

Tom Leinster and Mark W. Meckes

# The magnitude of a metric space: from category theory to geometric measure theory

## 5.1 Introduction

Magnitude is a numerical isometric invariant of metric spaces. Its definition arises by viewing a metric space as a kind of *enriched category* — an abstract structure that appears more algebraic than geometric in nature — and adapting a construction from the intersection of category theory and homotopy theory. One would hardly expect, from such a provenance, that magnitude would have any strong relationship to geometry as usually conceived. Surprisingly, however, magnitude turns out to encode many invariants from integral geometry and geometric measure theory, including volume, capacity, dimension, and intrinsic volumes. This paper will give an overview of the theory of magnitude, from its category-theoretic genesis to its connections with these geometric quantities.

We begin with a brief overview of the history of magnitude so far. The grandparent of magnitude is the Euler characteristic of a topological space, which is a natural analogue of the cardinality of a finite set. To each category there is associated a topological space called its *classifying space*. In [16], a formula was found for the Euler characteristic of the classifying space of a suitably nice finite category; applying this formula to less nice categories (for which the Euler characteristic of the classifying space need not exist) yielded a new cardinality-like invariant of categories, again called the *Euler characteristic* of a finite category.

Categories are a special case of a more general family of structures, *enriched categories*, which encompass both categories with additional structure (like linear categories) and, surprisingly, metric spaces. In [19, 23], the definition of Euler characteristic of a category was generalized to enriched categories, renamed *magnitude*, then re-specialized to finite metric spaces. The first paper to be written on magnitude [23] focused on the asymptotic behavior of the magnitudes of finite approximations to specific compact subsets of Euclidean space. The results there hinted strongly that magnitude is closely related to geometric quantities including volume and fractal dimension; numerical computations in [40] gave further evidence of these relationships.

---

**Tom Leinster:** School of Mathematics, University of Edinburgh, James Clerk Maxwell Building, Peter Guthrie Tait Road, Edinburgh EH9 3FD, United Kingdom, E-mail: Tom.Leinster@ed.ac.uk

**Mark W. Meckes:** Department of Mathematics, Applied Mathematics, and Statistics, Case Western Reserve University, 10900 Euclid Ave., Cleveland, Ohio 44106, U.S.A., email: mark.meckes@case.edu



© 2017 Tom Leinster and Mark W. Meckes

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 License.

In [41], a definition was proposed for the magnitude of certain compact metric spaces, and connections were found between magnitude and some intrinsic volumes of Riemannian manifolds. Shortly thereafter, the paper [19] appeared which laid out for the first time the general theory of the magnitude of finite metric spaces; and [27] which put the asymptotic approach of [23] for studying magnitude of compact spaces on firm footing, and showed that it also coincides with the definition used in [41].

The paper [28] introduced yet another equivalent approach to magnitude for compact spaces, which makes magnitude more accessible to a wide variety of analytic techniques. Using a result from potential theory, [28] showed in particular that magnitude can be used to recover the Minkowski dimension of a compact set in Euclidean space. Following the approach of [28], the paper [3] applied Fourier analysis to show that magnitude also recovers volume in Euclidean space, and applied PDE techniques to compute precisely magnitudes of Euclidean balls.

This paper aims to serve as a guide to the path from the definition of the Euler characteristic of a finite category, to the geometric results of [28] and [3] on magnitude in Euclidean space. It also includes a number of new results, in particular a significant partial result toward a conjecture from [19] relating magnitude in  $\ell_1^n$  to a family of intrinsic volumes adapted to the  $\ell_1$  metric, as well as generalizations of several regularity results for magnitude from Euclidean space to more general normed spaces. In order to reach the results of geometric interest as quickly as possible, we omit many results from the papers named above, and depart significantly at some points from the historical development of ideas. We give complete proofs only for the new results, and for a few known results for which we take a more direct approach than in previous papers.

Section 5.2 begins with the definition of the Euler characteristic of a finite category, and leads up to the magnitude of a finite metric space and its basic properties. Section 5.3 covers the definition of the magnitude of a compact space, its basic properties, and the results on magnitude of manifolds. Section 5.4 covers magnitude in (quasi)normed spaces, particularly  $\ell_1^n$  and Euclidean space, and contains the new results of this paper. Finally, in section 5.5, we discuss a number of open problems about magnitude.

Before moving on, we need to mention two threads in the story of magnitude which have been ignored above and will make only brief appearances in this paper. The first is the magnitude of a graph, viewed as a metric space with the shortest-path distance between vertices. This subject has been developed in [20], which in particular investigated its relationship to classical, combinatorial graph invariants, and [8], which found that the magnitude of graphs is the Euler characteristic associated to a graded homology theory for graphs. The second thread is the connection of magnitude to quantifying biodiversity and maximum entropy problems. This is actually related with the historically first appearance of the magnitude of a metric space in the literature, in [35], and was developed in [17, 22]; section 5.3.3 will take half a step in the direction of these connections.

## 5.2 Finite metric spaces

Here we explain the origins of the notion of magnitude. There is a simple combinatorial definition of the magnitude or Euler characteristic of a finite category (section 5.2.2), which extends in a natural way to a more general class of structures, the enriched categories (section 5.2.3). As we show, this general invariant is closely related to several existing invariants of size. Specializing it in a different direction gives the definition of the magnitude of a finite metric space (sections 5.2.4 and 5.2.5).

In order to do any of this, we first need to define the magnitude of a matrix.

### 5.2.1 The magnitude of a matrix

Recall that a **semiring** is a “ring without negatives”, that is, an abelian group (written additively) with an associative operation of multiplication that distributes over addition. Let  $k$  be a commutative semiring (always assumed to have a multiplicative identity 1) and  $A$  a finite set, and let  $Z \in k^{A \times A}$  be a square matrix over  $k$  indexed by the elements of  $A$ . A **weighting** on  $Z$  is a column vector  $w \in k^A$  satisfying  $Zw = e$ , where  $e$  is the column vector of 1s, and a **coweighting** on  $Z$  is a row vector  $v \in k^A$  satisfying  $vZ = e^T$ . That is,

$$\sum_{b \in A} Z(a, b)w_b = 1 \text{ for every } a \in A$$

and

$$\sum_{a \in A} v_a Z(a, b) = 1 \text{ for every } b \in A.$$

If  $w$  is a weighting and  $v$  a coweighting on  $Z$  then

$$\sum_{a \in A} w_a = e^T w = vZw = ve = \sum_{a \in A} v_a.$$

When  $Z$  admits both a weighting and a coweighting, we may therefore define the **magnitude**  $|Z|$  of  $Z$  to be the common quantity  $\sum_a w_a = \sum_a v_a$ , for any weighting  $w$  and coweighting  $v$ .

An important special case is when  $Z$  is invertible. Then  $Z$  has a unique weighting and a unique coweighting, and its magnitude is the sum of the entries of  $Z^{-1}$ :

$$|Z| = \sum_{a, b \in A} Z^{-1}(a, b). \quad (5.1)$$

An even more special case is that of positive definite matrices:

**Proposition 5.2.1.** *Let  $Z \in \mathbb{R}^A$  be a positive definite matrix. Then*

$$|Z| = \sup_{0 \neq x \in \mathbb{R}^A} \frac{(\sum_a x_a)^2}{x^T Z x},$$

and the supremum is attained exactly when  $x$  is a scalar multiple of the unique weighting on  $Z$ .

This follows swiftly from the Cauchy–Schwarz inequality [19, Proposition 2.4.3].

### 5.2.2 The Euler characteristic of a finite category

A category can be viewed as a directed graph (allowing multiple parallel edges) together with an associative, unital operation of composition. The vertices of the graph are the objects of the category, and for each pair  $(a, b)$  of vertices, the edges from  $a$  to  $b$  in the graph are the maps from  $a$  to  $b$  in the category, which form a set  $\text{Hom}(a, b)$ . Thus, composition defines a function  $\text{Hom}(a, b) \times \text{Hom}(b, c) \rightarrow \text{Hom}(a, c)$  for each  $a, b, c$ , and there is a loop  $1_a \in \text{Hom}(a, a)$  on each vertex  $a$ . Although in many categories of interest, the collections of objects and maps form infinite sets or even proper classes, we will be considering **finite categories**: those with only finitely many objects and maps.

Let  $\mathbf{A}$  be a finite category, with set of objects  $\text{ob } \mathbf{A}$ . The **Euler characteristic** of  $\mathbf{A}$  is the magnitude of the matrix  $Z_{\mathbf{A}} \in \mathbb{Q}^{\text{ob } \mathbf{A} \times \text{ob } \mathbf{A}}$  given by  $Z_{\mathbf{A}}(a, b) = \#\text{Hom}(a, b)$  (where  $\#$  denotes cardinality), whenever this magnitude is defined.

For example, if  $\mathbf{A}$  has no maps other than identities then  $Z_{\mathbf{A}}$  is the identity and the Euler characteristic of  $\mathbf{A}$  is simply the number of objects. More generally, any partially ordered set  $(P, \leq)$  gives rise to a category  $\mathbf{A}$  whose objects are the elements of  $P$ , and with one map  $a \rightarrow b$  when  $a \leq b$  and none otherwise. In a theory made famous by Rota [30], every finite partially ordered set  $P$  has associated with it a Möbius function  $\mu$ , which is defined on pairs  $(a, b)$  of elements of  $P$  such that  $a \leq b$ , and takes values in  $\mathbb{Z}$ . It generalizes the classical Möbius function, and the construction above for categories generalizes it further still:  $\mu(a, b) = Z_{\mathbf{A}}^{-1}(a, b)$  whenever  $a \leq b$ , and the definition of Euler characteristic of a category extends the existing definition for ordered sets [16, Proposition 4.5].

To any small category  $\mathbf{A}$  there is assigned a topological space, called its **classifying space**. The name “Euler characteristic” is largely justified by the following result.

**Theorem 5.2.2** ([16, Proposition 2.11]). *Let  $\mathbf{A}$  be a finite category. Under appropriate conditions (which imply, in particular, that the Euler characteristic of the classifying space of  $\mathbf{A}$  is defined), the Euler characteristic of the category  $\mathbf{A}$  is equal to the Euler characteristic of its classifying space.*

Euler characteristic for finite categories enjoys many properties analogous to those enjoyed by topological Euler characteristic [16, Section 2]. For instance, categorical Euler characteristic is invariant under equivalence (mirroring homotopy invariance in the topological setting), and is additive with respect to disjoint union of categories and

multiplicative with respect to products. There is even an analogue of the topological formula for the Euler characteristic of the total space of a fibration.

Schanuel [33] argued that Euler characteristic for topological spaces is closely analogous to cardinality for sets. For instance, it has analogous additivity and multiplicativity properties, it satisfies the inclusion-exclusion principle (under hypotheses), and, indeed, it reduces to cardinality for finite discrete spaces. Similarly, the results described above suggest that Euler characteristic for finite categories is the categorical analogue of cardinality.

### 5.2.3 Enriched categories

A **monoidal category** is a category  $\mathcal{V}$  equipped with an associative binary operation  $\otimes$  (which is formally a functor  $\mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$ ) and a unit object  $\mathbb{1} \in \mathcal{V}$ . The associativity and unit axioms are only required to hold up to suitably coherent isomorphism; see [26] for details.

Typical examples of monoidal categories  $(\mathcal{V}, \otimes, \mathbb{1})$  are the categories  $(\text{Set}, \times, \{*\})$  of sets with cartesian product and  $(\mathbf{FDVect}_K, \otimes, K)$  of finite-dimensional vector spaces over a field  $K$ . A less obvious example is the ordered set  $([0, \infty], \geq)$ . As a category, its objects are the nonnegative reals together with  $\infty$ , there is one map  $x \rightarrow y$  when  $x \geq y$ , and there are none otherwise. It is monoidal with  $\otimes = +$  and  $\mathbb{1} = 0$ .

Let  $\mathcal{V} = (\mathcal{V}, \otimes, \mathbb{1})$  be a monoidal category. The definition of category enriched in  $\mathcal{V}$ , or  $\mathcal{V}$ -category, is obtained from the definition of ordinary category by requiring that the hom-sets are no longer sets but objects of  $\mathcal{V}$ . Thus, a (small)  $\mathcal{V}$ -**category**  $\mathbf{A}$  consists of a set  $\text{ob } \mathbf{A}$  of objects, an object  $\text{Hom}(a, b)$  of  $\mathcal{V}$  for each  $a, b \in \text{ob } \mathbf{A}$ , and operations of composition and identity satisfying appropriate axioms [10]. The composition consists of a map

$$\text{Hom}(a, b) \otimes \text{Hom}(b, c) \rightarrow \text{Hom}(a, c)$$

in  $\mathcal{V}$  for each  $a, b, c \in \text{ob } \mathbf{A}$ , while the identities are provided by a map  $\mathbb{1} \rightarrow \text{Hom}(a, a)$  for each  $a \in \text{ob } \mathbf{A}$ .

- Examples 5.2.3.** 1. When  $\mathcal{V} = \text{Set}$  (with monoidal structure as above), a  $\mathcal{V}$ -category is an ordinary (small) category.
2. When  $\mathcal{V} = \mathbf{Vect}_K$ , a  $\mathcal{V}$ -category is a **linear category**, that is, a category in which each hom-set carries the structure of a vector space, and composition is bilinear.
3. When  $\mathcal{V} = [0, \infty]$ , a  $\mathcal{V}$ -category is a **generalized metric space** [14, 15]. That is, a  $\mathcal{V}$ -category consists of a set  $A$  of objects or points together with, for each  $a, b \in A$ , a real number  $\text{Hom}(a, b) = d(a, b) \in [0, \infty]$ , satisfying the axioms

$$d(a, b) + d(b, c) \geq d(a, c), \quad d(a, a) = 0$$

$(a, b, c \in A)$ . Such spaces are more general than classical metric spaces in three ways:  $\infty$  is permitted as a distance, the separation axiom  $d(a, b) = 0 \implies a = b$  is dropped, and, most significantly,  $d$  is not required to be symmetric.

4. The category  $\mathcal{V} = ([0, \infty], \geq)$  can alternatively be given the monoidal structure  $(\max, 0)$ . A  $\mathcal{V}$ -category is then a generalized ultrametric space, that is, a generalized metric space satisfying the stronger triangle inequality  $\max\{d(a, b), d(b, c)\} \geq d(a, c)$ .

To define the magnitude of an enriched category, we start with a monoidal category  $(\mathcal{V}, \otimes, \mathbb{1})$  together with a commutative semiring  $k$  and a map  $|\cdot| : \text{ob } \mathcal{V} \rightarrow k$ , with the property that  $|X| = |Y|$  whenever  $X \cong Y$ , and satisfying the multiplicativity axioms  $|X \otimes Y| = |X| \cdot |Y|$  and  $|\mathbb{1}| = 1$ .

