in A} e^{-d(a,x)} w(a)$$

is indeed the potential at each point $x \in X$. (The magnitude $|A|$ is, in this model, the total charge on A.) If $A \subseteq X$ is an infinite compact set, its weighting w represents a charge distribution, although we stress that w need not be given by a function or even a measure on A, but will in general be some more singular type of object. When X is Euclidean space, this charge distribution in fact turns out to be a distribution in the sense of Schwartz, as will be seen in the next section.

We can now give a finiteness condition and a variational formula for magnitude, dual to those in Theorem 3.4, in terms of the RKHS $\mathcal{H}$.

Theorem 4.1. *Let $A \subseteq X$ be compact. Then A possesses a weighting if and only if there exists a function $h \in \mathcal{H}$ such that $h \equiv 1$ on A. In that case,*

$$|A| = \inf \left\{ \|h\|_{\mathcal{H}}^2 \;\middle|\; h \in \mathcal{H} \text{ and } h \equiv 1 \text{ on } A \right\}, \tag{4.2}$$

and the infimum is uniquely attained by the potential function of A.

Proof. If A possesses a weighting w, then by definition the potential function $h = Zw$ of A lies in the set appearing in the right hand side of (4.2). Moreover, $|A| = \|w\|_{\mathcal{W}}^2 = \|h\|_{\mathcal{H}}^2$.

On the other hand, if $h \in \mathcal{H}$ and $h \equiv 1$ on A, then for each $\mu \in FM(A)$,

$$|\mu(A)|^2 = |\langle h, \mu \rangle|^2 \leq \|h\|_{\mathcal{H}}^2 \, \|\mu\|_{\mathcal{W}}^2 .$$

By Theorem 3.4, A possesses a weighting and $|A| \leq \|h\|_{\mathcal{H}}^2$. It furthermore follows that if A possesses a weighting then (4.2) holds.

Moreover, if the closed affine subspace

$$\{h \in \mathcal{H} \mid h \equiv 1 \text{ on } A\} = \bigcap_{a \in A} \{h \mid \langle h, \delta_a \rangle = 1\}$$

is nonempty then it contains a unique element of minimal norm, which must therefore be the potential function Zw of A. □

As observed already in Section 3, it is an open question whether an arbitrary compact PDMS A has finite magnitude. By Theorem 4.1, this is equivalent to asking whether the RKHS $\mathcal{H}_A$ contains the constant functions on A.

The next two results may be useful in computing magnitudes, as will be seen in Section 5.

Proposition 4.2. *Let $A \subseteq X$ be compact with weighting w. Then $|A| = \langle h, w\rangle$ for any $h \in \mathcal{H}$ such that $h \equiv 1$ on A.*

Proof. For any such h and any $\mu \in FM(A)$,

$$\langle h, \mu\rangle = \mu(A) = \left\langle \overline{Zw}, \mu\right\rangle = \langle \mu, w\rangle_{\mathcal{W}}.$$

Let μ_n be a sequence in $FM(A)$ which converges to w with respect to $\|\cdot\|_{\mathcal{W}}$. Then

$$\langle h, w\rangle = \lim_{n\to\infty} \langle h, \mu_n\rangle = \lim_{n\to\infty} \langle \mu_n, w\rangle_{\mathcal{W}} = \langle w, w\rangle_{\mathcal{W}} = |A|.$$

Alternatively, the proposition is equivalent to the claim that $|A| = \langle h, Zw\rangle_{\mathcal{H}}$ for any $h \in \mathcal{H}$ such that $h \equiv 1$ on A, where Zw is the potential function of A. In this form the statement follows from Theorem 4.1 and basic Hilbert space geometry. □

Corollary 4.3. *Let $A \subseteq X$ be compact with potential function h. Then $\langle h, g\rangle_{\mathcal{H}} = 0$ for any $g \in \mathcal{H}$ such that $g \equiv 0$ on A.*

Proof. For any such g, $g + h \equiv 1$ on A, and so by Proposition 4.2,

$$\langle h, g\rangle_{\mathcal{H}} = \langle h, g + h\rangle_{\mathcal{H}} - \langle h, h\rangle_{\mathcal{H}} = |A| - |A| = 0. \qquad \square$$

We consider next the role of measures which are not finitely supported. Let $M^{1/2}$ be the space of finite signed or complex Borel measures $\mu \in M(X)$ such that $\int \sqrt{d(x,y)}\, d|\mu|(y) < \infty$ for some (hence for any) $x \in X$. It is easy to verify that $h \mapsto \int h\, d\mu$ defines a bounded linear functional on $C^{1/2}$, and therefore on $\mathcal{H}$. Thus measures in $M^{1/2}$ define elements of $\mathcal{W}$, and in particular any $\mu \in M(A)$ defines an element of $\mathcal{W}_A$; the map Z acts on $\mu \in M^{1/2}$ according to (3.1). It is not clear, however, whether distinct measures in $M^{1/2}$ necessarily give rise to distinct elements of $\mathcal{W}$. Two equivalent formulations of this issue are that the form $\langle \cdot, \cdot\rangle_{\mathcal{W}}$ may be degenerate on $M^{1/2}$, and that Z may not be injective on $M^{1/2}$. This is why only finitely supported measures were used in defining $\mathcal{W}$ in the previous section. However, this does imply that in (3.5), one can extend the supremum to all $\mu \in M(A)$ such that $\|\mu\|_{\mathcal{W}} \neq 0$; this leads to another proof of the first part of [13, Theorem 2.4].

The above observation also lets us reproduce [13, Theorem 2.3], which shows that if A possesses a **weight measure** $\mu \in M(A)$—that is, $Z\mu(a) = \int e^{-d(a,b)}\, d\mu(b) = 1$ for each $a \in A$—then

$$|A| = \int (Z\mu)\, d\mu = \mu(A);$$

in other words, the obvious analogue of the formula (2.2) for the magnitude of a finite set does hold in this case. In [20] weight measures are used to define the magnitude of metric spaces which are not necessarily positive definite (this will be discussed in more detail in Section 7 below); by [13, Theorem 2.3] this coincides with the present definition for any PDMS which possesses a weight measure. Although it is not yet clear how to define weightings, and hence magnitude, for arbitrary metric spaces, it is clear that using weight measures is insufficient since in general the weighting of a PDMS is not given by a measure. Nevertheless, it appears to be reasonable to use this definition whenever a weight measure does exist.

We end this section by considering a quantity related to magnitude which is in some ways better behaved. For a compact (not necessarily positive definite) metric space A, the **maximum diversity** of A is

$$|A|_+ = \sup_{\mu \in P(A)} \left(\int \int e^{-d(a,b)} \; d\mu(a) \; d\mu(b) \right)^{-1}, \tag{4.3}$$

where $P(A)$ denotes the space of Borel probability measures on A. By renormalization, this is simply what one obtains by restricting the supremum in (3.5) to positive measures; thus we trivially have

$$|A|_+ \leq |A| \tag{4.4}$$

for any compact PDMS A. The name stems from the following interpretation of the quantity inside the supremum. Suppose that the points of a finite metric space A represent all the species present in an ecosystem, and $e^{-d(a,b)} \in (0,1]$ is viewed as the similarity between species a and b. If $\mu \in P(A)$ describes the relative abundances of the species, then

$$\int \int e^{-d(a,b)} \; d\mu(a) \; d\mu(b) \tag{4.5}$$

is the expected similarity of a pair of independently picked random organisms. The reciprocal of (4.5) quantifies, in a similarity-sensitive way, the diversity of the ecosystem; $|A|_+$ is thus the maximum possible diversity of the given collection of species when one considers all possible relative abundances.

We remark that the reciprocal of (4.5) is just one of an infinite family of diversity measures introduced in [7, 10]. It is shown in [7] that, under certain conditions, they all have the same maximum value. See [10] for a thorough discussion of these diversity measures in the context of theoretical ecology, and [7] for a proof of a subtler relationship between maximum diversity and magnitude.

