The Magnitude of Metric Spaces

Tom Leinster

Received: August 15, 2012

Communicated by Stefan Schwede

Abstract. Magnitude is a real-valued invariant of metric spaces, analogous to Euler characteristic of topological spaces and cardinality of sets. The definition of magnitude is a special case of a general categorical definition that clarifies the analogies between cardinality-like invariants in mathematics. Although this motivation is a world away from geometric measure, magnitude, when applied to subsets of $\mathbb{R}^n$, turns out to be intimately related to invariants such as volume, surface area, perimeter and dimension. We describe several aspects of this relationship, providing evidence for a conjecture (first stated in joint work with Willerton) that magnitude encodes all the most important invariants of classical integral geometry.

2010 Mathematics Subject Classification: 51F99 (primary), 18D20, 18F99, 28A75, 49Q20, 52A20, 52A38, 53C65 (secondary).
Keywords and Phrases: metric space, magnitude, enriched category, Möbius inversion, Euler characteristic of a category, finite metric space, convex set, integral geometry, valuation, intrinsic volume, fractal dimension, positive definite space, space of negative type.

Contents

Introduction

Many mathematical objects carry a canonical notion of size. Sets have cardinality, vector spaces have dimension, topological spaces have Euler characteristic, and probability spaces have entropy. This work adds a new item to the list: metric spaces have magnitude.

Already, several cardinality-like invariants are tied together by the notion of the Euler characteristic of a category [21]. This is a rational-valued invariant of finite categories. A network of theorems describes the close relationships between this invariant and established cardinality-like invariants, including the cardinality of sets and of groupoids [1], the Euler characteristic of topological spaces and of posets, and even the Euler characteristic of orbifolds. (That Euler characteristic deserves to be considered an analogue of cardinality was first made clear by Schanuel [41, 42].) These results attest that for categories, Euler characteristic is the fundamental notion of size.

Here we go further. Categories are a special case of the more general concept of enriched category. Much of ordinary category theory generalizes to the enriched setting, and this is true, in particular, of the Euler characteristic of categories. Rebaptizing Euler characteristic as 'magnitude' to avoid a potential ambiguity later, this gives a canonical definition of the magnitude of an enriched category. Metric spaces, as well as categories, are examples of enriched categories:

$$\text{(categories)} \subset \text{(enriched categories)} \supset \text{(metric spaces)}$$

[19, 20]. The analogy between categories and metric spaces can be understood in broad terms immediately. A category has objects; a metric space has points. For any two objects there is a set (the maps between them); for any two points there is a real number (the distance between them). For any three objects there is an operation of composition; for any three points there is a triangle inequality.

Having generalized the definition of magnitude (or Euler characteristic) from ordinary to enriched categories, we specialize it to metric spaces. This gives our invariant. The fundamental role of the Euler characteristic of categories strongly suggests that the magnitude of metric spaces should play a fundamental role too. Our faith is rewarded by a series of theorems showing that magnitude is intimately related to the classical invariants of integral geometry: dimension, perimeter, surface area, volume, ... This is despite the fact that no concept of measure or integration goes into the definition of magnitude; they arise spontaneously from the general categorical definition.

This, then, is part of the appeal of magnitude. It is motivated in abstract terms that seem to have nothing to do with geometry or measure, and defined in the very wide generality of enriched categories—yet when specialized to the context of metric spaces, it turns out to have a great deal to say about geometry and measure.

There is a further surprise. While the author's motivation was category-theoretic, magnitude had already arisen in work on the quantification of biological diversity. In 1994, Solow and Polasky [44] carried out a probabilistic analysis of the benefits of high biodiversity, and isolated a particular quantity that they called the 'effective number of species'. It is the same as our magnitude. This is no coincidence: the theoretical problem of how to maximize diversity can be solved using the concept of magnitude and some of the results presented here [23]. Indeed, under suitable circumstances, magnitude can be interpreted as maximum diversity, a cousin to maximum entropy.

Our first step is to define the magnitude of an enriched category (Section 1). This puts the notion of the magnitude of a metric space into a wide mathematical context, showing how analogous theories can be built in parts of mathematics far away from metric geometry. The reader interested only in geometry can, however, avoid these general considerations without logical harm, and begin at Section 2.