**Definition 5.2.4.** Let  $\mathbf{A}$  be a  $\mathcal{V}$ -category with only finitely many objects.

1. The **similarity matrix** of  $\mathbf{A}$  is the  $\text{ob } \mathbf{A} \times \text{ob } \mathbf{A}$  matrix  $Z_{\mathbf{A}}$  over  $k$  defined by  $Z_{\mathbf{A}}(a, b) = |\text{Hom}(a, b)|$ .
2. A **(co)weighting** on  $\mathbf{A}$  is a (co)weighting on  $Z_{\mathbf{A}}$ , and  $\mathbf{A}$  **has magnitude** if  $Z_{\mathbf{A}}$  does. Its **magnitude** is then  $|\mathbf{A}| = |Z_{\mathbf{A}}|$ .

**Examples 5.2.5.** 1. Let  $\mathcal{V}$  be the monoidal category  $(\mathbf{FinSet}, \times, \{*\})$  of finite sets. Let  $k = \mathbb{Q}$ , and for  $X \in \mathbf{FinSet}$ , let  $|X| \in \mathbb{Q}$  be the cardinality of  $X$ . Then we obtain a notion of magnitude for finite categories; it is exactly the Euler characteristic of section 5.2.2.

2. Let  $\mathcal{V}$  be the monoidal category  $\mathbf{FDVect}_K$  of finite-dimensional vector spaces over a field  $K$ . Let  $k = \mathbb{Q}$ , and for  $X \in \mathbf{FDVect}_K$ , put  $|X| = \dim X \in \mathbb{Q}$ . Then we obtain a notion of magnitude for linear categories with finitely many objects and finite-dimensional hom-spaces. As shown in [5], this invariant is closely related to the Euler form of an associative algebra, defined homologically.
3. Let  $\mathcal{V} = [0, \infty]$ , with monoidal structure  $(+, 0)$ . Let  $k = \mathbb{R}$ , and for  $x \in [0, \infty]$ , put  $|x| = e^{-x}$ . (We have little choice about this: the multiplicativity axioms force  $|x| = C^x$  for some constant  $C$ , at least assuming that  $|\cdot|$  is to be measurable. We will address the one degree of freedom here through the introduction of magnitude functions in the next section.) Then we obtain a notion of the magnitude  $|\mathbf{A}| \in \mathbb{R}$  of a finite metric space  $|\mathbf{A}|$ , examined in detail later.
4. Let  $\mathcal{V} = [0, \infty]$ , now with monoidal structure  $(\max, 0)$ . Let  $k = \mathbb{R}$ , and define  $|\cdot| : [0, \infty] \rightarrow \mathbb{R}$  to be either the indicator function of  $[0, 1]$  or that of  $[0, 1)$ . It is shown in Section 8 of [28] that these are essentially the only possibilities for  $|\cdot|$ , and that the resulting magnitude of a finite ultrametric space is simply the number of balls of radius 1 (closed or open, respectively) needed to cover it. It is also shown that this leads naturally to the notion of  $\varepsilon$ -entropy or  $\varepsilon$ -capacity.

The multiplicativity condition  $|X \otimes Y| = |X| \cdot |Y|$  on objects of  $\mathcal{V}$  has so far not been used. However, it implies a similar multiplicativity condition on categories enriched in  $\mathcal{V}$ . In the case of metric spaces, this reduces to Proposition 5.2.8 below; for the general statement, see [19, Proposition 1.4.3].

#### 5.2.4 The magnitude of a finite metric space

Concretely, the magnitude  $|A|$  of a finite metric space  $(A, d)$  is the magnitude of the matrix  $Z = Z_A \in \mathbb{R}^{A \times A}$  given by  $Z_A(a, b) = e^{-d(a, b)}$ , if that is defined. Taking advantage of the symmetry of  $Z_A$  to simplify slightly, this means the following. A vector  $w \in \mathbb{R}^A$  is a **weighting** for  $A$  if  $Z_A w = e$ , where  $e \in \mathbb{R}^A$  is the column vector of 1s, and if a weighting for  $A$  exists, then the **magnitude** of  $A$  is

$$|A| = \sum_{a \in A} w_a.$$

This is not a classical invariant, or one that appears to have previously been explored mathematically prior to the work cited in the introduction. Neither is it wholly new. In a probabilistic analysis of the benefits of highly diverse ecosystems, Solow and Polasky [35] derived a lower bound on the benefit and identified one term, which they called the “effective number of species”, as especially interesting. Although it was not thoroughly investigated in [35], this term is exactly our magnitude. The reader is referred to [17, 19, 21, 22] for more information about this connection.

Not every finite metric space possesses a weighting or, therefore, has well-defined magnitude. One large and important class of spaces which always does is the subject of section 5.2.5. The next two results give additional examples.

From now on, to simplify the statements of results, all metric spaces and all compact sets in a metric space are assumed to be nonempty.

**Proposition 5.2.6** ([23, Theorem 2] and [19, Proposition 2.1.3]). *Let  $(A, d)$  be a finite metric space, and suppose that whenever  $a, b \in A$  with  $a \neq b$ , we have  $d(a, b) > \log(\#A - 1)$ . Then  $A$  possesses a positive weighting, and  $|A|$  is therefore defined.*

A metric space  $(A, d)$  is called **homogeneous** if its isometry group acts transitively on the points of  $A$ .

**Proposition 5.2.7** ([36]; see also [19, Proposition 2.1.5]). *If  $(A, d)$  is a finite homogeneous metric space and  $a_0 \in A$  is any fixed point, then  $A$  possesses a positive weighting and*

$$|A| = \frac{(\#A)^2}{\sum_{a, b \in A} e^{-d(a, b)}} = \frac{\#A}{\sum_{a \in A} e^{-d(a, a_0)}}.$$

For metric spaces  $(A, d_A)$  and  $(B, d_B)$ , we denote by  $A \times_1 B$  the set  $A \times B$  equipped with the metric

$$d((a, b), (a', b')) = d_A(a, a') + d_B(b, b').$$

**Proposition 5.2.8** ([19, Proposition 2.3.6]). *Suppose that  $(A, d_A)$  and  $(B, d_B)$  are finite metric spaces with weightings  $w \in \mathbb{R}^A$  and  $v \in \mathbb{R}^B$  respectively. Then  $x \in \mathbb{R}^{A \times B}$  given by  $x_{(a,b)} = w_a v_b$  is a weighting for  $A \times_1 B$ , and  $|A \times_1 B| = |A| |B|$ .*

Proposition 5.2.8 has a generalization, Theorem 2.3.11 of [19], which is an analogue for magnitude of the formula for the Euler characteristic of the total space of a fibration.

As noted earlier, there is an arbitrary choice of scale implicit in the definition of magnitude: we could choose any other base for the exponent in place of  $e^{-1}$ . To deal with this, we will often work with the whole family of metric spaces  $\{tA\}_{t>0}$ , where  $tA$  denotes the metric space  $(A, td)$ . We will sometimes also let  $0A$  denote a one-point space. The (partially defined) function  $t \mapsto |tA|$  is called the **magnitude function** of  $A$ .

**Proposition 5.2.9** ([19, Proposition 2.2.6]). *Let  $(A, d)$  be a finite metric space.*

1.  $|tA|$  is defined for all but finitely many  $t > 0$ .
2. For sufficiently large  $t$ ,  $|tA|$  is an increasing function of  $t$ .
3.  $\lim_{t \rightarrow \infty} |tA| = \#A$ .

Proposition 5.2.9 supports the interpretation of the magnitude  $|tA|$  as the “effective number of points” in  $A$ , when viewed at a scale determined by  $t$ . (We recall Solow and Polasky’s interpretation of  $|A|$  as the “effective number of species”.) However, the hypotheses of the propositions above also highlight the counterintuitive behaviors that magnitude may exhibit. In particular, there exists a metric space  $A$  such that each of the following holds:

1.  $|tA|$  is undefined for some  $t > 0$ .
2.  $|tA|$  is decreasing for some  $t > 0$ .
3.  $|tA| < 0$  for some  $t > 0$ .
4. There exists a  $B \subseteq A$  such that  $|tB| > |tA|$  for some  $t > 0$ .

We need not look that hard to find such an ill-behaved space: the complete bipartite graph  $K_{3,2}$ , equipped with the shortest path metric, has all these unpleasant properties; see Example 2.2.7 of [19]. In the next section we will consider a class of spaces which avoids most of these pathologies.

We end this section by noting that the issue of scale can be dealt with in a more elegant way if  $A$  is the vertex set of a graph and  $d$  is the shortest path metric, or more generally, whenever  $d$  is integer-valued. By (5.1), in this situation  $|tA|$  is a rational function of  $q = e^{-t}$ . More directly, if one restricts attention to such spaces, the semiring  $k$  in



the previous section can be taken to be the ring  $\mathbb{Q}(q)$  of rational functions in a formal variable  $q$ . Then the matrix  $Z_A \in (\mathbb{Q}(q))^{A \times A}$  is always invertible, so the magnitude  $|A|$  is always defined as an element of  $\mathbb{Q}(q)$ ; see section 2 of [20].

### 5.2.5 Positive definite metric spaces

As noted in section 5.2.1, a positive definite matrix  $Z$  always has magnitude, given by Proposition 5.2.1. We will now explore the consequences of this observation for magnitude of metric spaces.

A finite metric space  $(A, d)$  is said to be **positive definite** if the associated matrix  $Z_A$  is positive definite, and is said to be of **negative type** if  $Z_{tA}$  is positive semidefinite for every  $t > 0$ . It can be shown [27, Theorem 3.3] that if  $(A, d)$  is of negative type, then in fact  $Z_{tA}$  is positive definite, and hence  $tA$  is a positive definite space. A general metric space is said to be positive definite or of negative type, respectively, if every finite subspace is.

The strange turn of terminology here is due to the negative sign in  $e^{-d}$ . Negative type has several other equivalent formulations, and is an important property in the theory of metric embeddings (see, e.g., [4, 6, 39]). The fact that negative type appears naturally when considering magnitude is a hint that magnitude does in fact connect with more classical topics in geometry.

The following result is an immediate consequence of Proposition 5.2.1 and the definition of magnitude.

**Proposition 5.2.10** ([19, Proposition 2.4.3]). *If  $A$  is a finite positive definite metric space, then the magnitude  $|A|$  is defined, and*

$$|A| = \max_{0 \neq x \in \mathbb{R}^A} \frac{(\sum_{a \in A} x_a)^2}{x^T Z_A x},$$

*and the supremum is attained exactly when  $x$  is a scalar multiple of the unique weighting on  $A$ .*

A first application of Proposition 5.2.10 is Proposition 5.2.6, which is proved by showing that for large enough  $t$ ,  $Z_{tA}$  is positive definite.

**Corollary 5.2.11** (Corollaries 2.4.4 and 2.4.5 of [19]). *If  $A$  is a finite positive definite metric space and  $\emptyset \neq B \subseteq A$ , then  $1 \leq |B| \leq |A|$ .*

Proposition 5.2.10 will also be one of our main tools in the extension of magnitude to compact spaces in section 5.3.

Proposition 5.2.10 and its consequences would be of little interest without a large supply of interesting examples of positive definite spaces. Many are collected in the following result; we refer to [27, Theorem 3.6] for references and further examples.

**Theorem 5.2.12.** *The following metric spaces are of negative type, and thus magnitude is defined for all their finite subsets.*

1.  $\ell_p^n$ , the set  $\mathbb{R}^n$  equipped with the metric derived from the  $\ell_p$ -norm, for  $n \geq 1$  and  $1 \leq p \leq 2$ ;
2. Lebesgue space  $L_p[0, 1]$ , for  $1 \leq p \leq 2$ ;
3. round spheres (with the geodesic distance);
4. real and complex hyperbolic space;
5. ultrametric spaces;
6. weighted trees.

Furthermore, some natural operations on positive definite spaces yield new positive definite spaces.

**Proposition 5.2.13** ([19, Lemma 2.4.2]).

1. Every subspace of a positive definite metric space is positive definite.
2. If  $A$  and  $B$  are positive definite metric spaces, then  $A \times_1 B$  is positive definite.

On the other hand, many spaces of geometric interest are *not* of negative type, and many natural operations fail to preserve positive definiteness; see [27, Section 3.2] for examples and references.

### 5.3 Compact metric spaces

Despite strong and growing interest in the geometry of finite metric spaces (see e.g. [25]), it is natural to try to define an invariant of metric spaces, like magnitude, more generally. The most obvious context is that of compact spaces. The general definition of the magnitude of an enriched category does not help us here; nonetheless, several strategies present themselves, including approximating a compact space by finite subspaces and generalizing the notion of a weighting to compact spaces. In section 5.3.1 we will see that there is a canonical (hence “correct”) extension of magnitude from finite metric spaces to compact positive definite spaces, which can be formulated in several ways. In section 5.3.2 we will investigate a generalization of weightings to compact spaces, and see that this approach to defining magnitude agrees with the former one. This approach is of more limited scope, but often gives the easiest approach to computing magnitude; using it, we will see that magnitude knows about at least some intrinsic volumes of certain Riemannian manifolds. Finally, section 5.3.3 will introduce another invariant, maximum diversity, which is closely related to magnitude, and will be a crucial tool in proving the connection between magnitude and Minkowski dimension.

### 5.3.1 Compact positive definite spaces

To justify the “correctness” of our definition of magnitude for compact positive definite spaces, we need a topology on the family of (isometry classes of) compact metric spaces. Recall that the **Hausdorff metric**  $d_H$  on the family of compact subsets of a metric space  $X$  is given by

$$d_H(A, B) = \max \left\{ \sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A) \right\}.$$

The **Gromov–Hausdorff distance** between two compact metric spaces  $A$  and  $B$  is

$$d_{GH}(A, B) = \inf d_H(\phi(A), \psi(B)),$$

where the infimum is over all metric spaces  $X$  and isometric embeddings  $\phi : A \rightarrow X$  and  $\psi : B \rightarrow X$ . This defines a metric on the family of isometry classes of compact metric spaces; see [7, Chapter 3].

The following result follows from the proof of [27, Theorem 2.6], although our definitions are organized rather differently in that paper. We give a more streamlined version of the argument from [27].