Although maximum diversity lacks the category-theoretic motivation of magnitude, it is in some ways more tractable than magnitude due to the fact that it can be represented in terms of *positive* measures. One easy property which follows immediately from (4.3) is that

$$|A|_+ \leq \exp(\operatorname{diam} A),$$

where $\operatorname{diam} A$ denotes the diameter of A. A subtler property is that $|A|_+$ is a continuous function of A with respect to the Gromov–Hausdorff topology (see [13, Proposition 2.11]; this result is stated for PDMSs but the proof applies to arbitrary compact metric spaces). Both of these properties fail in general for magnitude, as witnessed by Example 2.2.8 in [9], due to Willerton, of a PDMS A with six points such that $\lim_{t\to 0^+} |tA| = 6/5$. Nevertheless, we will see in Section 6 below that magnitude and maximum diversity have a deep relationship in Euclidean space. This will play a crucial role in our investigation of the asymptotic growth of the magnitude function in Section 7.

The maximum diversity of a compact PDMS can also be given a dual characterization analogous to Theorem 4.1.

Proposition 4.4. *Let X be a PDMS and let $A \subseteq X$ be compact. Then*

$$|A|_+ = \inf \left\{ \|h\|_{\mathcal{H}}^2 \;\middle|\; h \in \mathcal{H} \text{ and } h \geq 1 \text{ on } A \right\}. \tag{4.6}$$

Proof. For simplicity of exposition we assume in this proof that $\mathbb{F} = \mathbb{R}$; the complex case then follows since $\|\operatorname{Re} h\|_{\mathcal{H}} \le \|h\|_{\mathcal{H}}$.

By (4.3) and the duality between $\mathcal{W}$ and $\mathcal{H}$,

$$|A|_+^{-1/2} = \inf_{\mu \in P(A)} \|\mu\|_{\mathcal{W}} = \inf_{\mu \in P(A)} \sup_{\substack{g \in \mathcal{H} \\ \|g\|_{\mathcal{H}} \le 1}} \int g \, d\mu.$$

By a general minimax theorem (see e.g. [1, Theorem 2.4.1]), this is equal to

$$\sup_{\substack{g \in \mathcal{H} \\ \|g\|_{\mathcal{H}} \le 1}} \inf_{\mu \in P(A)} \int g \, d\mu = \sup_{\substack{g \in \mathcal{H} \\ \|g\|_{\mathcal{H}} \le 1}} \inf_{a \in A} g(a).$$

Since $\mathcal{H}$ contains strictly positive functions, the supremum is unchanged if g is assumed to be positive on A. Given a $g \in \mathcal{H}$ with $\inf_{a \in A} g(a) = c > 0$, define $h = \frac{1}{c} g$. The mapping $g \mapsto h$ defines a bijection

$$\{g \in \mathcal{H} \mid \|g\|_{\mathcal{H}} \le 1 \text{ and } g > 0 \text{ on } A\} \to \{h \in \mathcal{H} \mid h \ge 1 \text{ on } A\},$$

such that $\inf_{a \in A} g(a) = c = \|h\|_{\mathcal{H}}^{-1}$. It follows that

$$|A|_+^{-1/2} = \sup \left\{ \|h\|_{\mathcal{H}}^{-1} \;\middle|\; h \in \mathcal{H} \text{ and } h \ge 1 \text{ on } A \right\}. \qquad \square$$

5. Magnitude in Euclidean space

We now specialize to the case in which $X = \mathbb{R}^n$, equipped with the standard inner product $\langle \cdot, \cdot \rangle$ and the associated norm $\|\cdot\|$ and metric d. Some of the results of this section generalize to certain other normed or quasinormed spaces (cf. Sections 3 and 4 of [13]), but here we restrict to the Euclidean case, which is of most central interest and about which the most can be said.

The first task is to observe that the weighting space $\mathcal{W}$ and the RKHS $\mathcal{H}$ turn out to be well-known Sobolev-type spaces of distributions and functions, respectively. Define $F : \mathbb{R}^n \to \mathbb{R}$ by $F(x) := e^{-\|x\|}$. Then the Fourier transform of F is

$$\widehat{F}(x) = \frac{1}{(2\pi)^{n/2}} \int F(y) e^{-i\langle x, y\rangle} \, dy = \frac{n! \omega_n}{(2\pi)^{n/2}} \left(1 + \|x\|^2\right)^{-(n+1)/2}, \tag{5.1}$$

where $\omega_n = \pi^{n/2} / \Gamma\left(\frac{n}{2} + 1\right)$ is the volume of the unit ball in $\mathbb{R}^n$ [17, Theorem 1.14]. In this setting the map $Z : FM \to C^{1/2}$ is the convolution operator $\mu \mapsto F * \mu$, and therefore

$$\widehat{Z\mu}(x) = (2\pi)^{n/2} \widehat{F}(x) \widehat{\mu}(x) = n! \omega_n \left(1 + \|x\|^2\right)^{-(n+1)/2} \widehat{\mu}(x). \tag{5.2}$$

For $\alpha \in \mathbb{R}$, the **Bessel potential space** $H^\alpha = H^\alpha(\mathbb{R}^n)$ is the Hilbert space of tempered distributions

$$H^\alpha := \left\{ \varphi \in \mathcal{S}'(\mathbb{R}^n) \;\middle|\; \left(1 + \|\cdot\|^2\right)^{\alpha/2} \widehat{\varphi} \in L^2(\mathbb{R}^n) \right\}$$

equipped with the norm

$$\|\varphi\|_{H^\alpha} := \sqrt{\int_{\mathbb{R}^n} \left(1 + \|x\|^2\right)^\alpha \left|\widehat{\varphi(x)}\right|^2 \, dx}.$$

(See e.g. [4, Section 7.9]; our normalization for the Fourier transform—identified in (5.1)—differs from that of [4], but the H^α norms agree.) When $\alpha \ge 0$, $H^\alpha \subseteq L^2$, and so H^α is actually a space of functions. For each $\alpha > 0$, the space $\mathcal{S}$ of Schwartz functions is dense in H^α, and H^α and $H^{-\alpha}$ act as duals via the unique extension of the bilinear pairing

$$\langle f, \varphi \rangle = \varphi(f)$$

for $f \in \mathcal{S} \subseteq H^\alpha$ and $\varphi \in H^{-\alpha} \subseteq \mathcal{S}'$.

The following results show that when $X = \mathbb{R}^n$, the weighting space $\mathcal{W}$ may be identified with the space of distributions $H^{-(n+1)/2}$ and the RKHS $\mathcal{H}$ is the function space $H^{(n+1)/2}$. Thus, the potential function of a compact set $A \subseteq \mathbb{R}^n$ is in fact a so-called Bessel potential.

Proposition 5.1. *The inclusion map $FM \hookrightarrow H^{-(n+1)/2}$ extends to a bijection $\mathcal{W} \to H^{-(n+1)/2}$ such that $\|\mu\|_{\mathcal{W}} = \sqrt{n!\omega_n}\, \|\mu\|_{H^{-(n+1)/2}}$ for each $\mu \in FM$.*

Proof. For each $x \in \mathbb{R}^n$, $\left|\widehat{\delta_x}\right| \equiv (2\pi)^{-n/2}$. Since

$$\int_{\mathbb{R}^n} \left(1 + \|x\|^2\right)^{-(n+1)/2} \, dx < \infty,$$

as can be seen by integrating in polar coordinates, it follows that $\delta_x \in H^{-(n+1)/2}$, and therefore that $FM \subseteq H^{-(n+1)/2}$.

Now let $\mu = \sum_{j=1}^n c_j \delta_{x_j} \in FM$. By the Fourier inversion theorem,

$$\begin{aligned} n!\omega_n \|\mu\|_{H^{-(n+1)/2}}^2 &= (2\pi)^{n/2} \int_{\mathbb{R}^n} \widehat{F}\, |\widehat{\mu}|^2 \\ &= (2\pi)^{n/2} \int_{\mathbb{R}^n} \widehat{F}(y) \frac{1}{(2\pi)^n} \sum_{j,k=1}^n c_j \overline{c_k} e^{-i\langle y, x_j - x_k \rangle} \, dy \\ &= \sum_{j,k=1}^n c_j \overline{c_k} \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \widehat{F}(y) e^{i\langle y, x_k - x_j \rangle} \, dy \\ &= \sum_{j,k=1}^n c_j \overline{c_k} F(x_k - x_j) = \|\mu\|_{\mathcal{W}}^2 . \end{aligned}$$

Thus the inclusion $FM \hookrightarrow H^{-(n+1)/2}$ extends to a map $\mathcal{W} \to H^{-(n+1)/2}$ with the stated identity between norms, and which is therefore injective.