A topological space is not guaranteed to have a well-defined Euler characteristic unless it satisfies some finiteness condition. Similarly, the magnitude of an enriched category is defined under an assumption of finiteness; specializing to metric spaces, the definition of magnitude is just for *finite* spaces (Section 2). The magnitude of a finite metric space can be thought of as the 'effective number of points'. It deserves study partly because of its intrinsic interest, partly because of its applications to the measurement of diversity, and partly because it is used in the theory of magnitude of *infinite* metric spaces.

While categorical arguments do not (yet) furnish a definition of the magnitude of an infinite space, several methods for passing from finite to infinite immediately suggest themselves. Meckes [31] has shown that they are largely equivalent. Using the most elementary such method, coupled with some Fourier analysis, we produce evidence for the following conjectural principle:

magnitude encodes all the most important invariants of integral geometry

(Section 3). The most basic instance of this principle is the fact that a line segment of length t has magnitude $1 + t/2$, enabling one to recover length from magnitude. Less basic is the notion of the *magnitude dimension* of a space A, defined as the growth of the function $t \mapsto |tA|$; here tA is A scaled up by a factor of t, and $|tA|$ is its magnitude. We show, for example, that a subset of $\mathbb{R}^N$ with nonzero Lebesgue measure has magnitude dimension N. Magnitude dimension also appears to behave sensibly for fractals: for instance, Theorem 11 of [27] implies that the magnitude dimension of the ternary Cantor set is the same as its Hausdorff dimension (namely, $\log_3 2$).

It seems, moreover, that for any convex subset A of Euclidean space, all of the intrinsic volumes of A can be recovered from the function $t \mapsto |tA|$. This was first conjectured in [27], and appears below as Conjecture 3.5.10. In two dimensions, the conjectured formula is

$$|tA| = \tfrac{1}{2\pi}\operatorname{area}(A) \cdot t^2 + \tfrac{1}{4}\operatorname{perimeter}(A) \cdot t + \chi(A).$$

This resembles the theorem of Willerton [51] that for a compact homogeneous Riemannian 2-manifold A,

$$|tA| = \tfrac{1}{2\pi}\operatorname{area}(A) \cdot t^2 + \chi(A) + O(t^{-2})$$

as $t \to \infty$. This in turn resembles the celebrated tube formula of Weyl.

Review sections provide the necessary background on both enriched categories and integral geometry. No expertise in category theory or integral geometry is needed to read this paper.

Related work The basic ideas of this paper were first written up in a 2008 internet posting [22]. Several papers have already built on this. Leinster and Willerton [27] studied the large-scale asymptotics of the magnitude of subsets of Euclidean space, and stated the conjecture just mentioned. The precise form of that conjecture was motivated by numerical evidence and heuristic arguments found by Willerton [50]. Leinster [23] established magnitude as maximum diversity. Meckes [31] proved, *inter alia*, the equivalence of several definitions of the magnitude of compact metric spaces, and by using more subtle analytical methods than are used here, extended some of the results of Section 3 below. The magnitude of spheres is especially well understood [27, 51, 31].

In the literature on quantifying biodiversity, magnitude appears not only in the paper of Solow and Polasky [44], but also in later papers such as [38]. The approach to biodiversity measurement taken in [26] arose from the theory in the present paper. This relationship is explored further in [23] and Section 2 of [31].

Geometry as the study of metric structures is developed in the books of Blumenthal [4] and Gromov [11], among others; representatives of the theory of finite metric spaces are [4] and papers of Dress and collaborators [2, 6]. We will make contact with the theory of spaces of negative type, which goes back to Menger [32] and Schoenberg [43]. This connection has been exploited by Meckes [31]. It is notable that the complete bipartite graph $K_{3,2}$ appears as a minimal example in both [2] and Example 2.2.7 below.

NOTATION Given $N \in \mathbb{N} = \{0, 1, 2, \ldots\}$, we write $\mathbb{R}^N$ for real N-dimensional space as a set, topological space or vector space—but with no implied choice of metric except when $N = 1$. The metric on a metric space A is denoted by d or d_A. We write $\#X$ for the cardinality of a finite set X. When $\mathscr{C}$ is a category, $C \in \mathscr{C}$ means that C is an object of $\mathscr{C}$.