**Proposition 5.3.1.** *The quantity*

$$M(A) = \sup \left\{ |A'| \mid A' \subseteq A, A' \text{ finite} \right\} \quad (5.2)$$

*is lower semicontinuous as a function of  $A$  (taking values in  $[0, \infty]$ ), on the class of compact positive definite metric spaces equipped with the Gromov–Hausdorff topology.*

*Proof.* Suppose first that  $d_{GH}(A, B) < \delta$  for finite positive definite spaces  $A$  and  $B$ , and let  $w \in \mathbb{R}^A$  be a weighting for  $A$ . There is a function  $f : A \rightarrow B$  such that  $|d(f(a), f(a')) - d(a, a')| < 2\delta$  for all  $a, a' \in A$ . Define  $v \in \mathbb{R}^B$  by  $v_b = \sum_{a \in f^{-1}(b)} w_a$ , and  $Z_f \in \mathbb{R}^{A \times A}$  by  $Z_f(a, a') = e^{-d(f(a), f(a'))}$ . Then  $v^T Z_B v = w^T Z_f w$ , and so

$$\left| w^T Z_A w - v^T Z_B v \right| = \left| w^T (Z_A - Z_f) w \right| \leq \|w\|_1^2 \|Z_A - Z_f\|_\infty < 2 \|w\|_1^2 \delta.$$

Thus by Proposition 5.2.10,

$$|B| \geq \frac{(\sum_b v_b)^2}{v^T Z_B v} \geq \frac{(\sum_a w_a)^2}{w^T Z_A w + 2 \|w\|_1^2 \delta} = \frac{|A|^2}{|A| + 2 \|w\|_1^2 \delta} \geq |A| - 2 \|w\|_1^2 \delta. \quad (5.3)$$

Now for general  $A$ , assume for simplicity that  $M(A) < \infty$  (the case  $M(A) = \infty$  is handled similarly). Given  $\varepsilon > 0$ , pick a finite subset  $A' \subset A$  such that  $|A'| \geq M(A) - \varepsilon$ , and let  $w \in \mathbb{R}^{A'}$  be a weighting for  $A'$ . If  $d_{GH}(A, B) < \delta$ , then there is a finite subset  $B' \subseteq B$  such that  $d_{GH}(A', B') < \delta$ , and so by (5.3),

$$M(B) \geq |B'| \geq |A'| - 2 \|w\|_1^2 \delta \geq M(A) - \varepsilon - 2 \|w\|_1^2 \delta.$$

Therefore  $M(B) \geq M(A) - 2\varepsilon$  when  $d_{GH}(A, B)$  is sufficiently small.  $\square$

Corollary 5.2.11 implies that  $M(A) = |A|$  when  $A$  itself is finite and positive definite. Proposition 5.3.1 thus implies first of all that magnitude is lower semicontinuous (l.s.c.) on the class of finite positive definite metric spaces. It follows that there is a canonical extension of magnitude to the class of compact positive definite metric spaces, namely, the *maximal* l.s.c. extension. Proposition 5.3.1 furthermore implies that this extension is precisely the function  $M$  in (5.2). For a compact positive definite metric space  $(A, d)$ , we therefore define the **magnitude**  $|A|$  to be the value of the supremum  $M(A)$  in (5.2).

Thus magnitude is lower semicontinuous on the class of compact positive spaces. This cannot be improved to continuity in general, even for the class of finite spaces of negative type. Examples 2.2.8 and 2.4.9 in [19] discuss a space  $A$  of negative type with six points, such that  $|tA| = 6/(1 + 4e^{-t})$ ; thus  $\lim_{t \rightarrow 0^+} |tA| = 6/5$ , whereas the space  $tA$  itself converges to a one-point space. On the other hand, magnitude is continuous when restricted to certain classes of spaces, as we will see in Corollary 5.3.13 and Theorem 5.4.15 below.

Proposition 5.3.1, Proposition 5.2.8, and Corollary 5.2.11 yield the following results.

**Proposition 5.3.2** ([19, Lemma 3.1.3]). *If  $A$  is a compact positive definite metric space and  $\emptyset \neq B \subseteq A$ , then  $1 \leq |B| \leq |A|$ .*

**Proposition 5.3.3** ([27, Corollary 2.7]). *Let  $A$  be a compact positive definite metric space, and let  $\{A_k\}$  be any sequence of compact subsets of  $A$  such that  $A_k \xrightarrow{k \rightarrow \infty} A$  in the Hausdorff topology. Then  $|A| = \lim_{k \rightarrow \infty} |A_k|$ .*

**Proposition 5.3.4** ([19, Proposition 3.1.4]). *If  $A$  and  $B$  are compact positive definite metric spaces, then  $|A \times_1 B| = |A| |B|$ .*

Proposition 5.3.1 justifies the above definition of magnitude as the “correct” one for a compact positive definite space  $A$ . Nevertheless, for both aesthetic and practical reasons, it is desirable to be able to work directly with  $A$  itself, as opposed to approximations of  $A$  by finite subspaces. Two different, more direct approaches to defining magnitude for compact positive definite spaces were developed in [27, 28]. In essence, these papers introduced two different topologies on the space  $\{w \in \mathbb{R}^A \mid \text{supp } w \text{ is finite}\}$ . The topology used in [27] has the advantage of being more familiar, whereas the topology in [28] has the advantage of being better suited to the analysis of magnitude. In particular, the topology used in [28] can be dualized in a way that presents a new set of tools to study magnitude. In the pursuit of our goal of proceeding as quickly as possible to geometric results, here we will go straight to the dual version.

Recall that a **positive definite kernel** on a space  $X$  is a function  $K : X \times X \rightarrow \mathbb{C}$  such that, for every finite set  $A \subseteq X$ , the matrix  $[K(a, b)]_{a, b \in A} \in \mathbb{C}^{A \times A}$  is positive definite. Given a positive definite kernel on  $X$ , the **reproducing kernel Hilbert space**

(RKHS)  $\mathcal{H}$  on  $X$  with kernel  $K$  is the completion of the linear span of the functions  $k_x(y) = K(x, y)$  with respect to the inner product given by

$$\langle k_x, k_y \rangle_{\mathcal{H}} = K(x, y)$$

(see [2]). If  $f \in \mathcal{H}$ , then  $f(x) = \langle f, k_x \rangle_{\mathcal{H}}$  for every  $x \in X$ , and consequently

$$|f(x)| \leq \|f\|_{\mathcal{H}} \|k_x\|_{\mathcal{H}} = \|f\|_{\mathcal{H}} \sqrt{K(x, x)} \quad (5.4)$$

by the Cauchy–Schwarz inequality.

Now if  $(X, d)$  is a positive definite metric space, then  $K(x, y) = e^{-d(x, y)}$  is a positive definite kernel on  $X$ . We will refer to the corresponding RKHS as the RKHS  $\mathcal{H}$  for  $X$ .

**Theorem 5.3.5** ([28, Theorem 4.1 and Proposition 4.2]). *Let  $X$  be a positive definite metric space, and let  $A \subseteq X$  be compact. Then  $|A| < \infty$  if and only if there exists a function  $h \in \mathcal{H}$  such that  $h \equiv 1$  on  $A$ . In that case,*

$$|A| = \inf \left\{ \|h\|_{\mathcal{H}}^2 \mid h \in \mathcal{H}, h \equiv 1 \text{ on } A \right\}.$$

*The infimum is achieved for a unique function  $h$ . If  $f \in \mathcal{H}$  also satisfies  $f \equiv 1$  on  $A$ , then  $|A| = \langle f, h \rangle_{\mathcal{H}}$ .*

*Proof.* First observe that if  $w \in \mathbb{R}^B$  for a finite subset  $B \subseteq X$ , and  $f_w = \sum_{b \in B} w_b e^{-d(\cdot, b)}$ , then

$$w^T Z_B w = \sum_{a, b \in B} w_a e^{-d(a, b)} w_b = \|f_w\|_{\mathcal{H}}^2. \quad (5.5)$$

Now suppose that  $|A| < \infty$ . If  $B \subseteq A$  is finite and  $w \in \mathbb{R}^B$ , then by Proposition 5.2.10, (5.5), and the definition of  $|A|$ ,

$$\left( \sum_{b \in B} w_b \right)^2 \leq |B| \|f_w\|_{\mathcal{H}}^2 \leq |A| \|f_w\|_{\mathcal{H}}^2.$$

Thus the linear functional  $f_w \mapsto \sum_{b \in B} w_b$  on the subspace  $\{f_w \mid w \in \mathbb{R}^B, B \subseteq A \text{ finite}\} \subseteq \mathcal{H}$  has norm at most  $\sqrt{|A|}$ . Therefore there is a function  $h \in \mathcal{H}$  with  $\|h\|_{\mathcal{H}}^2 = |A|$  such that

$$\sum_{b \in B} w_b = \langle f_w, h \rangle_{\mathcal{H}} = \sum_{b \in B} w_b h(b)$$

for every  $f_w$ ; taking  $f_w = e^{-d(\cdot, a)}$  for  $a \in A$  yields  $h(a) = 1$ .

Next suppose that there exists an  $h \in \mathcal{H}$  such that  $h \equiv 1$  on  $A$ . Then for any finite subset  $B \subseteq A$  and  $w \in \mathbb{R}^B$ , by the Cauchy–Schwarz inequality,

$$\left| \sum_{b \in B} w_b \right| = |\langle h, f_w \rangle| \leq \|h\|_{\mathcal{H}} \|f_w\|_{\mathcal{H}}.$$

Equation (5.5) and Proposition 5.2.10 then imply that  $|B| \leq \|h\|_{\mathcal{H}}^2$ , and so by definition  $|A| \leq \|h\|_{\mathcal{H}}^2$ .

The above arguments prove both the “if and only if” statement and the infimum expression for  $|A|$ . The last two statements follow from elementary Hilbert space geometry.  $\square$

We will call the unique function  $h$  which achieves the infimum in Theorem 5.3.5 the **potential function** of  $A$ . Theorem 5.3.5 will prove its worth in sections 5.4.3 and 5.4.4 below.

For now, we consider what has happened to weightings, which were central to the original category-inspired definition of magnitude, but have vanished from the scene in Theorem 5.3.5. Weightings of finite subspaces of  $X$  are naturally identified with elements of the dual space  $\mathcal{H}^*$ , if we restrain ourselves from the usual impulse to identify  $\mathcal{H}^*$  with  $\mathcal{H}$  itself. We can then identify a weighting of a compact subspace  $A$  with finite magnitude as an element of  $\mathcal{H}^*$ , specifically the element of  $\mathcal{H}^*$  represented by the potential function  $h$ . See [28] for details.

### 5.3.2 Weight measures

Proposition 5.3.1 may justify the definition of magnitude adopted in the previous section as the canonical correct definition, but it has two deficiencies. First, it applies only to positive definite spaces, and second, it lies quite far from the original category-inspired definition, being fundamentally based instead on the reformulation in Proposition 5.2.10. The second drawback is to some extent addressed in the last paragraph of the previous section, though still only for positive definite spaces.

In this section we discuss another approach to defining magnitude for compact metric spaces, first used in [41], which more closely parallels the original definition for finite spaces.

A **weight measure** on a compact metric space  $(A, d)$  is a finite signed Borel measure  $\mu$  on  $A$  such that

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

for every  $a \in A$ .

A finite metric space  $A$  possesses a weight measure  $\mu$  if and only if it possesses a weighting  $w \in \mathbb{R}^A$ , with the correspondence given by  $w_a = \mu(\{a\})$ . The magnitude of  $A$  is in that case

$$|A| = \sum_{a \in A} w_a = \mu(A).$$

This suggests defining the magnitude of a compact metric space to be  $|A| = \mu(A)$  whenever  $A$  possesses a weight measure  $\mu$ . The following result shows that doing so agrees with the definition adopted in the previous section, whenever both definitions apply.

**Proposition 5.3.6** ([27, Theorem 2.3]). *Suppose that  $A$  is a compact positive definite metric space with weight measure  $\mu$ . Then  $|A| = \mu(A)$ .*

*Proof.* For any finite signed measure  $\mu$  on  $A$  and  $f \in \mathcal{H}$ ,

$$\left| \int f \, d\mu \right| \leq \|f\|_\infty \|\mu\|_{TV} \leq \|f\|_{\mathcal{H}} \|\mu\|_{TV}$$

by (5.4) (since  $K(x, x) = 1$  here), where  $\|\mu\|_{TV}$  denotes the total variation norm of  $\mu$ . Therefore  $f \mapsto \int f \, d\mu$  is a bounded linear functional on  $\mathcal{H}$ , represented by some  $g \in \mathcal{H}$ . So for each  $a \in A$ ,

$$1 = \int e^{-d(a,b)} \, d\mu(b) = \left\langle e^{-d(\cdot, b)}, g \right\rangle_{\mathcal{H}} = g(a).$$

Then by the last statement of Theorem 5.3.5, if  $h$  is the potential function of  $A$ , then

$$|A| = \langle g, h \rangle_{\mathcal{H}} = \int h \, d\mu = \mu(A). \quad \square$$

In fact it can be shown that  $g = h$  in the proof above.

We therefore define the **magnitude** of a compact metric space  $A$  with a weight measure  $\mu$  to be  $|A| := \mu(A)$ , with Proposition 5.3.6's assurance that when  $A$  is positive definite, this definition is consistent with the previous one.

A first nontrivial example is a compact interval  $[a, b] \subseteq \mathbb{R}$ . A straightforward computation (see [41, Theorem 2]) shows that

$$\mu_{[a,b]} = \frac{1}{2}(\delta_a + \lambda_{[a,b]} + \delta_b) \quad (5.6)$$

is a weight measure for  $[a, b]$ , where  $\delta_x$  denotes the point mass at  $x$  and  $\lambda_{[a,b]}$  denotes Lebesgue measure restricted to  $[a, b]$ . It follows that

$$|[a, b]| = 1 + \frac{b-a}{2}. \quad (5.7)$$

See [32] for a contention that (up to the  $\frac{1}{2}$  scaling factor) this is the “correct” size of an interval. In any case, the appearance of the length  $(b-a)$  gives the first compelling evidence that magnitude knows about genuinely “geometric” information for infinite spaces.

The following easy consequence of Fubini's theorem further extends the reach of Propositions 5.2.8 and 5.3.4.

**Proposition 5.3.7.** *If  $\mu_A$  and  $\mu_B$  are weight measures on compact metric spaces  $A$  and  $B$ , then  $\mu_A \otimes \mu_B$  is a weight measure on  $A \times_1 B$ , and so  $|A \times_1 B| = |A| |B|$ .*

The chief drawback to the definition of magnitude in terms of weight measures is that many interesting spaces do not possess weight measures. For example, the results of

[3] imply that balls in  $\ell_2^3$  do not possess weight measures (rather, their weightings turn out to be higher-order distributions), and numerical computations in [40] suggest that squares and discs in  $\ell_2^2$  also do not possess weight measures.