To show that this map is surjective, we need to show that FM is dense in $H^{-(n+1)/2}$. Equivalently, we need to show that any bounded linear functional on $H^{-(n+1)/2}$ which vanishes on FM is zero. Each bounded linear functional on $H^{-(n+1)/2}$ is represented by some $f \in H^{(n+1)/2}$. The evaluation of that linear functional on the point mass δ_x is $f(x)$; thus the vanishing of f on FM implies that $f = 0$. □

Corollary 5.2. *The map $Z : FM \to C^{1/2}$ extends to a bijection $Z : \mathcal{W} \to H^{(n+1)/2}$ such that $\|Z\mu\|_{H^{(n+1)/2}} = \sqrt{n!\omega_n}\, \|\mu\|_{\mathcal{W}}$ for each $\mu \in FM$. Thus $\mathcal{H} = H^{(n+1)/2}$ and $\|h\|_{H^{(n+1)/2}} = \sqrt{n!\omega_n}\, \|h\|_{\mathcal{H}}$ for each $h \in H^{(n+1)/2}$.*

Proof. This follows from Proposition 5.1, formula (5.2), and the fact that

$$\varphi \mapsto \mathcal{F}^{-1}\Big(\left(1 + \|\cdot\|^2\right)^{-(n+1)/2} \widehat{\varphi}\Big)$$

is an isometry of $H^{-(n+1)/2}$ onto $H^{(n+1)/2}$, where $\mathcal{F}$ denotes the Fourier transform on $\mathcal{S}'$, which follows immediately from the definition of the spaces H^α. □

Alternatively, in light of the comments in the previous section, Corollary 5.2 amounts to the known (but not easy to find explicitly stated) fact that $H^{(n+1)/2}$ is an RKHS with reproducing kernel $\frac{1}{n!\omega_n}e^{-\|x-y\|}$. (Note that H^α is only an RKHS when $\alpha > n/2$, and that there is no simple explicit formula for the reproducing kernel for most values of α.) Proposition 5.1 then follows from the duality between $H^{(n+1)/2}$ and $H^{-(n+1)/2}$.

If follows from Corollary 5.2 that the general fact that $\|h\|_{C^{1/2}} \le \|h\|_{\mathcal{H}}$ for $h \in \mathcal{H}$ is, in the Euclidean setting, a special case of the Sobolev embedding theorem.

Corollary 5.3. *Let $A \subseteq \mathbb{R}^n$ be compact. Then A possesses a weighting in $H^{-(n+1)/2}$, and*

$$|A| = \frac{1}{n!\omega_n} \inf\left\{\|h\|^2_{H^{(n+1)/2}} \mid h \in H^{(n+1)/2} \text{ and } h \equiv 1 \text{ on } A\right\}.$$

Proof. Any Schwartz function h with $h \equiv 1$ on A lies in $H^{(n+1)/2}$. Theorem 4.1 and Corollary 5.2 then imply the stated formula for $|A|$ and the existence of a weighting for A, which by Proposition 5.1 may be regarded as an element of $H^{-(n+1)/2}$. □

Corollary 5.3 contains the fact that each compact subset of $\mathbb{R}^n$ has finite magnitude, first proved in [9, Proposition 3.5.3]. However, unlike the upper bound on magnitude in [9, Lemma 3.5.2], from which Proposition [9, Proposition 3.5.3] was deduced, the infimum expression in Corollary 5.3 is sharp, and implies some new, sharper bounds. In the next result, the upper bound on $|tA|$ improves [9, Lemma 3.5.4], which implies a similar upper bound with an additional constant factor (depending on A). The lower bound is new.

Theorem 5.4. *Let $A \subseteq \mathbb{R}^n$ be compact and $t \ge 1$. Then*

$$\frac{|A|}{t} \le |tA| \le t^n |A|.$$

Proof. Let $h \in H^{(n+1)/2}$ be the potential function for A, and let $h_t(x) = h(x/t)$. By Corollary 5.3,

$$\begin{aligned} n!\omega_n |tA| &\le \int_{\mathbb{R}^n} \left(1 + \|x\|^2\right)^{(n+1)/2} \left|\widehat{h_t}(x)\right|^2 \, dx \\ &= \int_{\mathbb{R}^n} \left(1 + \|x\|^2\right)^{(n+1)/2} \left|t^n \widehat{h}(tx)\right|^2 \, dx \\ &= t^n \int_{\mathbb{R}^n} \left(1 + t^{-2}\|y\|^2\right)^{(n+1)/2} \left|\widehat{h}(y)\right|^2 \, dy \\ &\le t^n \int_{\mathbb{R}^n} \left(1 + \|y\|^2\right)^{(n+1)/2} \left|\widehat{h}(y)\right|^2 \, dy \\ &= t^n n!\omega_n |A|. \end{aligned}$$

Now let g be the potential function for tA, and let $g^t(x) = g(tx)$. By Corollary 5.3,

$$\begin{aligned} n!\omega_n |A| &\le \int_{\mathbb{R}^n} \left(1 + \|x\|^2\right)^{(n+1)/2} \left|\widehat{g^t}(x)\right|^2 \, dx \\ &= \int_{\mathbb{R}^n} \left(1 + \|x\|^2\right)^{(n+1)/2} \left|t^{-n}\widehat{g}(x/t)\right|^2 \, dx \\ &= t^{-n} \int_{\mathbb{R}^n} \left(1 + t^2\|y\|^2\right)^{(n+1)/2} \left|\widehat{g}(y)\right|^2 \, dy \\ &= t \int_{\mathbb{R}^n} \left(t^{-2} + \|y\|^2\right)^{(n+1)/2} \left|\widehat{g}(y)\right|^2 \, dy \end{aligned}$$

$$\leq t \int_{\mathbb{R}^n} \left(1 + \|y\|^2\right)^{(n+1)/2} |\widehat{g}(y)|^2 \; dy$$
$$= tn!\omega_n |tA| . \qquad \square$$

It is an open question whether the magnitude function of a compact subset of $\mathbb{R}^n$ (or more generally, a compact PDMS) must be nondecreasing. The lower bound in Theorem 5.4 is one partial result in that direction; for another see the discussion following Corollary 6.2 below.

Corollary 5.5. *If $A \subseteq \mathbb{R}^n$ is compact, then the magnitude function of A is continuous on $(0, \infty)$.*

Proof. It suffices by rescaling to prove continuity of $|tA|$ at $t = 1$. Theorem 5.4 immediately implies that $\lim_{t\to 1^+} |tA| = |A|$, and upon replacing t with $1/t$ it also implies that $\lim_{t\to 1^-} |tA| = |A|$. □

As mentioned in Section 3, it was proved in [13] that magnitude is lower semicontinuous on the class of compact PDMSs; this implies that the magnitude function of a compact space of negative type is lower semicontinuous. Proposition 2.2.6 of [9] implies that the magnitude function of a finite space of negative type is even analytic. It is unknown whether the magnitude function of a compact space of negative type is continuous for $t > 0$ in general, although [9, Example 2.2.8], mentioned above in Section 4, shows that the magnitude function can fail to be continuous at 0 if we define $0A$ to be a one-point metric space.

Corollary 5.3 can also be used to give an easy proof of the following special case of [9, Theorem 3.5.6]. (The methods of the present paper can also be used to prove [9, Theorem 3.5.6], which treats an arbitrary positive definite finite dimensional normed space, in full generality.)

Proposition 5.6. *If $A \subseteq \mathbb{R}^n$ is compact, then $|A| \geq \frac{\mathrm{vol}(A)}{n!\omega_n}$.*

Proof. Suppose $h \in H^{(n+1)/2}$ and $h \equiv 1$ on A. Then by Parseval's identity,

$$\|h\|^2_{H^{(n+1)/2}} \geq \left\|\widehat{h}\right\|^2_{L^2} = \|h\|^2_{L^2} \geq \mathrm{vol}(A),$$

and the result follows from Corollary 5.3. □

We next show that the potential function of a compact subset of Euclidean space satisfies a certain pseudodifferential equation, which reduces to a partial differential equation in odd dimensions. This equation amounts to the Euler–Lagrange equation for the variational problem described by Corollary 5.3.