ACKNOWLEDGEMENTS I have had countless useful conversations on magnitude with Mark Meckes and Simon Willerton. Their insights have played a very important role, and I thank them for it. I also thank John Baez, Neal Bez, Paul Blackwell, Yemon Choi, Christina Cobbold, David Corfield, Alastair Craw, Jacques Distler, Joe Fu, Anton Geraschenko, Martin Hyland, David Jordan, André Joyal, Joachim Kock, Christian Korff, Urs Schreiber, Josh Shadlen, Ivan Smith, David Speyer, and Terry Tao. Two web resources have been crucial to the progress of this work: The n-Category Café[1] and MathOverflow.[2] Parts of this work were carried out at the Centre de Recerca Matemàtica (Barcelona) and the School of Mathematics and Statistics at the University of Sheffield. I thank them for their hospitality. This work was supported by an EPSRC Advanced Research Fellowship.

1 ENRICHED CATEGORIES

This section describes the conceptual origins of the notion of magnitude.

We define the magnitude of an enriched category, in two steps. First we assign a number to every matrix; then we assign a matrix to every enriched category. We pause in between to recall some basic aspects of enriched category theory: the definitions, and how a metric space can be viewed as an enriched category.

1.1 THE MAGNITUDE OF A MATRIX

A *rig* (or semiring) is a ri*n*g without *n*egatives: a set k equipped with a commutative monoid structure $(+, 0)$ and a monoid structure $(\cdot, 1)$, the latter distributing over the former. For us, rig will mean *commutative* rig: one whose multiplication is commutative.

It will be convenient to use matrices whose rows and columns are indexed by abstract finite sets. Thus, for finite sets I and J, an $I \times J$ *matrix* over a rig k is a function $I \times J \to k$. The usual operations can be performed, e.g. an $H \times I$ matrix can be multiplied by an $I \times J$ matrix to give an $H \times J$ matrix. The

[1] http://golem.ph.utexas.edu/category
[2] http://mathoverflow.net

identity matrix is the Kronecker δ. An $I \times J$ matrix ζ has a $J \times I$ transpose ζ^*.

Given a finite set I, we write $u_I \in k^I$ for the column vector with $u_I(i) = 1$ for all $i \in I$.

DEFINITION 1.1.1 Let ζ be an $I \times J$ matrix over a rig k. A *weighting* on ζ is a column vector $w \in k^J$ such that $\zeta w = u_I$. A *coweighting* on ζ is a row vector $v \in k^I$ such that $v\zeta = u_J^*$.

A matrix may admit zero, one, or many (co)weightings, but their freedom is subject to the following constraint.

LEMMA 1.1.2 *Let ζ be an $I \times J$ matrix over a rig, let w be a weighting on ζ, and let v be a coweighting on ζ. Then*

$$\sum_{j \in J} w(j) = \sum_{i \in I} v(i).$$

PROOF $\sum_j w(j) = u_J^* w = v\zeta w = v u_I = \sum_i v(i)$. $\square$

We refer to the entries $w(j) \in k$ of a weighting w as *weights*, and similarly *coweights*. The lemma implies that if a matrix ζ has both a weighting and a coweighting, then the total weight is independent of the weighting chosen. This makes the following definition possible.

DEFINITION 1.1.3 A matrix ζ over a rig k *has magnitude* if it admits at least one weighting and at least one coweighting. Its *magnitude* is then

$$|\zeta| = \sum_j w(j) = \sum_i v(i) \in k$$

for any weighting w and coweighting v on ζ.

We will be concerned with square matrices ζ. If ζ is invertible then there are a unique weighting and a unique coweighting. (Conversely, if k is a field then a unique weighting or coweighting implies invertibility.) The weights are then the sums of the rows of ζ^{-1}, and the coweights are the sums of the columns. Lemma 1.1.2 is obvious in this case, and there is an easy formula for the magnitude:

LEMMA 1.1.4 *Let ζ be an invertible $I \times I$ matrix over a rig. Then ζ has a unique weighting w, given by $w(j) = \sum_i \zeta^{-1}(j,i)$ $(j \in I)$, and a unique coweighting given by the dual formula. Moreover,*

$$|\zeta| = \sum_{i,j \in I} \zeta^{-1}(j,i).$$

$\square$

Often our matrix ζ will be symmetric, in which case weightings and coweightings are essentially the same.