On the other hand, the following result can be interpreted as saying that compact positive definite spaces “almost” possess weight measures.

**Proposition 5.3.8** ([27, Theorems 2.3 and 2.4]). *If  $A$  is a compact positive definite metric space, then*

$$|A| = \sup \left\{ \frac{\mu(A)^2}{\int_A \int_A e^{-d(a,b)} d\mu(a) d\mu(b)} \mid \mu \in M(A), \int_A \int_A e^{-d(a,b)} d\mu(a) d\mu(b) \neq 0 \right\},$$

where  $M(A)$  denotes the space of finite signed Borel measures on  $A$ . The supremum is attained if and only if  $A$  possesses a weight measure; in that case it is attained precisely by scalar multiples of weight measures.

One positive result about the existence of weight measures is the following.

**Proposition 5.3.9** ([27, Lemma 2.8 and Corollary 2.10]). *Suppose  $(A, d)$  is a compact positive definite space, and that each finite  $A' \subseteq A$  possesses a weighting with positive components. Then  $A$  possesses a positive weight measure.*

The hypothesis of Proposition 5.3.9 is satisfied, for example, by all compact subsets of  $\mathbb{R}$  and by all compact ultrametric spaces (see Theorem 5.4.1 below and [19, Proposition 2.4.18]). Since Proposition 5.3.9 applies only to positive definite spaces, it does not extend the scope of magnitude beyond that of the previous section. Nevertheless, the existence of a positive weight measure makes it much easier to compute magnitude, and has other theoretical consequences which will come up in the next section.

The following generalization of Proposition 5.2.7 gives another large class of spaces which possess weight measures.

**Lemma 5.3.10** ([41, Theorem 1]). *Let  $A$  be a compact homogeneous metric space. Then  $A$  possesses a weight measure, which is a scalar multiple of the unique isometry-invariant probability measure  $\mu$  on  $A$ . Furthermore,*

$$|A| = \left( \int_A \int_A e^{-d(a,b)} d\mu(a) d\mu(b) \right)^{-1}.$$

Using Lemma 5.3.10, Willerton explicitly computed the magnitudes of round spheres with the geodesic metric: for  $n$  even, the magnitude of the  $n$ -sphere with radius  $R$  is

$$\frac{2}{1 + e^{-\pi R}} \left[ 1 + \left( \frac{R}{1} \right)^2 \right] \left[ 1 + \left( \frac{R}{3} \right)^2 \right] \cdots \left[ 1 + \left( \frac{R}{n-1} \right)^2 \right],$$



and there is a similar formula for odd  $n$ ; see [41, Theorem 7].

Lemma 5.3.10 is particularly useful in analyzing the magnitude function of a homogeneous space  $A$ , since it implies that  $tA$  possesses a weight measure for every  $t > 0$ , which is moreover independent of  $t$  (up to normalization). In the particular case of a homogeneous Riemannian manifold, Willerton proved the following asymptotic results. (We note that most homogeneous manifolds are *not* of negative type, so that  $tA$  need not be positive definite; see [12].)

**Theorem 5.3.11** ([41, Theorem 11]). *Suppose that  $(M, d)$  is an  $n$ -dimensional homogeneous Riemannian manifold equipped with its geodesic distance  $d$ . Then*

$$|tM| = \frac{1}{n! \omega_n} \left( \text{vol}(M) t^n + \frac{n+1}{6} \text{tsc}(M) t^{n-2} + O(t^{n-4}) \right) \quad \text{as } t \rightarrow \infty,$$

where  $\text{vol}$  denotes Riemannian volume,  $\text{tsc}$  denotes total scalar curvature, and  $\omega_n$  is the volume of the  $n$ -dimensional unit ball in  $\ell_2^n$ .

In particular, if  $M$  is a homogeneous Riemannian surface, then

$$|tM| = \frac{\text{area}(M)}{2\pi} t^2 + \chi(M) + O(t^{-2}) \quad \text{as } t \rightarrow \infty,$$

where  $\chi(M)$  denotes the Euler characteristic of  $M$ .

Theorem 5.3.11 shows in particular that the magnitude function of a homogeneous Riemannian manifold determines both its volume and its total scalar curvature.

We note that most Riemannian manifolds are neither homogeneous nor positive definite, and it is so far not clear how to define their magnitude.

### 5.3.3 Maximum diversity

Proposition 5.3.8 suggests considering, for a compact metric space  $(A, d)$ , the quantity

$$\begin{aligned} |A|_+ &:= \sup \left\{ \frac{\mu(A)^2}{\int_A \int_A e^{-d(a,b)} d\mu(a) d\mu(b)} \mid \mu \in M_+(A), \mu \neq 0 \right\} \\ &= \sup_{\mu \in P(A)} \left( \int_A \int_A e^{-d(a,b)} d\mu(a) d\mu(b) \right)^{-1}, \end{aligned} \quad (5.8)$$

where  $M_+(A)$  is the space of finite positive Borel measures on  $A$ , and  $P(A)$  is the space of Borel probability measures on  $A$ . We refer to  $|A|_+$  as the **maximum diversity** of  $A$ , for reasons that will be described shortly. Maximum diversity lacks the category-theoretic motivation of magnitude, but it turns out to have its own interesting interpretations, and to be both intimately related to magnitude and easier to analyze in certain respects.

Regarding interpretation, suppose that  $A$  is finite, the points of  $A$  represent species in some ecosystem, and that  $e^{-d(a,b)} \in (0, 1]$  represents the “similarity” of two species  $a, b \in A$ . If  $\mu \in P(A)$  gives the relative abundances of species, then

$$\left( \int_A \int_A e^{-d(a,b)} d\mu(a) d\mu(b) \right)^{-1}$$

gives a way of quantifying the “diversity” of the ecosystem which is sensitive to both the abundances of the species and the similarities between them; see [21] for extensive discussion of a much larger family of diversities that this fits into. It is this interpretation that motivates the name “maximum diversity”.

There are multiple connections between magnitude and maximum diversity. The most obvious is that, by Proposition 5.3.8,  $|A|_+ \leq |A|$  for any compact positive definite space  $A$ . Moreover, Proposition 5.3.8 implies that  $|A|_+ = |A|$  if  $A$  is positive definite and possesses a positive weight measure; Proposition 5.3.9 and Lemma 5.3.10 indicate some families of such spaces. Finally, as we will see in Corollary 5.4.23 below, if  $A \subseteq \ell_2^n$ , then the inequality  $|A|_+ \leq |A|$  can be reversed, up to a (dimension-dependent) multiplicative constant. We will see applications of all these connections below.

A more subtle connection between maximum diversity and magnitude, which we will not discuss here, is proved in the main result of [17, 22].

We now move on to ways in which maximum diversity is better behaved than magnitude. One is that the supremum in (5.8) is always achieved, unlike the one in Proposition 5.3.8. This is a consequence of the compactness of  $P(A)$  in the weak- $*$  topology; see [27, Proposition 2.9] (this fact is used in the proof of Proposition 5.3.9 above). Another is the following improvement, for maximum diversity, of Proposition 5.3.1.

**Proposition 5.3.12** ([27, Proposition 2.11]). *The maximum diversity  $|A|_+$  is continuous as a function of  $A$ , on the class of compact metric spaces equipped with the Gromov–Hausdorff topology.*

**Corollary 5.3.13** ([27, Corollary 2.12]). *The magnitude  $|A|$  is continuous as a function of  $A$ , on the class of compact positive definite metric spaces which possess positive weight measures, equipped with the Gromov–Hausdorff topology.*

*In particular, magnitude is continuous on the class of compact subsets of  $\mathbb{R}$ , and on the class of compact ultrametric spaces.*

The next result shows how the asymptotic behavior of  $|tA|_+$  is relatively easy to analyze. Recall that the **covering number**  $N(A, \varepsilon)$  is the minimum number of  $\varepsilon$ -balls required to cover  $A$ , and that the **Minkowski dimension** of  $A$  may be defined as

$$\dim_{\text{Mink}} A := \lim_{\varepsilon \rightarrow 0^+} \frac{N(A, \varepsilon)}{\log(1/\varepsilon)} \quad (5.9)$$

whenever this limit exists. The idea of the proof of Proposition 5.3.14 below is simply that when  $t$  is large and  $\varepsilon$  is small, the supremum over  $P(A)$  defining  $|tA|_+$  is approximately attained by a measure uniformly supported on the centers of a maximal family of disjoint  $\varepsilon$ -balls in  $A$ .

**Proposition 5.3.14** ([28, Theorem 7.1]). *If  $A$  is a compact metric space, then*

$$\lim_{t \rightarrow \infty} \frac{\log |tA|_+}{\log t} = \dim_{\text{Mink}} A. \quad (5.10)$$

Proposition 5.3.14 should be interpreted as saying that the limit on the left hand side of (5.10) exists if and only if  $\dim_{\text{Mink}} A$  exists. Moreover, if the limit is replaced with a  $\liminf$  or  $\limsup$ , the left hand side of (5.10) is equal to the so-called lower or upper Minkowski dimension of  $A$ , respectively, defined by modifying (5.9) in the same way.

Since  $|A|_+ \leq |A|$  for any compact positive definite space, Proposition 5.3.14 gives a lower bound for the growth rate of the magnitude function for a compact space of negative type. Moreover, in Euclidean space  $\ell_2^n$ , Proposition 3.12 and the rough equivalence of magnitude and maximum diversity mentioned above will be used to show that Minkowski dimension can be recovered from magnitude; see Theorem 5.4.24 below. (Proposition 7.5 of [28] proves the same fact for compact homogeneous metric spaces, using Lemma 5.3.10 above.)

## 5.4 Magnitude in normed spaces

In this section we will specialize magnitude to compact subsets of finite-dimensional vector spaces with translation-invariant metrics. It is in these settings that we find the strongest connections between magnitude and geometry. In section 5.4.1, we find a quite complete description of the magnitude of an arbitrary compact set  $A \subseteq \mathbb{R}$ ; in particular,  $|A|$  depends only on the Lebesgue measure of  $A$  and the sizes of the “gaps” in  $A$  (Corollary 5.4.3). In section 5.4.2, we show that in  $\ell_1^n$ , magnitude can be used to recover  $\ell_1$  analogues of the classical intrinsic volumes of a convex body (Theorem 5.4.6). In section 5.4.3, we apply Fourier analysis to the study of magnitude, when  $\mathbb{R}^n$  is equipped with a norm (or more generally, a  $p$ -norm) which makes it a positive definite metric space. In particular, we find that magnitude is continuous on convex bodies in such spaces (Theorem 5.4.15). Finally, in section 5.4.4, we specialize these tools to the most familiar normed space, the Euclidean space  $\ell_2^n$ . In that setting the Fourier-analytic perspective of section 5.4.3 uncovers connections with partial differential equations and potential theory. Among other results, we will see that in Euclidean space, magnitude knows about volume (Theorem 5.4.14) and Minkowski dimension (Theorem 5.4.24), although there are frustratingly few compact sets in  $\ell_2^n$  whose exact magnitudes are known (see Theorem 5.4.21).

Corollary 5.4.3 and the material of section 5.4.2 are new. Most of the results of section 5.4.3 were previously proved for Euclidean space, but are new in the generality discussed here.

#### 5.4.1 Magnitude in $\mathbb{R}$

In the real line  $\mathbb{R}$ , magnitude can be analyzed in great detail thanks to the order structure underlying the metric structure. Namely, if  $a < b < c$ , then  $Z(a, c) = Z(a, b)Z(b, c)$ , where we recall that  $Z(a, b) = e^{-d(a, b)}$ . This simple fact lies behind the proof of the next result.

**Theorem 5.4.1** ([23, Theorem 4] and [19, Proposition 2.4.13]). *Given real numbers  $a_1 < a_2 < \dots < a_N$ , the weighting  $w$  of  $A = \{a_1, \dots, a_N\}$  is given by*

$$w_{a_i} = \frac{1}{2} \left( \tanh \frac{a_i - a_{i-1}}{2} + \tanh \frac{a_{i+1} - a_i}{2} \right)$$

for  $2 \leq i \leq N - 1$ , and

$$w_{a_1} = \frac{1}{2} \left( 1 + \tanh \frac{a_2 - a_1}{2} \right), \quad w_{a_N} = \frac{1}{2} \left( 1 + \tanh \frac{a_N - a_{N-1}}{2} \right).$$

Consequently,

$$|A| = 1 + \sum_{i=2}^N \tanh \frac{a_i - a_{i-1}}{2}.$$

Theorem 5.4.1, together with Proposition 5.3.3, was used to give the first derivation of the magnitude of an interval; see [23, Theorem 7] and [19, Theorem 3.2.2].

As mentioned above, by Proposition 5.3.9, Theorem 5.4.1 implies that every compact subset of  $\mathbb{R}$  possesses a weight measure. Furthermore, as noted in Corollary 5.3.13, this implies that magnitude on  $\mathbb{R}$  is continuous with respect to the Gromov–Hausdorff topology.

The last part of the following corollary appears, with additional technical assumptions, as [41, Lemma 3].

**Corollary 5.4.2.** *Suppose that  $A, B \subseteq \mathbb{R}$  are compact with  $a = \sup A \leq \inf B = b$ . Then*

$$|A \cup B| = |A| + |B| - 1 + \tanh \frac{b - a}{2}.$$

Consequently, if  $C \subseteq \mathbb{R}$  is compact and  $[a, b] \subseteq C$ , then

$$|C \setminus (a, b)| = |C| - \frac{b - a}{2} + \tanh \frac{b - a}{2}.$$

*Proof.* The first claim follows immediately from Theorem 5.4.1 in the case that  $A$  and  $B$  are finite, and then follows for general compact sets by continuity. The second equality

follows by writing  $C = A \cup [a, b] \cup B$ , where  $A = C \cap (-\infty, a]$  and  $B = C \cap [b, \infty)$ , then applying the first equality twice and (5.7).  $\square$

Corollary 5.4.2, together with continuity and the knowledge of the magnitude of a compact interval, can be used to compute the magnitude of any compact set  $A \subseteq \mathbb{R}$ , since  $A$  can be written as

$$A = [a, b] \setminus \bigcup_i (a_i, b_i), \quad (5.11)$$

where  $\{(a_i, b_i)\}$  is a finite or countable collection of disjoint subintervals of  $[a, b]$ .

**Corollary 5.4.3.** *If  $A \subseteq \mathbb{R}$  is compact, then*

$$|A| = 1 + \frac{\text{vol}_1 A}{2} + \sum_i \tanh \frac{b_i - a_i}{2},$$

where  $a_i$  and  $b_i$  are as in (5.11).