Proposition 5.7. *Let $A \subseteq \mathbb{R}^n$ be compact with potential function h. Then $(I-\Delta)^{(n+1)/2}h = 0$ in the distributional sense on $\mathbb{R}^n \setminus A$.*

Proof. Let $g : \mathbb{R}^n \to \mathbb{R}$ be a smooth function with compact support contained in $\mathbb{R}^n \setminus A$. By Corollaries 4.3 and 5.2,

$$\left\langle (I-\Delta)^{(n+1)/2}h, g\right\rangle = \left\langle \mathcal{F}\big((I-\Delta)^{(n+1)/2}h\big), \overline{\widehat{g}}\right\rangle = \left\langle \left(1 + \|\cdot\|^2\right)^{(n+1)/2}\widehat{h}, \overline{\widehat{g}}\right\rangle$$
$$= \langle h, g\rangle_{H^{(n+1)/2}} = 0. \qquad \square$$

Corollary 5.8. *Let $A \subseteq \mathbb{R}^n$ be compact. Then the potential function for A is C^∞ on $\mathbb{R}^n \setminus A$.*

Proof. The pseudodifferential operator $(I-\Delta)^{(n+1)/2}$ is elliptic, and the corollary follows from classical elliptic regularity theory; see e.g. [18, Corollary 4.5]. For a more elementary treatment in the case that n is odd (so that $(I-\Delta)^{(n+1)/2}$ is actually a differential operator), see the corollary to Theorem 8.12 in [15]. □

Proposition 5.9. *Let $A \subseteq \mathbb{R}^n$ be compact with potential function h. Then the weighting w of A is the distribution $w = \frac{1}{n!\omega_n}(I-\Delta)^{(n+1)/2}h$.*

Proof. This follows from Corollary 5.2, (5.2), and the fact that $h = Zw$. □

As an application of the above results, we compute the potential function, weighting, and magnitude of an interval $A = [0,\ell] \subseteq \mathbb{R}$. Corollary 5.8 implies that the potential function h of A is smooth on $(-\infty, 0)$ and on (ℓ, ∞), and then Proposition 5.7 shows that $h'' = h$ (in the classical sense) on those intervals. By definition, the potential function h is equal to 1 on A, and is continuous by Proposition 3.2. Furthermore, since $h \in H^1(\mathbb{R})$, it decays to 0 at $\pm\infty$ (see e.g. [4, Corollary 7.9.4]). This boundary value problem has the unique solution

$$h(x) = \begin{cases} e^x & \text{if } x < 0, \\ 1 & \text{if } 0 \le x \le \ell, \\ e^{-x+\ell} & \text{if } x > \ell. \end{cases} \tag{5.3}$$

By Proposition 5.9, the weighting of A is the distribution

$$w = \frac{1}{2}(h - h'') = \frac{1}{2}\big(\lambda_A + \delta_0 + \delta_\ell\big),$$

where λ_A denotes Lebesgue measure on the interval A. Since w is a measure, Proposition 4.2 implies that $|A| = w(A) = 1 + \frac{\ell}{2}$.

The magnitude of an interval was first found in [11, Theorem 7] using approximation by finite sets, as justified by the results of [13]; see also [9, Theorem 3.2.2]. The weight measure of an interval was given, and proved to be a weight measure, in [20, Theorem 2], though it was not computed from more basic data as above. We note that Lemma 2.8 and Corollary 2.10 of [13] imply that the weighting of *any* compact subset of $\mathbb{R}$ is a measure, although numerical computations in [19] indicate this is unlikely to be true in higher dimensions.

6. Magnitude, diversity, and capacity

Experts in potential theory will have found several of the definitions and results of Sections 3 and 4 very familiar, and recognized by Section 5 that their specializations to Euclidean space are rather classical. In this section, we make this connection explicit, and note that a deep result about equivalence of capacities in Euclidean space implies an important relationship between magnitude and maximum diversity, which will be vital in our analysis of the growth of magnitude functions in Section 7 below.

Many of the definitions and results of potential theory have complicated histories of successive generalizations. We will rely on the book [1] as a source, referring the reader there for original references.

For $\alpha > 0$, the **Bessel kernel** $G_\alpha : \mathbb{R}^n \to \mathbb{R}$ is defined as the function such that

$$\widehat{G_\alpha}(x) = (2\pi)^{-n/2}\big(1 + \|x\|^2\big)^{-\alpha/2}; \tag{6.1}$$

see [1, Section 1.2] for its basic properties. (Again, our normalization for the Fourier transform differs from that of [1], but the normalizations of G_α and of the norm on H^α—denoted by $L^{\alpha,2}$ in [1]—are the same.) The **Bessel capacity** of order α of a compact set $A \subseteq \mathbb{R}^n$ may defined in the following dual ways:

$$\begin{aligned} C_\alpha(A) &:= \inf\left\{ \|f\|_{H^\alpha}^2 \;\middle|\; f \in H^\alpha \text{ and } f \ge 1 \text{ on } A \right\} \\ &= \sup_{\mu \in P(A)} \|G_\alpha * \mu\|_{L^2}^{-2} = \sup_{\mu \in P(A) \cap H^{-\alpha}} \|\mu\|_{H^{-\alpha}}^{-2}; \end{aligned} \tag{6.2}$$

see Definition 2.2.6 and Theorem 2.2.7 of [1]. (There is a more general L^p version of the Bessel capacity, a subject of *nonlinear* potential theory, which we need not consider here.) By (4.3), (5.1), and (6.1), it follows that for a compact set $A \subseteq \mathbb{R}^n$,

$$|A|_+ = \frac{1}{n!\omega_n} C_{(n+1)/2}(A). \tag{6.3}$$

Furthermore, the special case $\alpha = (n+1)/2$ of the equality in (6.2) is the same as the special case of Proposition 4.4 for Euclidean space. A corollary of (6.3) is that if $A \subseteq \mathbb{R}$ is compact, then $|A| = \frac{1}{2} C_1(A)$. This follows since $|A| = |A|_+$ for compact $A \subseteq \mathbb{R}$, by [13, Lemma 2.8].

Another classical type of capacity[1] of a compact set $A \subseteq \mathbb{R}^n$ is

$$N_\alpha(A) := \inf\left\{ \|f\|_{H^\alpha}^2 \;\middle|\; f \in \mathcal{S} \text{ and } f \equiv 1 \text{ on a neighborhood of } A \right\}; \tag{6.4}$$

see [1, Definition 2.7.1]. Note that $C_\alpha(A) \le N_\alpha(A)$ trivially. Theorem 4.1 indicates that magnitude in $\mathbb{R}^n$ is closely related to the capacity $N_{(n+1)/2}$, and in fact that for a compact set $A \subseteq \mathbb{R}^n$,

$$|A| \le \frac{1}{n!\omega_n} N_{(n+1)/2}(A). \tag{6.5}$$

There is also a dual formulation of N_α, given in [1, Theorem 2.7.2], which closely parallels Theorem 3.4.

As noted in Section 4, maximum diversity lacks the category-theoretic motivation behind the definition of magnitude, but it is in many ways easier to study than magnitude due to its representation in terms of positive measures. It is interesting to note that an analogous situation appears in potential theory: C_α is simpler to analyze than N_α since it can also be represented in terms of positive measures, but N_α, whose dual formulation requires signed measures or more general distributions, arises naturally in certain applications, cf. [1, p. 47].

The following deep result gives the crucial relationship between the two capacities C_α and N_α which allows C_α to be used in some situations in which N_α appears more naturally.

Proposition 6.1 ([1, Corollary 3.3.4])**.** *For each $\alpha > 0$ and positive integer n there exists a constant $\kappa(\alpha, n) > 0$ such that*

$$C_\alpha(A) \le N_\alpha(A) \le \kappa(\alpha, n) C_\alpha(A)$$

for every compact set $A \subseteq \mathbb{R}^n$.

This proposition essentially specializes to the following result about magnitude in Euclidean space.