1.2 Background on enriched categories

Here we review two standard notions: monoidal category, and category enriched in a monoidal category.

A *monoidal category* is a category $\mathscr{V}$ equipped with an associative binary operation $\otimes$ (which is formally a functor $\mathscr{V} \times \mathscr{V} \to \mathscr{V}$) and a unit object $\mathbb{1} \in \mathscr{V}$. The associativity and unit axioms are only required to hold up to suitably coherent isomorphism; see [28] for details.

Examples 1.2.1 i. $\mathscr{V}$ is the category **Set** of sets, $\otimes$ is cartesian product $\times$, and $\mathbb{1}$ is a one-element set $\{\star\}$.

ii. $\mathscr{V}$ is the category **Vect** of vector spaces over some field K, the product $\otimes$ is the usual tensor product $\otimes_K$, and $\mathbb{1} = K$.

iii. A poset can be viewed as a category in which each hom-set has at most one element. In particular, consider the poset $([0, \infty], \geq)$ of nonnegative reals together with infinity. The objects of the resulting category are the elements of $[0, \infty]$, there is one map $x \to y$ when $x \geq y$, and there are none otherwise. This is a monoidal category with $\otimes = +$ and $\mathbb{1} = 0$.

iv. Let **2** be the category of Boolean truth values [19]: there are two objects, $\mathtt{f}$ ('false') and $\mathtt{t}$ ('true'), and a single non-identity map, $\mathtt{f} \to \mathtt{t}$. Taking $\otimes$ to be conjunction and $\mathbb{1} = \mathtt{t}$ makes **2** monoidal. Then **2** is a monoidal subcategory of **Set**, identifying $\mathtt{f}$ with $\emptyset$ and $\mathtt{t}$ with $\{\star\}$. It is also a monoidal subcategory of $[0, \infty]$, identifying $\mathtt{f}$ with ∞ and $\mathtt{t}$ with 0.

Let $\mathscr{V} = (\mathscr{V}, \otimes, \mathbb{1})$ be a monoidal category. The definition of category enriched in $\mathscr{V}$, or $\mathscr{V}$-category, is obtained from the definition of ordinary category by asking that the hom-sets are no longer sets but objects of $\mathscr{V}$. Thus, a (small) *$\mathscr{V}$-category* $\mathbf{A}$ consists of a set $\mathrm{ob}\,\mathbf{A}$ of objects, an object $\mathrm{Hom}(a, b)$ of $\mathscr{V}$ for each $a, b \in \mathrm{ob}\,\mathbf{A}$, and operations of composition and identity satisfying appropriate axioms [16]. The operation of composition consists of a map

$$\mathrm{Hom}(a, b) \otimes \mathrm{Hom}(b, c) \to \mathrm{Hom}(a, c)$$

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

There is an accompanying notion of enriched functor. Given $\mathscr{V}$-categories $\mathbf{A}$ and $\mathbf{A}'$, a *$\mathscr{V}$-functor* $F \colon \mathbf{A} \to \mathbf{A}'$ consists of a function $\mathrm{ob}\,\mathbf{A} \to \mathrm{ob}\,\mathbf{A}'$, written $a \mapsto F(a)$, together with a map

$$\mathrm{Hom}(a, b) \to \mathrm{Hom}(F(a), F(b))$$

in $\mathscr{V}$ for each $a, b \in \mathrm{ob}\,\mathbf{A}$, satisfying suitable axioms [16]. We write $\mathscr{V}$-**Cat** for the category of $\mathscr{V}$-categories and $\mathscr{V}$-functors.