Another proof of Corollary 5.4.3 can be given using [19, Proposition 3.2.3]. As an application of Corollary 5.4.3, we obtain the magnitude of the length  $\ell$  ternary Cantor set  $C_\ell$  (see [23, Theorem 10], [41, Theorem 4]):

$$|C_\ell| = 1 + \frac{1}{2} \sum_{i=1}^{\infty} \tanh \frac{\ell}{2 \cdot 3^i}.$$

## 5.4.2 Magnitude in the $\ell_1$ -norm

The magnitude of subsets of  $\mathbb{R}^n$  is generally most tractable when we equip  $\mathbb{R}^n$  with the  $\ell_1$ -norm. Although that may not be the norm of primary geometric interest, it provides a testing ground for questions that are more difficult to settle in Euclidean space.

We have already seen that  $\ell_1^n$ , like  $\ell_2^n$ , is of negative type (Theorem 5.2.12). The key difference is Proposition 5.3.4, the multiplicativity of magnitude with respect to the  $\ell_1$  product. Since we already know the magnitude of intervals, this immediately allows us to calculate the magnitude of boxes in  $\ell_1^n$ . Unions of boxes can then be used to approximate more complex subsets, as we shall see.

Explicitly, a box  $\prod_{i=1}^n [a_i, a_i + L_i]$  in  $\ell_1^n$  has magnitude  $\prod_{i=1}^n (1 + L_i/2)$ . It follows that  $|tA| \rightarrow 1$  as  $t \rightarrow 0^+$  for boxes  $A$ . But then monotonicity of magnitude (Proposition 5.3.2) implies a more general result:

**Proposition 5.4.4.** *If  $A \subseteq \ell_1^n$  is compact, then  $\lim_{t \rightarrow 0^+} |tA| = 1$ .*

(In  $\ell_2^n$ , this is much harder to prove; see Theorem 5.4.18.) Proposition 5.4.4 and Theorem 5.4.17 together imply that the magnitude function  $t \mapsto |tA|$  is continuous on  $[0, \infty)$ .

Our formula for the magnitude of a box in  $\ell_1^n$  can be rewritten in terms of the intrinsic volumes  $V_0, V_1, \dots$  (defined in, for instance, Chapter 7 of [11] or Chapter 4 of [34]). Recall that  $V_i(A)$  is the canonical  $i$ -dimensional measure of a convex set  $A \subseteq \mathbb{R}^n$ , and that the intrinsic volumes are characterized by Steiner's polynomial formula

$$\text{vol}(A + r\mathcal{B}^n) = \sum_{i=0}^n \omega_{n-i} V_i(A) r^{n-i}$$

(Proposition 9.2.2 of [11] or Equation 4.1 of [34]), where  $\mathcal{B}^n$  is the unit Euclidean  $n$ -ball and  $\omega_j = \text{vol}(\mathcal{B}^j)$ . For boxes  $A \subseteq \ell_1^n$ , the formula above can be rewritten as

$$|A| = \sum_{i=0}^n \frac{V_i(A)}{2^i}, \quad (5.12)$$

either by direct calculation or by noting that  $|[0, L]| = 1 + V_1([0, L])/2$  and using the multiplicative property of the intrinsic volumes (Theorem 9.7.1 of [11]). Hence the magnitude function of a box  $A$  is a polynomial

$$|tA| = \sum_{i=0}^n \frac{V_i(A)}{2^i} t^i$$

whose coefficients are (up to known factors) the intrinsic volumes of  $A$ , and whose degree is its dimension. In particular, the magnitude function of a box determines all of its intrinsic volumes and its dimension.

In fact, such a result is true for a much larger class of subsets of  $\ell_1^n$  than just boxes. To show this, we must adapt the classical notion of intrinsic volume to  $\ell_1^n$ , following [18].

First recall that a metric space  $A$  is **geodesic** if for any  $a, b \in A$  there exists a distance-preserving map  $\gamma: [0, d(a, b)] \rightarrow A$  such that  $\gamma(0) = a$  and  $\gamma(d(a, b)) = b$ . The geodesic subsets of  $\ell_2^n$  are the convex sets. The geodesic subsets of  $\ell_1^n$ , called the  **$\ell_1$ -convex sets** [18], include the convex sets and much else besides (such as L shapes). In this setting, there is a Steiner-type theorem in which balls are replaced by cubes (Theorem 6.2 of [18]): for any  $\ell_1$ -convex compact set  $A \subseteq \ell_1^n$ , writing  $\mathcal{C}^n = [-1/2, 1/2]^n$ ,

$$\text{vol}(A + r\mathcal{C}^n) = \sum_{i=0}^n V'_i(A) r^{n-i} \quad (5.13)$$

where  $V'_0(A), \dots, V'_n(A)$  depend only on  $A$ .

The functions  $V'_0, V'_1, \dots$  on the class of  $\ell_1$ -convex compact sets are called the  **$\ell_1$ -intrinsic volumes** [18]. They are valuations (that is, finitely additive), continuous with respect to the Hausdorff metric, and invariant under isometries of  $\ell_1^n$ . There is a well-developed integral geometry of  $\ell_1$ -convex sets [18], closely parallel to the classical integral geometry of convex sets; for instance, there is a Hadwiger-type theorem for  $\ell_1$ -intrinsic volumes.

Although the intrinsic and  $\ell_1$ -intrinsic volumes are not in general equal, they coincide for boxes  $A$ , giving

$$|A| = \sum_{i=0}^n \frac{V'_i(A)}{2^i}, \quad |tA| = \sum_{i=0}^n \frac{V'_i(A)}{2^i} t^i \quad (5.14)$$

(the latter because  $V'_i$  is homogeneous of degree  $i$ ). It is this relationship, not (5.12), that generalizes from boxes to a much larger class of sets.

**Conjecture 5.4.5** ([19, Conjecture 3.4.10]). *For all compact  $\ell_1$ -convex sets  $A \subseteq \ell_1^n$ ,*

$$|A| = \sum_{i=0}^n \frac{V'_i(A)}{2^i}.$$

We will prove the following parts of this conjecture:

**Theorem 5.4.6.**

1.  $|A| \leq \sum_{i=0}^n 2^{-i} V'_i(A)$  for all compact  $\ell_1$ -convex sets  $A \subseteq \ell_1^n$ .
2.  $|A| = \sum_{i=0}^n 2^{-i} V'_i(A)$  for all convex bodies  $A \subseteq \ell_1^n$ .
3.  $|A| = \sum_{i=0}^2 2^{-i} V'_i(A)$  for all compact convex sets  $A \subseteq \ell_1^2$ .

(A **convex body** is a compact convex set with nonempty interior.)

For the proof, we will use some special classes of box. A **pixel** in  $\mathbb{R}^n$  is a unit cube  $\prod_{i=1}^n [a_i, a_i + 1]$  with integer coordinates  $a_i$ . More generally, a **subpixel** is a box  $\prod_{i=1}^n [a_i, b_i]$  with  $a_i \in \mathbb{Z}$  and  $b_i \in \{a_i, a_i + 1\}$ . Note that the intersection of two subpixels is either a subpixel or empty.

Equation (5.6) and Proposition 5.3.7 imply that for any box  $B = \prod_i [a_i, b_i]$  in  $\ell_1^n$ , the product measure  $\mu_B = \prod_{i=1}^n \mu_{[a_i, b_i]}$  is a weight measure on  $B$ .

**Lemma 5.4.7.** *There is a unique function*

$$\begin{array}{ccc} \{\text{finite unions of subpixels in } \mathbb{R}^n\} & \rightarrow & \{\text{signed Borel measures on } \mathbb{R}^n\} \\ A & \mapsto & \mu_A \end{array}$$

*extending the definition above for subpixels and satisfying  $\text{supp } \mu_A \subseteq A$ ,  $\mu_\emptyset = 0$ , and  $\mu_{A \cup B} = \mu_A + \mu_B - \mu_{A \cap B}$  whenever  $A$  and  $B$  are finite unions of subpixels.*

*Proof.* By the extension theorem of Groemer (Theorem 6.2.1 of [34]), it suffices to show that for any subpixels  $B_1, \dots, B_m$  such that  $B_1 \cup \dots \cup B_m$  is a subpixel,

$$\mu_{B_1 \cup \dots \cup B_m} = \sum_{k \geq 0} (-1)^k \sum_{1 \leq j_0 < \dots < j_k \leq m} \mu_{B_{j_0} \cap \dots \cap B_{j_k}}.$$

But  $B_1 \cup \dots \cup B_m$  is only a subpixel if some  $B_j$  contains all the others, and in that case the sum telescopes and the proof is trivial.  $\square$

A subset  $A$  of  $\ell_1^n$  is **1-pixelated** if it is a finite union of pixels; then  $\lambda A$  is said to be  **$\lambda$ -pixelated**. A set is **pixelated** if it is  $\lambda$ -pixelated for some  $\lambda > 0$ .

**Proposition 5.4.8.** *Let  $A \subseteq \ell_1^n$  be an  $\ell_1$ -convex pixelated set. Then  $\mu_A$  as given in Lemma 5.4.7 is a weight measure on  $A$ .*

*Proof.* We may harmlessly assume that  $A$  is 1-pixelated. The result holds when either  $n = 0$  or  $A$  is a single pixel. So, we may assume inductively that  $n \geq 1$ , that  $A$  contains at least two pixels, and that the result holds for  $\ell_1$ -convex 1-pixelated sets of smaller dimension or fewer pixels than  $A$ .

Fix  $a \in A$ . We may assume without loss of generality that at least two of the pixels in  $A$  differ in their last coordinates, that  $\sup_{b \in A} b_n = 1$ , and that  $a$  belongs to some pixel of  $A$  whose center has negative last coordinate. Write  $A_-$  for the union of the pixels in  $A$  whose centers have negative last coordinates, and similarly  $A_+$ . Thus,  $a \in A_-$  and the center of every pixel in  $A_+$  has last coordinate  $1/2$ . Both  $A_-$  and  $A_+$  are  $\ell_1$ -convex 1-pixelated sets (by Lemma 3.3 of [18]), and  $A_- \cap A_+$  is a finite union of subpixels (though need not be pixelated).

We have to show that

$$\int_{\mathbb{R}^n} Z(a, b) d\mu_A(b) = 1.$$

Since  $\mu_A = \mu_{A_+} + \mu_{A_-} - \mu_{A_- \cap A_+}$  and  $\mu_{A_-}$  is a weight measure on  $A_-$  (by inductive hypothesis), an equivalent statement is that

$$\int_{\mathbb{R}^n} Z(a, b) d\mu_{A_+}(b) = \int_{\mathbb{R}^n} Z(a, b) d\mu_{A_- \cap A_+}(b). \quad (5.15)$$

Write  $\pi: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$  for orthogonal projection onto the first  $(n-1)$  coordinates, and write  $a' = (\pi(a), 0) = (a_1, \dots, a_{n-1}, 0)$ . Then  $Z(a, b) = Z(a, a')Z(a', b)$  for  $b \in A_+$ , so (5.15) is equivalent to

$$\int_{\mathbb{R}^n} Z(a', b) d\mu_{A_+}(b) = \int_{\mathbb{R}^n} Z(a', b) d\mu_{A_- \cap A_+}(b).$$

We analyze each side in turn. First,  $A_+ = (\pi A_+) \times [0, 1]$ , so it follows from Proposition 5.3.7 that  $\mu_{A_+} = \mu_{\pi A_+} \otimes \mu_{[0,1]}$ . Using this and the fact that  $\mu_{[0,1]}$  is a weight measure on  $[0, 1]$ , we find that the left-hand side is equal to

$$\int_{\mathbb{R}^{n-1}} Z(\pi(a), c) d\mu_{\pi A_+}(c). \quad (5.16)$$

Next,  $\mu_{A_- \cap A_+}$  is supported on  $\mathbb{R}^{n-1} \times \{0\}$ , and  $\pi(A_- \cap A_+) = \pi A_- \cap \pi A_+$  (by Corollary 2.5 of [18]), which together imply that the right-hand side is equal to

$$\int_{\mathbb{R}^{n-1}} Z(\pi(a), c) d\mu_{\pi A_- \cap \pi A_+}(c). \quad (5.17)$$



Hence it suffices to show that the integrals (5.16) and (5.17) are equal. Since  $\mu_{\pi A} = \mu_{\pi A_-} + \mu_{\pi A_+} - \mu_{\pi A_- \cap \pi A_+}$ , an equivalent statement is that

$$\int_{\mathbb{R}^{n-1}} Z(\pi(a), c) d\mu_{\pi A}(c) = \int_{\mathbb{R}^{n-1}} Z(\pi(a), c) d\mu_{\pi A_-}(c). \quad (5.18)$$

But  $\pi A$  and  $\pi A_-$  are 1-pixelated sets of dimension  $n - 1$ , and are  $\ell_1$ -convex (by Corollary 1.12 of [18]), so our inductive hypothesis implies that  $\mu_{\pi A}$  and  $\mu_{\pi A_-}$  are weight measures on them. Since  $\pi(a) \in \pi A_- \subseteq \pi A$ , both sides of (5.18) are equal to 1, completing the proof.  $\square$

Our proof of Theorem 5.4.6 rests on the following result:

**Proposition 5.4.9.**  $|A| = \sum_{i=0}^n 2^{-i} V'_i(A)$  for all pixelated  $\ell_1$ -convex sets  $A \subseteq \ell_1^n$ .

*Proof.* Assume that  $A$  is 1-pixelated, and write  $A$  as a union  $\bigcup_{j=1}^m B_j$  of pixels. Also write  $W = \sum_{i=0}^n 2^{-i} V'_i$ ; then  $|B| = W(B)$  whenever  $B$  is a box or the empty set. Propositions 5.3.6 and 5.4.8 together with the valuation property of  $W$  give

$$\begin{aligned} |A| &= \mu_A(\mathbb{R}^n) = \sum_{k \geq 0} (-1)^k \sum_{1 \leq j_0 < \dots < j_k \leq m} \mu_{B_{j_0} \cap \dots \cap B_{j_k}}(\mathbb{R}^n) \\ &= \sum_{k \geq 0} (-1)^k \sum_{1 \leq j_0 < \dots < j_k \leq m} |B_{j_0} \cap \dots \cap B_{j_k}| \\ &= \sum_{k \geq 0} (-1)^k \sum_{1 \leq j_0 < \dots < j_k \leq m} W(B_{j_0} \cap \dots \cap B_{j_k}) = W(A), \end{aligned}$$

as required.  $\square$

*Proof of Theorem 5.4.6.* For part (1), let  $A \subseteq \ell_1^n$  be a compact  $\ell_1$ -convex set. For each  $\lambda > 0$ , let  $A_\lambda$  be the smallest  $\lambda$ -pixelated set containing  $A$ . Then  $A_\lambda$  is  $\ell_1$ -convex (by Proposition 3.1 of [18]), and  $A_\lambda \rightarrow A$  in the Hausdorff metric as  $\lambda \rightarrow 0$ . The result now follows from Proposition 5.4.9, continuity of the  $\ell_1$ -intrinsic volumes, and the monotonicity of magnitude (Proposition 5.3.2).