[1] The author has been unable to find a standard name for this capacity in the literature.

Corollary 6.2. *For each positive integer n there exists a $\kappa_n > 0$ such that for each compact set $A \subseteq \mathbb{R}^n$,*

$$\left|A\right|_+ \le \left|A\right| \le \kappa_n \left|A\right|_+ .$$

Proof. Setting $\alpha = (n+1)/2$, this follows immediately from (4.4), (6.3), (6.5), and Proposition 6.1. ☐

As mentioned in Section 5 above, it is an open question whether the magnitude function of a compact set $A \subseteq \mathbb{R}^n$ is nondecreasing. One simple consequence of Corollary 6.2 is a partial result in this direction. It follows immediately from (4.3) that the **diversity function** $t \mapsto \left|tA\right|_+$ is nondecreasing for $t > 0$, and Corollary 6.2 implies that

$$\left|tA\right|_+ \le \left|tA\right| \le \kappa_n \left|tA\right|_+ ; \tag{6.6}$$

thus the magnitude function of A is bounded above and below by multiples of a nondecreasing function. In particular, for $t \ge s$ we have that

$$\left|tA\right| \ge \left|tA\right|_+ \ge \left|sA\right|_+ \ge \kappa_n^{-1} \left|sA\right| ,$$

which roughly says that the magnitude function of A never decreases very much. This complements the lower bound in Theorem 5.4, which yields a sharper estimate when t is close to s but a weaker estimate when $t \gg s$. The estimates in (6.6) are most powerful when $t \to \infty$, as will be exploited in Corollary 7.4 below.

At this point the reader may imagine that the theory of magnitude in Euclidean space is already well-explored, albeit under a different name, in the literature on capacities of sets. In fact this is quite far from the truth, and not simply—or even primarily—because of the slight difference between the sets of functions appearing in Corollary 5.3 and (6.4). In typical applications, capacities are used to control the size of "exceptional" sets, and the principal interest is in sets of capacity 0. Proposition 6.1 on the equivalence of C_α and N_α is applied mainly via its corollary that $C_\alpha(A) = 0$ if and only if $N_\alpha(A) = 0$ (see [1, Section 2.9] for discussion and references). However, for any $\alpha > n/2$, picking μ in the supremum in (6.2) to be a point mass shows that $N_\alpha(A) \ge C_\alpha(A) > 0$ for each nonempty compact set $A \subseteq \mathbb{R}^n$. From the point of view of traditional applications of capacities, capacities of order $\alpha > n/2$—including magnitude and maximum diversity, for which one can easily check that $\left|A\right| \ge \left|A\right|_+ \ge 1$ for any compact PDMS A—are thus somewhat pathological.

On the other hand, from the perspective that $\left|A\right|$ is an effective number of points of A (see Section 2), it is perfectly natural that $\left|A\right| \ge 1$ always, and the principal interest is in the magnitude of large sets, particularly in the growth of the magnitude function $t \mapsto \left|tA\right|$ as $t \to \infty$. It is thus quite interesting that Proposition 6.1, which was motivated by applications involving sets of capacity 0, turns out to be a vital ingredient of the proof of Corollary 7.4 below, which is the main result about the growth of magnitude functions in Euclidean space.

7. Dimensions

We now turn to the investigation of the asymptotic growth of the magnitude function. Suppose that A is a compact metric space of negative type. The **upper magnitude dimension** of A is

$$\overline{\dim}_{\mathrm{Mag}} A := \limsup_{t \to \infty} \frac{\log \left|tA\right|}{\log t}$$

and the **lower magnitude dimension** of A is

$$\underline{\dim}_{\mathrm{Mag}} A := \liminf_{t\to\infty} \frac{\log |tA|}{\log t}.$$

When $\overline{\dim}_{\mathrm{Mag}} A = \underline{\dim}_{\mathrm{Mag}} A$, or equivalently $\lim_{t\to\infty} \frac{\log |tA|}{\log t}$ exists, the **magnitude dimension** $\dim_{\mathrm{Mag}} A$ is equal to this limit.

Magnitude dimensions of various subsets of Euclidean space were investigated in [11, 19, 9, 20]. For example, precise asymptotics of the magnitude function—which in particular yield the magnitude dimension—were found for line segments [11, Theorem 7], the Cantor set [11, Theorem 11], and spheres in $\mathbb{R}^n$ [20, Theorem 13]. The magnitude dimension of the Sierpinski gasket was approximated numerically in [19, Section 4]. Theorem 3.5.5 of [9] (which is sharpened by Theorem 5.4 above) implies that subsets of $\mathbb{R}^n$ have upper magnitude dimension of at most n; and [9, Theorem 3.5.6], reproved for Euclidean space as Proposition 5.6 above, implies that subsets of $\mathbb{R}^n$ with positive volume have magnitude dimension equal to n. Theorems 3.4.8 and 3.5.6 of [9] extend these last two facts to the ℓ^1 metric on $\mathbb{R}^n$, and [13, Theorems 4.4 and 4.5] extend them to all ℓ^p metrics on $\mathbb{R}^n$ for $1 \le p \le 2$, and in modified form to some related quasinorms[2]. In all these cases the magnitude dimension was found to agree with classical notions of dimension like Hausdorff or Minkowski dimension (which agree with each other for all these examples).

The main result of this section, Corollary 7.4, unifies and generalizes all the results about magnitude dimension in Euclidean space by proving that it is always equal to Minkowski dimension. Toward this goal, we next recall the definition of Minkowski dimensions for arbitrary compact metric spaces and prove a new characterization of them in terms of maximum diversity.

Let A be any compact metric space. For $\varepsilon > 0$, the **packing number** $M(A, \varepsilon)$ is the maximum number of disjoint closed ε-balls in A, and the **covering number** $N(A, \varepsilon)$ is the minimum number of closed ε-balls needed to cover A. It is an easy exercise to prove that

$$N(A, 2\varepsilon) \le M(A, \varepsilon) \le N(A, \varepsilon). \tag{7.1}$$

The quantities $\log N(A, \varepsilon)$ and $\log M(A, \varepsilon)$ are called the ε-**entropy** and ε-**capacity** of A, respectively.

The **upper Minkowski dimension** of A is

$$\overline{\dim}_{\mathrm{Mink}} A := \limsup_{\varepsilon\to 0^+} \frac{\log N(A, \varepsilon)}{\log(1/\varepsilon)} = \limsup_{\varepsilon\to 0^+} \frac{\log M(A, \varepsilon)}{\log(1/\varepsilon)}$$

and the **lower Minkowski dimension** of A is

$$\underline{\dim}_{\mathrm{Mink}} A := \liminf_{\varepsilon\to 0^+} \frac{\log N(A, \varepsilon)}{\log(1/\varepsilon)} = \liminf_{\varepsilon\to 0^+} \frac{\log M(A, \varepsilon)}{\log(1/\varepsilon)}.$$

The equalities between the two expressions for each of these dimensions follow from the equivalence (7.1) of packing and covering numbers. When $\overline{\dim}_{\mathrm{Mink}} A = \underline{\dim}_{\mathrm{Mink}} A$, or equivalently $\lim_{\varepsilon\to 0^+} \frac{\log N(A,\varepsilon)}{\log(1/\varepsilon)}$ exists, the **Minkowski dimension** $\dim_{\mathrm{Mink}} A$ is equal to this limit.

The **upper**, **lower**, and ordinary **diversity dimension** of an arbitrary compact metric space A are defined analogously to magnitude dimensions, using the maximum diversity

[2] The last paragraph of the published version of [13] misstates the consequences for magnitude dimension of those results when the quasinorm in question has homogeneity of a degree smaller than 1. A corrected version will be posted to the arXiv.