Examples 1.2.2 i. Let $\mathscr{V} = $ **Set**. Then $\mathscr{V}$-**Cat** is the category **Cat** of (small) categories and functors.

ii. Let $\mathscr{V} = \mathbf{Vect}$. Then $\mathscr{V}$-**Cat** is the category of *linear categories* or *algebroids*: categories equipped with a vector space structure on each hom-set, such that composition is bilinear.

iii. Let $\mathscr{V} = [0, \infty]$. Then, as observed by Lawvere [19, 20], a $\mathscr{V}$-category is a *generalized metric space*. That is, a $\mathscr{V}$-category consists of a set A of objects or points together with, for each $a, b \in A$, a real number $\mathrm{Hom}(a, b) = d(a, b) \in [0, \infty]$, satisfying the axioms

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

$(a, b, c \in A)$. Such spaces are more general than classical metric spaces in three ways: ∞ is permitted as a distance, the separation axiom $d(a, b) = 0 \Rightarrow a = b$ is dropped, and, most significantly, the symmetry axiom $d(a, b) = d(b, a)$ is dropped.

A $\mathscr{V}$-functor $f\colon A \to A'$ between generalized metric spaces A and A' is a *distance-decreasing map*: one satisfying $d(a, b) \geq d(f(a), f(b))$ for all $a, b \in A$. Hence $[0, \infty]$-**Cat** is the category **MS** of generalized metric spaces and distance-decreasing maps. The isomorphisms in **MS** are the isometries.

iv. Let $\mathscr{V} = \mathbf{2}$. A $\mathscr{V}$-category is a set equipped with a preorder (a reflexive transitive relation), which up to equivalence of $\mathscr{V}$-categories is the same thing as a poset.

The embedding $\mathbf{2} \hookrightarrow \mathbf{Set}$ of monoidal categories induces an embedding $\mathbf{2}$-**Cat** $\hookrightarrow$ **Set**-**Cat**; this is the embedding **Poset** $\hookrightarrow$ **Cat** of Example 1.2.1(iii). Similarly, the embedding $\mathbf{2} \hookrightarrow [0, \infty]$ induces an embedding **Poset** $\hookrightarrow$ **MS**: as observed in [19], a poset can be understood as a non-symmetric metric space whose distances are all 0 or ∞.

1.3 The magnitude of an enriched category

Here we meet the definition on which the rest of this work is built.

Having already defined the magnitude of a matrix, we now assign a matrix to each enriched category. To do this, we assume some further structure on the base category $\mathscr{V}$. In fact, we assume that we have a notion of size for *objects of* $\mathscr{V}$. This, then, will lead to a notion of size for *categories enriched in* $\mathscr{V}$.

Let $\mathscr{V}$ be a monoidal category. We will suppose given a rig k and a monoid homomorphism

$$|\cdot|\colon (\mathrm{ob}\, \mathscr{V} / \cong, \otimes, \mathbb{1}) \to (k, \cdot, 1).$$

(This is, deliberately, the same symbol as for magnitude; no confusion should arise.) The domain here is the monoid of isomorphism classes of objects of $\mathscr{V}$.

Examples 1.3.1 i. When $\mathscr{V}$ is the monoidal category **FinSet** of finite sets, we take $k = \mathbb{Q}$ and $|X| = \#X$.

ii. When $\mathscr{V}$ is the monoidal category **FDVect** of finite-dimensional vector spaces, we take $k = \mathbb{Q}$ and $|X| = \dim X$.

iii. When $\mathscr{V} = [0, \infty]$, we take $k = \mathbb{R}$ and $|x| = e^{-x}$. (If $|\cdot|$ is to be measurable[3] then the only possibility is $|x| = C^x$ for some constant $C \geq 0$.)

iv. When $\mathscr{V} = \mathbf{2}$, we take $k = \mathbb{Z}$, $|\mathtt{f}| = 0$ and $|\mathtt{t}| = 1$. This is a restriction of the functions $|\cdot|$ of (i) and (iii) along the embeddings $\mathbf{2} \hookrightarrow \mathbf{FinSet}$ and $\mathbf{2} \hookrightarrow [0, \infty]$ of Example 1.2.1(iv).

Write $\mathscr{V}$**-cat** (with a small 'c') for the category whose objects are the $\mathscr{V}$-categories with *finite* object-sets and whose maps are the $\mathscr{V}$-functors between them.