For (2), let  $A \subseteq \ell_1^n$  be a compact convex set with 0 in its interior. Given  $\varepsilon > 0$ , we can choose  $\alpha < 1$  such that  $d_H(\alpha A, A) < \varepsilon$ . But by convexity,  $\alpha A$  is a subset of the interior of  $A$ , so we can choose  $\lambda > 0$  such that  $\alpha A_\lambda \subseteq A$ . Thus, we have a pixelated  $\ell_1$ -convex subset  $B = \alpha A_\lambda$  of  $A$  satisfying  $d_H(B, A) < \varepsilon$ . Arguing as in part (1) but approximating from the inside rather than the outside, we obtain the opposite inequality  $|A| \geq \sum \frac{V'_i(A)}{2^i}$ . (Alternatively, use Theorem 5.4.15 below.)

For (3), the only nontrivial case remaining is that of a line segment, which is straightforward.  $\square$

### 5.4.3 The Fourier-analytic perspective

In the real line, the study of magnitude is facilitated by the order structure of  $\mathbb{R}$ ; in  $\ell_1^n$  we can exploit the algebraic structure of  $\ell_1$  products. In general normed spaces the most obvious special feature is translation-invariance. It will therefore come as no surprise that Fourier analysis is our key tool in that setting. This approach was developed in [28] for  $\ell_2^n$ , but with some additional effort we can work not only with more general norms but with the broader class of  $p$ -(quasi)norms for  $0 < p \leq 1$ .

Let  $0 < p \leq 1$ . A  **$p$ -norm** on a real vector space  $V$  is a function  $\|\cdot\| : V \rightarrow \mathbb{R}$  such that

- $\|v\| \geq 0$  for every  $v \in V$ , with equality only if  $v = 0$ ;
- $\|tv\| = |t| \|v\|$  for every  $t \in \mathbb{R}$  and  $v \in V$ ;
- $\|v + w\|^p \leq \|v\|^p + \|w\|^p$  for every  $v, w \in V$ .

Thus a 1-normed space is simply a normed space. A principal example of a  $p$ -normed space for  $p < 1$  is  $L_p[0, 1]$  with  $\|f\| = \left(\int_0^1 |f(x)|^p dx\right)^{1/p}$ .

If  $(V, \|\cdot\|)$  is a  $p$ -normed space, then  $d_p(v, w) = \|v - w\|^p$  is a metric on  $V$ . Conversely, if  $d$  is any translation-invariant, symmetric, positively homogeneous metric on a real vector space  $V$ , then  $\|v\| = d(v, 0)$  defines a  $p$ -norm on  $V$ , where  $p \in (0, 1]$  is the degree of homogeneity of  $d$ .

The following classical result, which goes back to Lévy [24] (see also [13, Theorem 6.6]), identifies which finite-dimensional  $p$ -normed spaces are positive definite metric spaces (and hence, by homogeneity, of negative type).

**Theorem 5.4.10.** *Let  $0 < p \leq 1$ , let  $\|\cdot\|$  be a  $p$ -norm on  $\mathbb{R}^n$ , and equip  $\mathbb{R}^n$  with the metric  $d_p(x, y) = \|x - y\|^p$ . Then  $(\mathbb{R}^n, d_p)$  is a positive definite metric space if and only if there is a linear map  $T : \mathbb{R}^n \rightarrow L_p[0, 1]$  such that  $\|Tx\|_p = \|x\|$  for every  $x \in \mathbb{R}^n$ .*

Theorem 5.4.10 implies in particular that  $L_p[0, 1]$  and  $\ell_p^n$  are positive definite with the metric  $d_p$  for  $0 < p \leq 1$ . We recall from Theorem 5.2.12 that  $L_q[0, 1]$  and  $\ell_q^n$  are also positive definite, with the usual metric, for  $1 \leq q \leq 2$ .

To simplify the statements of results:

For the rest of this section,  $\|\cdot\|$  will always denote a  $p$ -norm on  $\mathbb{R}^n$  such that  $(\mathbb{R}^n, d_p)$  is a positive definite metric space.

We will make use of the function  $F_p : \mathbb{R}^n \rightarrow \mathbb{R}$  defined by  $F_p(x) = e^{-\|x\|^p}$ , and denote by  $\mathcal{B} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$  the unit ball of  $\|\cdot\|$ . For  $f \in L_1(\mathbb{R}^n)$ , we adopt the convention that the Fourier transform of  $f$  is given by  $\widehat{f}(x) = \int_{\mathbb{R}^n} f(y) e^{-2\pi i \langle x, y \rangle} dy$ .

A key observation is that  $F_p$  is the Fourier transform of a  $p$ -stable probability distribution. Proposition 5.4.11 collects some crucial facts which follow from results from the literature on stable random processes.

**Proposition 5.4.11.**

1. There is a constant  $c > 0$  (depending on the  $p$ -norm  $\|\cdot\|$ ) such that  $\widehat{F}_p(x) \geq c(1 + \|x\|_2)^{-(1+p)n}$  for every  $x \in \mathbb{R}^n$ .
2. For each  $x \in \mathbb{R}^n$ ,  $\widehat{F}_p(tx)$  is nonincreasing as a function of  $t \geq 0$ . In particular,  $\|\widehat{F}_p\|_\infty = \widehat{F}_p(0) = \Gamma(\frac{n}{p} + 1) \text{vol } \mathcal{B}$ .

*Proof.* It follows from Theorem 5.4.10 and Bochner's theorem that  $\widehat{F}_p$  is the density of a  $p$ -stable distribution  $\mu$  on  $\mathbb{R}^n$ .

By a theorem of Lévy (see [13, Lemma 6.4]), there is a symmetric measure  $\sigma$  on  $S^{n-1}$  such that

$$\|x\|^p = \int_{S^{n-1}} |\langle x, \theta \rangle|^p d\sigma(\theta);$$

since  $\|x\| \neq 0$  for  $x \neq 0$ , the support of  $\sigma$  is not contained in any proper subspace of  $\mathbb{R}^n$ . Then  $\sigma$  is a positive scalar multiple of the spherical part of the Lévy measure of  $\mu$  (cf. [31, Section 14]). Since  $\sigma$  is symmetric and not supported in a proper subspace of  $\mathbb{R}^n$ , the linear span of its support is all of  $\mathbb{R}^n$ , and [38, Theorem 1.1(iii)] then implies the first claim.

Corollary 4.2 of [9] implies that every symmetric stable distribution on  $\mathbb{R}^n$  is *unimodal* in the sense defined in [9] and hence *n-unimodal* in the sense defined in [29] (see discussion on p. 80 and p. 84 of [9]). The second claim then follows from [29, Theorem 6].  $\square$

As in section 5.3.1, for a finite set  $B \subseteq \mathbb{R}^n$  and  $w \in \mathbb{R}^B$ , we write  $f_w(x) = \sum_{b \in B} w_b F_p(x - b)$ . Recall that the RKHS  $\mathcal{H}$  of the metric space  $(\mathbb{R}^n, d_p)$  is the completion of the span of such functions  $f_w$  with respect to the norm given by

$$\|f_w\|_{\mathcal{H}}^2 = \sum_{a, b \in B} w_a w_b F_p(a - b) = \int_{\mathbb{R}^n} \widehat{F}_p(x) \left| \sum_{b \in B} w_b e^{2\pi i \langle x, b \rangle} \right|^2 dx = \int_{\mathbb{R}^n} \frac{1}{\widehat{F}_p(x)} |\widehat{f}_w(x)|^2 dx.$$

Observe that the Fourier inversion theorem may be used here since  $\widehat{F}_p$  is the density of a random variable, hence integrable.

From here, standard arguments imply the following.

**Proposition 5.4.12.** *The RKHS of  $(\mathbb{R}^n, d_p)$  is*

$$\mathcal{H} = \left\{ f \in L_2(\mathbb{R}^n) \left| \int_{\mathbb{R}^n} \frac{1}{\widehat{F}_p(x)} |\widehat{f}(x)|^2 dx < \infty \right. \right\},$$

*with norm given by*

$$\|f\|_{\mathcal{H}}^2 = \int_{\mathbb{R}^n} \frac{1}{\widehat{F}_p(x)} |\widehat{f}(x)|^2 dx.$$

*The Schwartz space  $\mathcal{S}(\mathbb{R}^n)$  is contained in  $\mathcal{H}$ .*

The dual space of  $\mathcal{H}$  is naturally identified with the space of tempered distributions

$$\left\{ \phi \in \mathcal{S}'(\mathbb{R}^n) \mid \widehat{\phi} \in L_2(\widehat{F_p}(x) dx) \right\}.$$

Thus weightings for compact subsets of  $(\mathbb{R}^n, d_p)$  can be identified as tempered distributions satisfying a weak smoothness condition, although we will not make use of this fact here. Note that, since  $\widehat{F_p}$  is integrable, this space of distributions includes all finite signed measures on  $\mathbb{R}^n$ , so that weight measures fit gracefully into this perspective.

This concrete identification of the RKHS of  $(\mathbb{R}^n, d_p)$ , together with Proposition 5.4.11, make it possible to use Fourier analysis to prove a number of nice properties of magnitude in these spaces, including the following fundamental fact.

**Proposition 5.4.13.** *Let  $A \subseteq (\mathbb{R}^n, d_p)$  be compact. Then*

$$\frac{\text{vol } A}{\Gamma(\frac{n}{p} + 1) \text{vol } \mathcal{B}} \leq |A| < \infty.$$

*Proof.* By Proposition 5.4.12,  $\mathcal{H}$  contains functions which are uniformly equal to 1 on  $A$ , so the finiteness follows from Theorem 5.3.5.

For the lower bound, let  $h$  be the potential function of  $A$ . By Theorem 5.3.5, Proposition 5.4.12, Proposition 5.4.11(2), and Plancherel's theorem,

$$|A| = \|h\|_{\mathcal{H}}^2 \geq \frac{\|\widehat{h}\|_2^2}{\Gamma(\frac{n}{p} + 1) \text{vol } \mathcal{B}} = \frac{\|h\|_2^2}{\Gamma(\frac{n}{p} + 1) \text{vol } \mathcal{B}} \geq \frac{\text{vol } A}{\Gamma(\frac{n}{p} + 1) \text{vol } \mathcal{B}}. \quad \square$$

The finiteness statement in Proposition 5.4.13 was proved in Theorem 3.4.8 and Proposition 3.5.3 of [19] for  $\ell_1^n$  and  $\ell_2^n$ , and in somewhat greater generality in [27, Theorem 4.3]. The lower bound was proved in [19, Theorem 3.5.6] for  $p = 1$  and [27, Theorem 4.5] for the general case.<sup>5.1</sup> The proof here follows the approach used in [28] for  $\ell_2^n$  (see Proposition 5.6 and the remarks following Corollary 5.3 there).

We now consider the behavior of magnitude functions in  $(\mathbb{R}^n, d_p)$ . We must be careful about a subtle notational issue when  $p < 1$ . Recall that for a metric space  $(A, d)$  and  $t > 0$ , we denote by  $tA$  the metric space  $(A, td)$ , which in the present context is different from the usual interpretation of  $tA$ . Therefore we will introduce the notation  $t \cdot A = \{ta \mid a \in A\}$  for  $A \subseteq \mathbb{R}^n$ . Note that when  $A \subseteq \mathbb{R}^n$  is equipped with the metric  $d_p(x, y) = \|x - y\|^p$  associated to a  $p$ -norm, the metric space  $tA$  is isometric to the set  $t^{1/p} \cdot A \subseteq \mathbb{R}^n$  equipped with  $d_p$ .

The next result shows that magnitude knows about volume in all finite-dimensional positive definite  $p$ -normed spaces. This generalizes [3, Theorem 1] for Euclidean space  $\ell_2^n$ .

<sup>5.1</sup> Theorem 4.5 in the published version of [27] is misstated in the case  $p < 1$ ; see the current arXiv version for a correct statement.

**Theorem 5.4.14.** *If  $A \subseteq (\mathbb{R}^n, d_p)$  is compact, then*

$$\lim_{t \rightarrow \infty} \frac{|tA|}{t^{n/p}} = \lim_{t \rightarrow \infty} \frac{|t \cdot A|}{t^n} = \frac{\text{vol } A}{\Gamma(\frac{n}{p} + 1) \text{vol } \mathcal{B}}.$$

*Proof.* Proposition 5.4.13 implies that

$$|t \cdot A| \geq \frac{\text{vol}(t \cdot A)}{\Gamma(\frac{n}{p} + 1) \text{vol } \mathcal{B}} = \frac{t^n \text{vol } A}{\Gamma(\frac{n}{p} + 1) \text{vol } \mathcal{B}}$$

for every  $t > 0$ . Now suppose that  $h \in \mathcal{H}$  satisfies  $h \equiv 1$  on  $A$ , and let  $h_t(x) = h(x/t)$ . Then by Theorem 5.3.5 and Proposition 5.4.12,

$$\frac{|t \cdot A|}{t^n} \leq \int_{\mathbb{R}^n} \frac{1}{\widehat{F_p}(x)} |\widehat{h}_t(x)|^2 dx = \int_{\mathbb{R}^n} \frac{1}{\widehat{F_p}(x/t)} |\widehat{h}(x)|^2 dx. \quad (5.19)$$

Proposition 5.4.11(2), the monotone convergence theorem, and Plancherel's theorem imply that

$$\lim_{t \rightarrow \infty} \int_{\mathbb{R}^n} \frac{1}{\widehat{F_p}(x/t)} |\widehat{h}(x)|^2 dx = \frac{\|h\|_2^2}{\Gamma(\frac{n}{p} + 1) \text{vol } \mathcal{B}}.$$

By Theorem 5.4.12, there exist functions  $h \in \mathcal{H}$  with  $h \equiv 1$  on  $A$  such that  $\|h\|_2^2$  is arbitrarily close to  $\text{vol } A$  (cf. the proof of [3, Theorem 1]), which completes the proof.  $\square$

The next theorem is the major known continuity result (as opposed to mere *semicontinuity*) for magnitude.