$|tA|_+$ in place of the magnitude $|tA|$; we denote them by $\underline{\dim}_{\mathrm{Div}}$, $\overline{\dim}_{\mathrm{Div}}$, and $\dim_{\mathrm{Div}}$ respectively. Observe that by (4.4), for any compact space A of negative type,

$$\underline{\dim}_{\mathrm{Div}} A \le \underline{\dim}_{\mathrm{Mag}} A \qquad \text{and} \qquad \overline{\dim}_{\mathrm{Div}} A \le \overline{\dim}_{\mathrm{Mag}} A. \tag{7.2}$$

Theorem 7.1. *For any compact metric space A, $\underline{\dim}_{\mathrm{Div}} A = \underline{\dim}_{\mathrm{Mink}} A$ and $\overline{\dim}_{\mathrm{Div}} A = \overline{\dim}_{\mathrm{Mink}} A$. Consequently, $\dim_{\mathrm{Div}} A$ is defined if and only if $\dim_{\mathrm{Mink}} A$ is defined, and in that case $\dim_{\mathrm{Div}} A = \dim_{\mathrm{Mink}} A$.*

Proof. We prove first that $\underline{\dim}_{\mathrm{Div}} A \le \underline{\dim}_{\mathrm{Mink}} A$ and $\overline{\dim}_{\mathrm{Div}} A \le \overline{\dim}_{\mathrm{Mink}} A$. Let $\varepsilon > 0$, $t > 0$, and $\mu \in P(A)$ be given. Observe that for each $a \in A$,

$$\int e^{-td(a,b)} \, d\mu(b) \ge \int_{B(a,\varepsilon)} e^{-td(a,b)} \, d\mu(b) \ge e^{-t\varepsilon} \mu(B(a,\varepsilon)). \tag{7.3}$$

Therefore, by Jensen's inequality,

$$\left(\int \int e^{-td(a,b)} \, d\mu(a) \, d\mu(b) \right)^{-1} \le e^{t\varepsilon} \left(\int \mu(B(a,\varepsilon)) \, d\mu(a) \right)^{-1} \le e^{t\varepsilon} \int \frac{1}{\mu(B(a,\varepsilon))} \, d\mu(a).$$

Now let $N = N(A, \varepsilon/2)$, and let $a_1, \dots, a_N \in A$ such that $A = \bigcup_{j=1}^N B(a_j, \varepsilon/2)$. Suppose for the moment that $\mu(B(a_j, \varepsilon/2)) > 0$ for each j. If $a \in B(a_j, \varepsilon/2)$ then $B(a_j, \varepsilon/2) \subseteq B(a, \varepsilon)$, and so

$$\begin{aligned} \int \frac{1}{\mu(B(a,\varepsilon))} \, d\mu(a) &\le \sum_{j=1}^N \int_{B(a_j,\varepsilon/2)} \frac{1}{\mu(B(a,\varepsilon))} \, d\mu(a) \le \sum_{j=1}^N \int_{B(a_j,\varepsilon/2)} \frac{1}{\mu(B(a_j,\varepsilon/2))} \, d\mu(a) \\ &= \sum_{j=1}^N \frac{\mu(B(a_j,\varepsilon/2))}{\mu(B(a_j,\varepsilon/2))} = N. \end{aligned}$$

If $\mu(B(a_j, \varepsilon/2) = 0$ for some j, the sums above should be restricted to those j for which $\mu(B(a_j, \varepsilon/2)) > 0$, and one still obtains the upper bound of N. Altogether, we have that

$$|tA|_+ = \sup_{\mu \in P(A)} \left(\int \int e^{-td(a,b)} \, d\mu(a) \, d\mu(b) \right)^{-1} \le e^{t\varepsilon} N(A, \varepsilon/2).$$

Setting $\varepsilon = 2/t$, this suffices to prove that $\underline{\dim}_{\mathrm{Div}} A \le \underline{\dim}_{\mathrm{Mink}} A$ and $\overline{\dim}_{\mathrm{Div}} A \le \overline{\dim}_{\mathrm{Mink}} A$.

We next prove that $\underline{\dim}_{\mathrm{Mink}} A \le \underline{\dim}_{\mathrm{Div}} A$ and $\overline{\dim}_{\mathrm{Mink}} A \le \overline{\dim}_{\mathrm{Div}} A$. Let $\varepsilon > 0$ and $t > 0$ be given. Let $M = M(A, \varepsilon)$, let $a_1, \dots, a_M$ be the centers of disjoint closed ε-balls in A, and define

$$\mu := \frac{1}{M} \sum_{j=1}^M \delta_{a_j} \in P(A).$$

For each $a \in A$, there is at most one a_j in $B(a, \varepsilon)$, and so

$$\int e^{-td(a,b)} d\mu(b) = \frac{1}{M} \sum_{j=1}^M e^{-td(a,a_j)} \le \frac{1}{M} + e^{-t\varepsilon}.$$

It follows that

$$\frac{1}{|tA|_+} \le \frac{1}{M(A,\varepsilon)} + e^{-t\varepsilon}.$$

Now define $\varepsilon(t) := \frac{\log(2|tA|_+)}{t}$ for $t \ge 1$, so that $M(A, \varepsilon(t)) \le 2 |tA|_+$. We will prove below that $\varepsilon(t)$ is a continuous and strictly decreasing function of t; for now, assume this to be the case. If $\varepsilon(t)$ is bounded from below by a positive constant, then $\underline{\dim}_{\mathrm{Div}} A = \infty$ and the desired inequalities hold trivially. We may thus assume that $\varepsilon(t) \to 0$ as $t \to \infty$, and so

$$\frac{M(A, \varepsilon(t))}{\log(1/\varepsilon(t))} \le \frac{\log(2\left|tA\right|_+)}{\log(1/\varepsilon(t))} = \frac{\log\left|tA\right|_+}{\log t}\frac{\log t}{\log(1/\varepsilon(t))} + o(1)$$

as $t \to \infty$. Now

$$\frac{\log(1/\varepsilon(t))}{\log t} = 1 - \frac{\log\log(2\left|tA\right|_+)}{\log t},$$

and so if $t_n \to \infty$ such that $\frac{\log|t_nA|_+}{\log t_n}$ is bounded above, then $\frac{\log(1/\varepsilon(t_n))}{\log t_n} \to 1$ and thus

$$\frac{M(A, \varepsilon(t_n))}{\log(1/\varepsilon(t_n))} \le \frac{\log\left|t_nA\right|_+}{\log t_n}\big(1 + o(1)\big)$$

as $n \to \infty$. If $\underline{\dim}_{\mathrm{Div}} A < \infty$ then $\frac{\log|t_nA|_+}{\log t_n}$ is bounded above for some sequence $t_n \to \infty$, and so

$$\underline{\dim}_{\mathrm{Mink}} A \le \liminf_{n\to\infty} \frac{M(A, \varepsilon(t_n))}{\log(1/\varepsilon(t_n))} \le \liminf_{n\to\infty} \frac{\log\left|t_nA\right|_+}{\log t_n} = \underline{\dim}_{\mathrm{Div}} A.$$

If $\overline{\dim}_{\mathrm{Div}} A < \infty$, then $\frac{\log|tA|_+}{\log t}$ is bounded above for all $t \ge 1$, and since $\varepsilon : [1, \infty) \to (0, \varepsilon(1)]$ is bijective, we similarly obtain that $\overline{\dim}_{\mathrm{Mink}} A < \overline{\dim}_{\mathrm{Div}} A$.

It remains to show that $\varepsilon(t)$ is continuous and strictly decreasing on $(1, \infty)$. The continuity follows from [13, Proposition 2.11]. By definition,

$$\left|tA\right|_+^{1/t} = \sup_{\mu \in P(A)} \left\| e^{-d(\cdot,\cdot)} \right\|_{L^t(\mu\otimes\mu)}^{-1}.$$

By either Hölder's inequality or Jensen's inequality, for each fixed $\mu \in P(A)$, $\left\| e^{-d(\cdot,\cdot)} \right\|_{L^t(\mu\otimes\mu)}$ is a nondecreasing function of t. Thus $\left|tA\right|_+^{1/t}$ is the supremum of a family of nonincreasing functions of t, hence nonincreasing, and so $2^{1/t}\left|tA\right|_+^{1/t}$ is a strictly decreasing function of t. The claim follows since the logarithm is a strictly increasing function. □

In the setting of Euclidean space, Theorem 7.1 amounts to a characterization of Minkowski dimension in terms of Bessel capacities $C_{(n+1)/2}$. There are well-known relationships between the Hausdorff dimension of sets in $\mathbb{R}^n$ and Bessel capacities C_α for $\alpha \le n/2$ (see [1, Section 5.1]); this connection between Bessel capacities and Minkowski dimension appears to be new.

Theorem 7.1 yields the following comparison between magnitude dimension and Minkowski dimension for general spaces of negative type.

Corollary 7.2. *Let A be a compact metric space of negative type. Then $\underline{\dim}_{\mathrm{Mink}} A \le \underline{\dim}_{\mathrm{Mag}} A$ and $\overline{\dim}_{\mathrm{Mink}} A \le \overline{\dim}_{\mathrm{Mag}} A$.*

Proof. This follows immediately from Theorem 7.1 and (7.2). □

For certain classes of compact PDMSs, magnitude is always equal to maximum diversity, for example for subsets of $\mathbb{R}$, ultrametric spaces (see Section 8 for the definition), or homogeneous PDMSs; see [13, Lemma 2.8]. (Note that all subsets of $\mathbb{R}$ and ultrametric spaces are of negative type, but not all homogeneous metric spaces are; see Theorem 3.3 and subsequent remarks in [13].) For spaces of negative type with this property, Theorem 7.1 implies that magnitude dimensions and Minkowski dimensions agree. For example, we have the following consequence for ultrametric spaces. (We omit subsets of $\mathbb{R}$ and homogeneous spaces here since they will be covered by Corollary 7.4 and Proposition 7.5 below.)

Corollary 7.3. *If A is a compact ultrametric space, then $\overline{\dim}_{\mathrm{Mag}} A = \overline{\dim}_{\mathrm{Mink}} A$ and $\underline{\dim}_{\mathrm{Mag}} A = \underline{\dim}_{\mathrm{Mink}} A$. Consequently, $\dim_{\mathrm{Mag}} A$ is defined if and only if $\dim_{\mathrm{Mink}} A$ is defined, and in that case $\dim_{\mathrm{Mag}} A = \dim_{\mathrm{Mink}} A$.*

When combined with Proposition 6.1, Theorem 7.1 has the deeper consequence is that magnitude dimensions and Minkowski dimensions always agree in Euclidean space, fully explaining the various such agreements observed both rigorously and numerically in [9, 11, 19, 20].

Corollary 7.4. *If $A \subseteq \mathbb{R}^n$ is compact, then $\overline{\dim}_{\mathrm{Mag}} A = \overline{\dim}_{\mathrm{Mink}} A$ and $\underline{\dim}_{\mathrm{Mag}} A = \underline{\dim}_{\mathrm{Mink}} A$. Consequently, $\dim_{\mathrm{Mag}} A$ is defined if and only if $\dim_{\mathrm{Mink}} A$ is defined, and in that case $\dim_{\mathrm{Mag}} A = \dim_{\mathrm{Mink}} A$.*

Proof. For a compact set $A \subseteq \mathbb{R}^n$, (6.6) implies that $\overline{\dim}_{\mathrm{Mag}} A = \overline{\dim}_{\mathrm{Div}} A$ and $\underline{\dim}_{\mathrm{Mag}} A = \underline{\dim}_{\mathrm{Div}} A$, and the result follows from Theorem 7.1. □

It remains an open question whether magnitude dimension is equal to Minkowski dimension for each compact metric space of negative type.

Another approach to defining magnitude for infinite spaces, mentioned in Section 4 above, is to define a **weight measure** for a compact metric space A to be a signed measure $\mu \in M(A)$ such that, for each $a \in A$,

$$\int e^{-d(a,b)} \, d\mu(b) = 1;$$

and then define $|A| := \mu(A)$ whenever μ is a weight measure for A. This clearly extends the original definition 2.1 for the magnitude of finite spaces, and, as discussed in Section 4, can be proved to coincide with Definition 3.3 whenever A is a compact PDMS which possesses a weight measure. If tA possesses a weight measure for each $t > 0$, then we define the magnitude function and upper, lower, and ordinary magnitude dimensions of A as before.

This definition of magnitude is useful in particular when A is a compact homogeneous metric space (i.e., the isometry group acts transitively on the points of A). In this case there exists a unique isometry-invariant probability measure $\mu \in P(A)$ (see, e.g., [14, Theorem 1.3]), which is also isometry-invariant on tA for each $t > 0$. Theorem 1 of [20] then shows that an appropriate scalar multiple of μ is a weight measure for tA, and for each $a \in A$,

$$|tA| = \left(\int e^{-td(a,b)} \, d\mu(b) \right)^{-1}. \tag{7.4}$$

Using this definition of magnitude, precise asymptotics for the magnitude function of a compact homogeneous Riemannian manifold were found in [20, Theorem 11]; these imply that for such manifolds, the magnitude dimension equals the usual dimension. Similar arguments as in the proof of Theorem 7.1 generalize this fact—with Minkowski dimension in place of the dimension of a manifold—to arbitrary compact homogeneous metric spaces. The existence of an invariant weight measure takes the place in this setting of the equivalence of capacities from Proposition 6.1.

Proposition 7.5. *If A is a compact homogeneous metric space then $\overline{\dim}_{\mathrm{Mag}} A = \overline{\dim}_{\mathrm{Mink}} A$ and $\underline{\dim}_{\mathrm{Mag}} A = \underline{\dim}_{\mathrm{Mink}} A$. Consequently, $\dim_{\mathrm{Mag}} A$ is defined if and only if $\dim_{\mathrm{Mink}} A$ is defined, and in that case $\dim_{\mathrm{Mag}} A = \dim_{\mathrm{Mink}} A$.*

Proof. Let μ be the unique isometry-invariant probability measure on A. Let $N = N(A,\varepsilon)$, and let $a_1,\dots,a_N \in A$ such that $A = \bigcup_{j=1}^N B(a_j,\varepsilon)$. Then for each $a \in A$,

$$1 = \mu(A) \le \sum_{j=1}^N \mu(B(a_j,\varepsilon)) = N\mu(B(a,\varepsilon)).$$

Together with (7.3) and (7.4), this implies that

$$|tA| \le e^{t\varepsilon} N(A,\varepsilon).$$

Setting $\varepsilon = 1/t$, this suffices to prove that $\underline{\dim}_{\mathrm{Mag}} A \le \underline{\dim}_{\mathrm{Mink}} A$ and $\overline{\dim}_{\mathrm{Mag}} A \le \overline{\dim}_{\mathrm{Mink}} A$.

Similarly, if $M = M(A,\varepsilon)$ and $a_1,\dots,a_M \in A$ are the centers of disjoint balls of radius ε, then for each $a \in A$,

$$1 = \mu(A) \ge \sum_{j=1}^M \mu(B(a_j,\varepsilon)) = M\mu(B(a,\varepsilon)).$$

Therefore,

$$\begin{aligned}\int e^{-td(a,b)}\ d\mu(b) &= \int_{B(a,\varepsilon)} e^{-td(a,b)}\ d\mu(b) + \int_{A\setminus B(a,\varepsilon)} e^{-td(a,b)}\ d\mu(b) \\ &\le \mu(B(a,\varepsilon)) + e^{-t\varepsilon} \le \frac{1}{M(A,\varepsilon)} + e^{-t\varepsilon}.\end{aligned}$$

Together with (7.4), this implies that

$$\frac{1}{|tA|} \le \frac{1}{M(A,\varepsilon)} + e^{-t\varepsilon},$$

and the proof is completed as in the second half of the proof of Theorem 7.1. □

8. Afterword: ultramagnitude of ultrametric spaces

As discussed in the introduction, the magnitude of a finite metric space is a special case of the more general notion of the magnitude of a finite enriched category, presented in [9, Section 1]. Besides the cases of ordinary categories (for which, as mentioned earlier, magnitude is known as Euler characteristic, and is related to more classical invariants of that name) and of metric spaces, the magnitude of enriched categories has mostly not yet been very fully explored (see [8] for a discussion). In this section, we work out another special case, that of ultrametric spaces. We will see that this leads to an extremely simple notion of the size of an ultrametric space, whose theory is similar to, but drastically simpler than, the theory of magnitude of metric spaces. In particular, the notions of packings, coverings, and Minkowski dimensions, which played central roles in the previous section, come immediately out of this theory.