**Theorem 5.4.15.** *Denote by  $\mathcal{K}_n$  the class of nonempty compact subsets of  $\mathbb{R}^n$ , equipped with the Hausdorff metric  $d_H$  induced by  $d_p$ , and suppose that  $A \in \mathcal{K}_n$  is star-shaped with respect to some point in its interior. Then magnitude, as a function  $\mathcal{K}_n \rightarrow \mathbb{R}$ , is continuous at  $A$ .*

*Proof.* By Proposition 5.3.1, we only need to show that magnitude is upper semicontinuous at  $A$ . Letting  $h$  be the potential function of  $A$ , (5.19) and Proposition 5.4.12 imply that  $|t \cdot A| \leq t^n |A|$  for  $t \geq 1$ . By translation-invariance, we may assume that  $A$  is star-shaped about 0 and  $r^{1/p} \cdot \mathcal{B} \subseteq A$  for some  $r > 0$ . Now if  $B \in \mathcal{K}_n$  and  $d_H(A, B) < \varepsilon$ , then

$$B \subseteq A + \varepsilon^{1/p} \cdot \mathcal{B} \subseteq \left(1 + \left(\frac{\varepsilon}{r}\right)^{1/p}\right) \cdot A,$$

and so  $|B| \leq \left(1 + \left(\frac{\varepsilon}{r}\right)^{1/p}\right)^n |A|$ . Thus magnitude is upper semicontinuous at  $A$ .  $\square$

The family of sets  $A$  in Theorem 5.4.15 is slightly larger than what are sometimes called “star bodies”, and of course includes all convex bodies. It is unknown, however, whether magnitude is continuous when restricted to compact convex sets which are not required to have nonempty interior.

The final result in this section shows that, in positive definite  $p$ -normed spaces, magnitude can be computed from potential functions simply by integrating, as opposed to computing the (more complicated)  $\mathcal{H}$ -norm.

**Theorem 5.4.16.** *Let  $A \subseteq (\mathbb{R}^n, d_p)$  be compact, and suppose that the potential function  $h \in \mathcal{H}$  of  $A$  is integrable. Then*

$$|A| = \frac{1}{\Gamma(\frac{n}{p} + 1) \operatorname{vol} \mathcal{B}} \int_{\mathbb{R}^n} h(x) \, dx.$$

*Proof.* Fix an even function  $f \in \mathcal{S}(\mathbb{R}^n)$  with  $f \equiv 1$  on some open neighborhood of the origin. Set  $f_k(x) = f(x/k)$  and  $\phi_k = \widehat{f_k / \widehat{F_p}}$  for  $k \in \mathbb{N}$ . Then  $\phi_k \in L_1(\mathbb{R}^n)$  and

$$\|\widehat{\phi_k}\|_\infty \leq \|\phi_k\|_1 = \int_{\mathbb{R}^n} \frac{1}{\widehat{F_p}(x/k)} |\widehat{f}(x)| \, dx \leq \int_{\mathbb{R}^n} \frac{1}{\widehat{F_p}(x)} |\widehat{f}(x)| \, dx < \infty$$

by Proposition 5.4.11(2). Furthermore, for every  $x \in \mathbb{R}^n$ ,

$$\widehat{\phi_k}(x) = \int_{\mathbb{R}^n} e^{-2\pi i \langle x, y/k \rangle} \frac{\widehat{f}(y)}{\widehat{F_p}(y/k)} \, dy \xrightarrow{k \rightarrow \infty} \int_{\mathbb{R}^n} \frac{\widehat{f}(y)}{\widehat{F_p}(0)} \, dy = \frac{f(0)}{\Gamma(\frac{n}{p} + 1) \operatorname{vol} \mathcal{B}} = \frac{1}{\Gamma(\frac{n}{p} + 1) \operatorname{vol} \mathcal{B}}$$

by the dominated convergence theorem.

By the last part of Theorem 5.3.5, for sufficiently large  $k$ ,

$$|A| = \langle h, f_k \rangle_{\mathcal{H}} = \int_{\mathbb{R}^n} \widehat{h}(x) \phi_k(x) \, dx = \int_{\mathbb{R}^n} h(x) \widehat{\phi_k}(x) \, dx$$

by Parseval's identity, and the claim now follows by the dominated convergence theorem.  $\square$

#### 5.4.4 Magnitude in Euclidean space

Finally, we specialize the tools of section 5.4.3 to the setting of Euclidean space  $\ell_2^n$ , where they become even more powerful, allowing one to prove much more refined results about continuity, asymptotics, and exact values of magnitude than in more general normed spaces.

We will write simply  $F(x) = e^{-\|x\|_2}$ , and let  $\mathcal{B}_R^n = \{x \in \mathbb{R}^n \mid \|x\|_2 \leq R\}$ . In this setting we have the explicit formula

$$\widehat{F}(x) = \frac{n! \omega_n}{(1 + 4\pi^2 \|x\|_2^2)^{(n+1)/2}}, \quad (5.20)$$

where  $\omega_n = \operatorname{vol}_n(\mathcal{B}_1^n)$  (see [37, Theorem 1.14]). This implies that the RKHS  $\mathcal{H}$  for  $\ell_2^n$  is the classical Sobolev space

$$H^{(n+1)/2}(\mathbb{R}^n) = \left\{ f \in L_2(\mathbb{R}^n) \left| \int_{\mathbb{R}^n} (1 + 4\pi^2 \|x\|_2^2)^{(n+1)/2} |\widehat{f}(x)|^2 \, dx < \infty \right. \right\},$$

and that  $\|f\|_{\mathcal{H}}^2 = \frac{1}{n! \omega_n} \|f\|_{H^{(n+1)/2}}^2$ .

A first application of this observation is the following, proved for  $\ell_2^n$  in [28, Corollary 5.5].

**Theorem 5.4.17.** *If  $A$  is a compact subset of  $\ell_1^n$  or  $\ell_2^n$ , then the magnitude function  $t \mapsto |tA|$  is continuous on  $(0, \infty)$ .*

*Sketch of proof.* For  $\ell_2^n$ , using (5.20) one can show that  $|tA| \geq \frac{1}{t} |A|$  for  $t \geq 1$ , along the lines of (5.19). For  $\ell_1^n$ , if we let  $G(x) = e^{-\|x\|_1} = \prod_{i=1}^n e^{-|x_i|}$ , then the  $n = 1$  case of (5.20) implies that

$$\widehat{G}(x) = \prod_{i=1}^n \frac{2}{1 + 4\pi^2 x_i^2},$$

and a similar argument yields that  $|tA| \geq t^{-n} |A|$  for  $t \geq 1$ .

In either case, (5.19) shows that  $|tA| \leq t^n |A|$ . Together, these estimates imply that the magnitude function of  $A$  is continuous on  $(0, \infty)$ ; see Theorem 5.4 and Corollary 5.5 of [28].  $\square$

The most significant consequence of (5.20) is that when  $n$  is odd,  $1/\widehat{F}$  is the symbol of a differential operator on  $\mathbb{R}^n$ . In particular, when  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is smooth,

$$\|f\|_{H^{(n+1)/2}}^2 = \int_{\mathbb{R}^n} f(x) [(I - \Delta)^{(n+1)/2} f](x) dx, \quad (5.21)$$

where  $I$  is the identity operator and  $\Delta$  is the Laplacian on  $\mathbb{R}^n$ . This opens the door to using differential equations techniques to study magnitude. A first application is the proof of the following result.

**Theorem 5.4.18** ([3, Theorem 1]). *If  $A \subseteq \ell_2^n$  is compact, then  $\lim_{t \rightarrow 0^+} |tA| = 1$ .*

*Sketch of proof.* By Proposition 5.3.2, it suffices to show that  $\limsup_{R \rightarrow 0^+} |\mathcal{B}_R^n| \leq 1$ ; it further suffices, by embedding  $\ell_2^n$  in  $\ell_2^{n+1}$  if necessary, to assume that  $n$  is odd. For  $0 < R < 1$  we can choose smooth functions  $f_R$  such that

$$f_R(x) = \begin{cases} 1 & \text{if } \|x\|_2 \leq R, \\ e^R e^{-\|x\|_2} & \text{if } \|x\|_2 \geq \sqrt{R} \end{cases}$$

and the derivatives of  $f_R$  are sufficiently small for  $R \leq \|x\|_2 \leq \sqrt{R}$  that, using (5.21),

$$\|f_R\|_{H^{(n+1)/2}}^2 = \|f_0\|_{H^{(n+1)/2}}^2 + o(1) = n! \omega_n + o(1)$$

when  $R \rightarrow 0$ ; see the proof of [3, Theorem 1]. By Theorem 5.3.5, this completes the proof.  $\square$

Together with Theorem 5.4.17, this shows that the magnitude function of a compact  $A \subseteq \ell_2^n$  is continuous on  $[0, \infty)$ . Recall that this result is false for general metric spaces

$A$  of negative type [19, Example 2.2.8], but it does also hold for  $A \subseteq \ell_1^n$  (Proposition 5.4.4). A monotone convergence argument would prove the same result in a  $p$ -normed space if  $\sup_x \widehat{F}_p(x)/\widehat{F}_p(2x) < \infty$ .

More significantly, we obtain the following conditions on the potential function of a compact set  $A \subseteq \ell_2^n$ , which provide the starting point for the only known approach for explicit computation of magnitude for a convex body in  $\ell_2^n$  when  $n > 1$ . This result follows by considering the Euler–Lagrange equation of the minimization problem in Theorem 5.3.5, and applying elliptic regularity.

**Theorem 5.4.19** (Proposition 5.7 and Corollary 5.8 of [28]). *Suppose that  $n$  is odd and  $A \subseteq \ell_2^n$  is compact. Then the potential function  $h$  of  $A$  is  $C^\infty$  on  $\mathbb{R}^n \setminus A$ , and satisfies*

$$(I - \Delta)^{(n+1)/2} h(x) = 0 \quad (5.22)$$

on  $\mathbb{R}^n \setminus A$ .

To indicate the usefulness of this observation, we show how Theorem 5.4.19 can be used to quickly compute the magnitude of an interval in  $[a, b] \subseteq \mathbb{R}$ . By Theorem 5.4.19, the potential function  $h$  satisfies  $h - h'' = 0$  outside  $[a, b]$ . The boundary conditions  $h(x) = 1$  for  $a \leq x \leq b$  and  $h(x) \rightarrow 0$  when  $|x| \rightarrow \infty$  (since  $h \in H^1(\mathbb{R})$ ) imply that

$$h(x) = \begin{cases} e^{x-a} & \text{if } x < a, \\ 1 & \text{if } a \leq x \leq b, \\ e^{b-x} & \text{if } x > b. \end{cases}$$

Then by Theorem 5.4.16,

$$|[a, b]| = \frac{1}{2} \int_{\mathbb{R}} h(x) dx = 1 + \frac{b-a}{2},$$

in agreement with (5.7). A more involved, but still elementary computation yields another proof of Corollary 5.4.3.

For higher dimensions, Barceló and Carbery [3] analyzed the minimization problem in more depth, and proved the following result using standard techniques of the theory of partial differential equations.

**Proposition 5.4.20** (See Proposition 2 and Lemma 4 of [3]). *Suppose that  $n$  and  $m$  are positive integers, and  $A \subseteq \mathbb{R}^n$  is a convex body.*

1. *There is a unique function  $f \in H^m(\mathbb{R}^n)$  such that*

$$(I - \Delta)^m f(x) = 0 \text{ on } \mathbb{R}^n \setminus A$$

*weakly and  $f \equiv 1$  on  $A$ .*

2. *If  $\partial A$  is piecewise  $C^1$  and  $f \in H^m(\mathbb{R}^n)$ , then all derivatives of  $f$  up to order  $m - 1$  vanish on  $\partial A$  (in the sense of traces of Sobolev functions).*



Together with Theorems 5.4.16 and 5.4.19, Proposition 5.4.20 reduces the computation of magnitudes (in many cases) to the solution of a PDE boundary value problem. In general, of course, solving a PDE boundary value problem is no simple matter. But in the case that  $A = \mathcal{B}_R^n$  is a Euclidean ball, rotational symmetry reduces the partial differential equation to an ordinary differential equation on  $[R, \infty)$ , albeit of high degree. Barceló and Carbery gave an algorithm for solving the resulting ODE boundary value problem, and hence determining the potential function  $h$  of  $\mathcal{B}_R^n$ , for every odd dimension  $n$  and radius  $R > 0$ . From there, Theorem 5.4.16 can be used to compute the magnitude  $\mathcal{B}_R^n$ . (In [3] the magnitude was found by computing  $\|h\|_{H^{(n+1)/2}}^2$  using (5.21), since Theorem 5.4.16 had not yet been proved; Theorem 5.4.16 makes the computation much simpler.) This approach yields the following.

**Theorem 5.4.21** (Theorems 2, 3, and 4 of [3]). *For every  $R > 0$ ,*

$$|\mathcal{B}_R^3| = 1 + 2R + R^2 + \frac{1}{6}R^3$$

*and*

$$|\mathcal{B}_R^5| = \frac{24 + 72R^2 + 35R^3 + 9R^4 + R^5}{8(R+3)} + \frac{1}{120}R^5.$$

*In general, when  $n$  is odd, the magnitude  $|\mathcal{B}_R^n|$  is a rational function of  $R > 0$  with rational coefficients.*

Barceló and Carbery also give an explicit formula for  $|\mathcal{B}_R^7|$ . We recall that they also determined the asymptotics of  $|\mathcal{B}_R^n|$  when  $R \rightarrow 0$  and  $R \rightarrow \infty$  in [3, Theorem 1], stated above in Theorems 5.4.18 and 5.4.14. To date, odd-dimensional balls are the only convex bodies in Euclidean space whose exact magnitudes are known.

It was previously conjectured in [23] that for a compact convex set  $A \subseteq \ell_2^n$ ,

$$|A| = \sum_{i=0}^n \frac{V_i(A)}{i! \omega_i}, \quad (5.23)$$

where  $V_0, \dots, V_n$  denotes the classical intrinsic volumes. Theorem 5.4.21 implies that (5.23) holds for balls in  $\ell_2^3$ , but is false in dimensions  $n \geq 5$ .

To put this conjecture in context, observe that (5.23) is a Euclidean version of Conjecture 5.4.5 in  $\ell_1^n$ . Note that  $2^i = i! \omega'_i$ , where  $\omega'_i$  denotes the volume of the unit ball in  $\ell_1^i$ , tightening the analogy between (5.23) and Conjecture 5.4.5. At the time that (5.23) was proposed, it was known to hold for  $n = 1$ , and was supported by numerical computations in  $n = 2$  [40]. Furthermore, some cases of Conjecture 5.4.5 in  $\ell_1^n$  (contained in Theorem 5.4.6) were known to be true.