An **ultrametric space** is a metric space (A,d) which satisfies the strengthened triangle inequality

$$d(a,c) \le \max\bigl\{d(a,b), d(b,c)\bigr\}$$

for each $a,b,c \in A$. That is, one obtains the definition of an ultrametric space by replacing the binary operation $+$ in the definition of a metric space with the binary operation max. Of course, ultrametric spaces are in particular metric spaces, and are even always positive definite (see [9, Proposition 2.4.18] or [13, Theorem 3.6]). Thus, one can speak of the magnitude of a compact ultrametric space, as we have done in Corollary 7.3 above. However, one obtains a different notion if one appropriately substitutes the operation of max for the operation $+$, not in Definition 2.1 itself, but in the category-theoretic considerations which motivate that definition. To distinguish this new notion from the magnitude of A when thought of simply as a metric space, we will call it “ultramagnitude”.

Definition 2.1 of the magnitude of a metric space A is built around the matrix $\zeta(a,b) = e^{-d(a,b)}$, the motivation for which was not explained in this paper. We will now explain just one part of its motivation—the part which needs to be modified for ultrametric spaces. The reader is referred to [9, Section 1] and the references therein for the full definition of magnitude of an enriched category and the category-theoretic background on which it is built, or to [11, Section 1.1] for a very brief summary. For brevity's sake, we will bring up here only the essential minimum, noting for experts that whereas a metric space is a category enriched over the monoidal category $\big(([0,\infty), \ge), +, 0\big)$, an ultrametric space is a category enriched over $\big(([0,\infty), \ge), \max, 0\big)$.

When specialized to metric spaces, Leinster's definition of the magnitude of an enriched category calls for a function $\Phi : [0,\infty) \to \mathbb{R}$ such that

$$\Phi(x+y) = \Phi(x)\Phi(y)$$

for each $x, y \in [0,\infty)$. One then defines $\zeta(a,b) = \Phi(d(a,b))$. Here the domain $[0,\infty)$ is the set of possible distances (the objects of the enriching category) and the operation $+$ in $x+y$ is the same operation appearing in the triangle inequality (the tensor product in the enriching category). If Φ is to be Lebesgue measurable, we must have $\Phi(x) = \alpha^x$ for some $\alpha \ge 0$ (see [3]); the choice of $\alpha = e^{-1}$ is the arbitrary choice of scale in Definition 2.1 which is addressed by considering magnitude functions.

To adapt Definition 2.1 to ultrametric spaces, we instead need a function $\Psi : [0,\infty) \to \mathbb{R}$ such that

$$\Psi\big(\max\{x,y\}\big) = \Psi(x)\Psi(y)$$

for each $x, y \in [0,\infty)$. Setting $x = y$ shows that Ψ can take only values in $\{0,1\}$. If $\Psi(x) = 0$ then $\Psi(y) = 0$ for any $y > x$ as well, and so Ψ must be the indicator function of some interval $[0,\beta]$ or $[0,\beta)$. We choose $\beta = 1$, which, like the choice $\alpha = e^{-1}$ above, amounts to a convenient but arbitrary choice of scale. We furthermore pick Ψ to be the indicator function of $[0,1]$, which amounts to a choice to work with closed balls as opposed to open balls. The general definition of the magnitude of a finite enriched category then specializes to ultrametric spaces in the following way.

Definition 8.1. *Given a finite ultrametric space (A,d), define the matrix $\xi \in \mathbb{R}^{A\times A}$ by*

$$\xi(a,b) := \begin{cases} 1 & \text{if } d(a,b) \le 1, \\ 0 & \text{if } d(a,b) > 1. \end{cases}$$

A vector $w \in \mathbb{R}^A$ is an **ultraweighting** *for A if for each $a \in A$,*

$$(\xi w)(a) = \sum_{b\in A} \xi(a,b) w(b) = 1.$$

If A possesses an ultraweighting w, then the **ultramagnitude** *of A is*

$$|A|_U := \sum_{a\in A} w(a).$$

Observe that the matrix ξ used in Definition 8.1 is a discretization of the matrix ζ from Definition 2.1. This suggests that magnitude should reflect finer information than ultramagnitude.

The entire theory of the ultramagnitude of finite ultrametric spaces can be summed up in the following result.

Proposition 8.2. *Let A be a finite ultrametric space. Then $|A|_U = N(A,1) = M(A,1)$.*

Proof. It is an easy exercise to check that the closed balls of radius 1 in an ultrametric space A form a partition of A, which must thus consist of $N(A,1) = M(A,1)$ distinct balls. An ultraweighting for A is given by

$$w(a) = \big(\#B(a,1)\big)^{-1},$$

where # denotes cardinality. Therefore,

$$|A|_U = \sum_{a\in A} w(a) = \sum_{\text{distinct } B(a,1)} 1 = N(A,1). \qquad \square$$

Given Proposition 8.2, it is simple to extend the definition of ultramagnitude to compact ultrametric spaces in a natural way: we simply let $|A|_U = N(A,1)$. (One can also arrive at this definition, with rather more effort, by appropriately modifying the approach of either [13] or Section 3 above, but we will not pursue this here.) For $t > 0$, it follows that $|tA|_U = N(A,1/t) = M(A,1/t)$. Thus the **ultramagnitude function** $t \mapsto |tA|_U$ contains precisely the same information as the ε-entropy or ε-capacity of a compact ultrametric space, and it follows trivially that "ultramagnitude dimension" is equal to Minkowski dimension.

References

[1] D. R. Adams and L. I. Hedberg. *Function Spaces and Potential Theory*, volume 314 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1996.

[2] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.

[3] M. Fréchet. Pri la funkcia ekvacio $f(x+y) = f(x) + f(y)$. *L'Enseignement Mathématique*, 15:390–393, 1913.

[4] L. Hörmander. *The Analysis of Linear Partial Differential Operators. I: Distribution Theory and Fourier Analysis*. Springer Study Edition. Springer-Verlag, Berlin, second edition, 1990.

[5] F. W. Lawvere. Metric spaces, generalized logic, and closed categories. *Rend. Sem. Mat. Fis. Milano*, 43:135–166 (1974), 1973.

[6] T. Leinster. The Euler characteristic of a category. *Doc. Math.*, 13:21–49, 2008.

[7] T. Leinster. A maximum entropy theorem with applications to the measurement of biodiversity. Preprint available at `http://arxiv.org/abs/0910.0906`, 2009.

[8] T. Leinster. The magnitude of an enriched category. Post at The n-Category Café, `http://golem.ph.utexas.edu/category/2011/06/the_magnitude_of_an_enriched_c.html`, 2011.

[9] T. Leinster. The magnitude of metric spaces. *Doc. Math.*, 18:857–905, 2013.

[10] T. Leinster and C. Cobbold. Measuring diversity: the importance of species similarity. *Ecology*, 93:477–489, 2012.

[11] T. Leinster and S. Willerton. On the asymptotic magnitude of subsets of Euclidean space. *Geom. Dedicata*, 164(1):287–310, 2013.

[12] E. H. Lieb and M. Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2001.

[13] M. Meckes. Positive definite metric spaces. *Positivity*, 17(3):733–757, 2013.

[14] V. D. Milman and G. Schechtman. *Asymptotic Theory of Finite-Dimensional Normed Spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986.

[15] W. Rudin. *Functional Analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition, 1991.

[16] A. R. Solow and S. Polasky. Measuring biological diversity. *Environmental and Ecological Statistics*, 1:95–107, 1994.

[17] E. M. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*, volume 32 of *Princeton Mathematical Series*. Princeton University Press, Princeton, N.J., 1971.

[18] F. Trèves. *Introduction to Pseudodifferential and Fourier Integral Operators. Vol. 1*. Plenum Press, New York, 1980.

[19] S. Willerton. Heuristic and computer calculations for the magnitude of metric spaces. Preprint available at `http://arxiv.org/abs/0910.5500`, 2009.

[20] S. Willerton. On the magnitude of spheres, surfaces and other homogeneous spaces. To appear in *Geom. Dedicata*. Preprint available at `http://arxiv.org/abs/1005.4041`, 2010.

Case Western Reserve University, Cleveland, OH 44106, U.S.A.
E-mail address: `mark.meckes@case.edu`