Several interesting questions remain open, most obviously whether (5.23) holds for  $n \leq 4$ . Noting that (5.23) is equivalent to

$$|tA| = \sum_{i=0}^n \frac{V_i(A)}{i! \omega_i} t^i, \quad (5.24)$$

Proposition 5.4.13 and Theorem 5.4.14 say that (5.24) is true to top order for sets of positive volume when  $t \rightarrow \infty$ , and Theorem 5.4.18 shows that (5.23) predicts the correct behavior when  $t \rightarrow 0$ . One could ask whether (5.24) is approximately true in some sharper asymptotic senses. Note that Theorem 5.3.11 is a Riemannian analogue of a weak asymptotic version of (5.24). Barceló and Carbery also raise the question of whether (5.23) holds if magnitude is replaced by a suitable modification which coincides with magnitude in  $\ell_2^n$  for  $n \leq 3$ . We mention another related question in section 5.5.

The final major consequence of the concrete identification of  $\mathcal{H}$  for Euclidean space is the realization that magnitude and maximum diversity, in the setting of  $\ell_2^n$ , are actually classical notions of capacity, well-known in potential theory. The formal similarity between magnitude and maximum diversity on the one hand, and capacity on the other, is clear from the definitions (cf. section 1.1 of [3]). But in  $\ell_2^n$ , magnitude and maximum diversity almost precisely reproduce classically studied forms of capacity.

Specifically, (5.20) and [1, Theorem 2.2.7] imply that for a compact set  $A \subseteq \ell_2^n$ ,  $|A|_+ = \frac{1}{n! \omega_n} C_{(n+1)/2}(A)$ , where

$$C_\alpha(A) := \inf \left\{ \|f\|_{H^\alpha}^2 \mid f \in \mathcal{S}(\mathbb{R}^n), f \geq 1 \text{ on } A \right\},$$

is the **Bessel capacity** of  $A$  of order  $\alpha$ . An alternative notion of capacity, which naturally arises in the study of removability of singularities (see [1, Section 2.7]), is

$$N_\alpha(A) := \inf \left\{ \|f\|_{H^\alpha}^2 \mid f \in \mathcal{S}(\mathbb{R}^n), f \geq 1 \text{ on a neighborhood of } A \right\}.$$

By Theorem 5.3.5,  $|A| \leq \frac{1}{n! \omega_n} N_{(n+1)/2}(A)$ . In fact one would expect  $|A| = \frac{1}{n! \omega_n} N_{(n+1)/2}(A)$ ; this appears not to be the case for arbitrary compact  $A$ , but happily the comparison we have will be enough for our purpose.

Before moving on, we pause to observe that, although we have just seen that magnitude and its cousin maximum diversity fit into classical families of capacities, they both just fail to fit into the parameter range which is of relevance for classical applications. As alluded to above, capacities are frequently used to quantify “exceptional” sets; sets of capacity 0 are a frequent substitute for sets of measure 0 when studying singularities. However,  $C_\alpha(A)$  and  $N_\alpha(A)$  are bounded below by positive constants whenever  $\alpha > n/2$ . So from the point of view of classical potential theory, magnitude and maximum diversity are rather pathological. Nevertheless, the following result from potential theory, whose main classical application is to show that  $C_\alpha(A) = 0$  if and only if  $N_\alpha(A) = 0$ , also applies in our setting.

**Proposition 5.4.22** ([1, Theorem 3.3.4]). *For each  $n$  and each  $\alpha > 0$  there is a constant  $\kappa_{n,\alpha} \geq 1$  such that, for every compact set  $A \subseteq \ell_2^n$ ,*

$$C_\alpha(A) \leq N_\alpha(A) \leq \kappa_{n,\alpha} C_\alpha(A).$$

**Corollary 5.4.23** ([28, Corollary 6.2]). *For each  $n$  there is a constant  $\kappa_n \geq 1$  such that, for every compact set  $A \subseteq \ell_2^n$ ,*

$$|A|_+ \leq |A| \leq \kappa_n |A|_+.$$

The significance of Corollary 5.4.23 is that, although maximum diversity is no easier to compute explicitly than magnitude, in some ways its rough behavior is easier to analyze. For example, it is natural to conjecture that the magnitude function  $t \mapsto |tA|$  is increasing for a compact space  $A$  of negative type. It is unknown whether this is true. On the other hand, it is obvious that  $t \mapsto |tA|_+$  is increasing, and Corollary 5.4.23 therefore implies that the magnitude function of a compact set  $A \subseteq \ell_2^n$  is at least bounded above and below by constant multiples of an increasing function.

A more substantial consequence of Corollary 5.4.23 is the following result, which, like Theorem 5.4.14, shows that the category-theoretically inspired notion of magnitude turns out to encode quantities of fundamental importance in geometry.

**Theorem 5.4.24** ([28, Corollary 7.4]). *If  $A \subseteq \ell_2^n$  is compact, then*

$$\lim_{t \rightarrow \infty} \frac{\log |tA|}{\log t} = \dim_{\text{Mink}} A.$$

Theorem 5.4.24, which should be interpreted in the same sense as Proposition 5.3.14, follows immediately from Proposition 5.3.14 and Corollary 5.4.23. Another interesting aspect of this result is that, as noted above, classically Proposition 5.4.22 is of interest primarily for sets of capacity 0, or more generally for small sets; here it is instead applied to large sets.

## 5.5 Open problems

There are many interesting open problems about magnitude. These include extending partial results discussed above, as well as some quite basic questions about the behavior of magnitude. We mention several of them below.

1. Does every compact positive definite space (or space of negative type) have finite magnitude?

Proposition 5.4.13 implies that every compact subset of a finite dimensional positive definite normed (or  $p$ -normed) space has finite magnitude, so that the obvious place to look for a counterexample is in infinite dimensions. Essentially the only infinite-dimensional spaces whose magnitudes are known are boxes in  $\ell_1$ , which just miss being a counterexample:

$$\left| \prod_{i=1}^{\infty} [0, r_i] \right| = \prod_{i=1}^{\infty} \left( 1 + \frac{r_i}{2} \right).$$

The condition  $\|r\|_1 < \infty$ , which both guarantees that this infinite-dimensional box lies in  $\ell_1$  and is compact, is also equivalent to the finiteness of the product on the right-hand side.

2. Is magnitude continuous on the class of compact sets in a positive definite normed (or  $p$ -normed) space? What if we assume the space is finite-dimensional, or we restrict to geodesic sets, or convex sets?

Recall that magnitude is continuous on convex bodies in a finite-dimensional positive definite  $p$ -normed space (Theorem 5.4.15), but is not continuous on the class of compact spaces of negative type (Examples 2.2.8 and 2.4.9 of [19]).

3. Is Conjecture 5.4.5 true? Is it at least true for compact convex sets  $A \subseteq \ell_1^n$ ?

In light of Theorem 5.4.6, Conjecture 5.4.5 is equivalent to the continuity of magnitude on compact, geodesic (i.e.,  $\ell_1$ -convex) sets in  $\ell_1^n$ . Similarly, if magnitude is continuous on compact convex sets in  $\ell_1^n$ , then Theorem 5.4.6 would imply that Conjecture 5.4.5 holds for compact convex sets.

4. Does the magnitude function of a convex body  $A \subseteq \ell_2^n$  determine its intrinsic volumes? What about a homogeneous compact Riemannian manifold?
5. Does it hold that

$$|t(A \cup B)| + |t(A \cap B)| - |tA| - |tB| \rightarrow 0$$

as  $t \rightarrow \infty$  for compact, convex sets  $A, B \subseteq \ell_2^n$  (or in more general normed spaces) such that  $A \cup B$  is convex?

For convex bodies in  $\ell_1^n$ , the left-hand side of the above is 0 for every  $t$ , as a consequence of Theorem 5.4.6; the same would be true in  $\ell_2^n$  if (5.23) were true.

6. Does Theorem 5.4.24 hold for arbitrary compact spaces of negative type?

Theorem 5.4.24 applies to compact subsets of  $\ell_2^n$ . As mentioned earlier, Proposition 7.5 of [28] shows that the conclusion of Theorem 5.4.24 also holds for compact homogeneous metric spaces. In addition, Theorem 5.4.14 implies that the conclusion of Theorem 5.4.24 holds for compact subsets of positive  $n$ -dimensional volume in an  $n$ -dimensional positive definite  $p$ -normed space; hence it holds, for example, for compact convex sets in any positive definite  $p$ -normed space.

## Bibliography

- [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] J. A. Barceló and A. Carbery: On the magnitudes of compact sets in Euclidean spaces. arXiv:1507.02502, 2015. To appear in *Amer. J. Math.*
- [4] Y. Benyamini and J. Lindenstrauss: *Geometric Nonlinear Functional Analysis. Vol. 1*, volume 48 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2000.

- [5] J. Chuang, A. King, and T. Leinster: On the magnitude of a finite dimensional algebra. *Theory and Applications of Categories*, 31:63–72, 2016.
- [6] M. M. Deza and M. Laurent: *Geometry of Cuts and Metrics*, volume 15 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 1997.
- [7] M. Gromov: *Metric Structures for Riemannian and Non-Riemannian Spaces*. Birkhäuser, Boston, 2001.
- [8] R. Hepworth and S. Willerton: Categorifying the magnitude of a graph. To appear in *Homology, Homotopy, and Applications*, 2015.
- [9] M. Kanter: Unimodality and dominance for symmetric random vectors. *Trans. Amer. Math. Soc.*, 229:65–85, 1977.
- [10] G. M. Kelly: *Basic Concepts of Enriched Category Theory*, volume 64 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1982. Also *Reprints in Theory and Applications of Categories* 10 (2005), 1–136.
- [11] D. A. Klain and G.-C. Rota: *Introduction to Geometric Probability*. Lezioni Lincee. Cambridge University Press, Cambridge, 1997.
- [12] S. L. Kulkarni: Does negative type characterize the round sphere? *Proc. Amer. Math. Soc.*, 135(11):3695–3702 (electronic), 2007.
- [13] A. Koldobsky: *Fourier Analysis in Convex Geometry*, volume 116 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2005.
- [14] F. W. Lawvere: Metric spaces, generalized logic and closed categories. *Rendiconti del Seminario Matematico e Fisico di Milano*, XLIII:135–166, 1973. Also *Reprints in Theory and Applications of Categories* 1:1–37, 2002.
- [15] F. W. Lawvere: Taking categories seriously. *Revista Colombiana de Matemáticas*, XX:147–178, 1986. Also *Reprints in Theory and Applications of Categories* 8:1–24, 2005.
- [16] T. Leinster: The Euler characteristic of a category. *Documenta Mathematica*, 13:21–49, 2008.
- [17] T. Leinster: A maximum entropy theorem with applications to the measurement of biodiversity. arXiv:0910.0906, 2009.
- [18] T. Leinster: Integral geometry for the 1-norm. *Advances in Applied Mathematics*, 49:81–96, 2012.
- [19] T. Leinster: The magnitude of metric spaces. *Documenta Mathematica*, 18:857–905, 2013.
- [20] T. Leinster: The magnitude of a graph. To appear in *Math. Proc. Cam. Phil. Soc.*, 2014.
- [21] T. Leinster and C. Cobbold: Measuring diversity: the importance of species similarity. *Ecology*, 93:477–489, 2012.
- [22] T. Leinster and M. W. Meckes: Maximizing diversity in biology and beyond. *Entropy*, 18(3):88, 2016.
- [23] T. Leinster and S. Willerton: On the asymptotic magnitude of subsets of Euclidean space. *Geom. Dedicata*, 164(1):287–310, 2013.
- [24] P. Lévy: *Théorie de l'addition de variables aléatoires*. Gauthier-Villars, Paris, 1937.
- [25] N. Linial: Finite metric-spaces—combinatorics, geometry and algorithms. In *Proceedings of the International Congress of Mathematicians, Vol. III (Beijing, 2002)*, pages 573–586. Higher Ed. Press, Beijing, 2002.
- [26] S. Mac Lane: *Categories for the Working Mathematician*. Graduate Texts in Mathematics 5. Springer, New York, 1971.
- [27] M. W. Meckes: Positive definite metric spaces. *Positivity*, 17:733–757, 2013. Corrected version available at [arxiv.org/abs/1012.5863v5](http://arxiv.org/abs/1012.5863v5).
- [28] M. W. Meckes: Magnitude, diversity, capacities, and dimensions of metric spaces. *Potential Analysis*, 42:549–572, 2015.
- [29] R. A. Olshen and L. J. Savage: A generalized unimodality. *J. Appl. Probability*, 7:21–34, 1970.

- [30] G.-C. Rota: On the foundations of combinatorial theory I: theory of Möbius functions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 2:340–368, 1964.
- [31] K.-I. Sato: *Lévy Processes and Infinitely Divisible Distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2013. Translated from the 1990 Japanese original, Revised edition of the 1999 English translation.
- [32] S. H. Schanuel: What is the length of a potato? An introduction to geometric measure theory. In *Categories in Continuum Physics*, Lecture Notes in Mathematics 1174. Springer, Berlin, 1986.
- [33] S. H. Schanuel: Negative sets have Euler characteristic and dimension. In *Category Theory (Como, 1990)*, Lecture Notes in Mathematics 1488, pages 379–385. Springer, Berlin, 1991.
- [34] R. Schneider: *Convex Bodies: the Brunn–Minkowski Theory*, volume 151 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, expanded edition, 2014.
- [35] A. R. Solow and S. Polasky: Measuring biological diversity. *Environmental and Ecological Statistics*, 1:95–107, 1994.
- [36] D. Speyer: Re: Metric spaces. Comment at [golem.ph.utexas.edu/category/2008/02/metric\\_spaces.html](http://golem.ph.utexas.edu/category/2008/02/metric_spaces.html), 2008.
- [37] 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.
- [38] T. Watanabe: Asymptotic estimates of multi-dimensional stable densities and their applications. *Trans. Amer. Math. Soc.*, 359(6):2851–2879 (electronic), 2007.
- [39] J. H. Wells and L. R. Williams: *Embeddings and Extensions in Analysis*. Springer-Verlag, New York-Heidelberg, 1975. *Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 84*.
- [40] S. Willerton: Heuristic and computer calculations for the magnitude of metric spaces. [arxiv.org/abs/0910.5500](https://arxiv.org/abs/0910.5500), 2009.
- [41] S. Willerton: On the magnitude of spheres, surfaces and other homogeneous spaces. *Geom. Dedicata*, 168:291–310, 2014.