DEFINITION 1.3.2 Let $\mathbf{A} \in \mathscr{V}\text{-}\mathbf{cat}$.

i. The *similarity matrix* of $\mathbf{A}$ is the $\operatorname{ob}\mathbf{A} \times \operatorname{ob}\mathbf{A}$ matrix $\zeta_{\mathbf{A}}$ over k defined by $\zeta_{\mathbf{A}}(a, b) = |\mathrm{Hom}(a, b)|$ $(a, b \in \mathbf{A})$.

ii. A *(co)weighting* on $\mathbf{A}$ is a (co)weighting on $\zeta_{\mathbf{A}}$.

iii. $\mathbf{A}$ *has magnitude* if $\zeta_{\mathbf{A}}$ does; its *magnitude* is then $|\mathbf{A}| = |\zeta_{\mathbf{A}}|$.

iv. $\mathbf{A}$ *has Möbius inversion* if $\zeta_{\mathbf{A}}$ is invertible; its *Möbius matrix* is then $\mu_{\mathbf{A}} = \zeta_{\mathbf{A}}^{-1}$.

Magnitude is, then, a partially-defined function $|\cdot| \colon \mathscr{V}\text{-}\mathbf{cat} \dashrightarrow k$.

EXAMPLES 1.3.3 i. When $\mathscr{V} = \mathbf{FinSet}$, we obtain a notion of the magnitude $|\mathbf{A}| \in \mathbb{Q}$ of any (suitable) finite category $\mathbf{A}$. This is also called the *Euler characteristic* of $\mathbf{A}$ and written as $\chi(\mathbf{A})$ [21]. There are theorems relating it to the Euler characteristics of topological spaces, of graphs, of posets and of orbifolds, the cardinality of sets, and the order of groups.

Very many finite categories have Möbius inversion, or are equivalent to some category with Möbius inversion; all such categories have Euler characteristic. (This includes all finite posets, groupoids, monoids, categories containing no nontrivial idempotents, and categories admitting an epi-mono factorization system.) The Möbius matrix $\mu_{\mathbf{A}}$ is a generalization of Rota's Möbius function for posets [40], which in turn generalizes the classical Möbius function on integers. For details, see [21], and for further material on Euler characteristic and Möbius inversion for categories, see Berger and Leinster [3], Fiore, Lück and Sauer [7, 8], Jacobsen and Møller [14], Leinster [25], and Noguchi [33, 34, 35, 36].

[3]I thank Mark Meckes for pointing out that the more obvious hypothesis of continuity can be weakened to measurability [9]. In fact it suffices to assume that $|\cdot|$ is bounded on some set of positive measure [18].

ii. Similarly, taking $\mathscr{V} = \mathbf{FDVect}$ gives an invariant $\chi(\mathbf{A}) = |\mathbf{A}| \in \mathbb{Q}$ of linear categories $\mathbf{A}$ with finitely many objects and finite-dimensional hom-spaces.

iii. Taking $\mathscr{V} = [0, \infty]$ gives the notion of the magnitude $|A| \in \mathbb{R}$ of a (generalized) finite metric space A. This is the main subject of this paper.

iv. Taking $\mathscr{V} = \mathbf{2}$ gives a notion of the magnitude $|A| \in \mathbb{Z}$ of a finite poset A. Under the name of Euler characteristic, this goes back to Rota [40]; see [46] for a modern account. It is *always* defined. Indeed, every poset has Möbius inversion, and the Möbius matrix is the Möbius function of Rota mentioned in (i).

    We have noted that a poset can be viewed as a category, or alternatively as a generalized metric space. The three notions of magnitude are compatible: the magnitude of a poset is the same as that of the corresponding category or generalized metric space.

v. Let $\mathscr{V}$ be a category of topological spaces in which every object has a well-defined Euler characteristic (e.g. finite CW-complexes). Taking $|X|$ to be the Euler characteristic of a space X, we obtain a notion of the magnitude or Euler characteristic of a topologically-enriched category.

The definition of the magnitude of a $\mathscr{V}$-category $\mathbf{A}$ is independent of the composition and identities in $\mathbf{A}$, so could equally well be made in the generality of $\mathscr{V}$-graphs. (A *$\mathscr{V}$-graph* $\mathbf{G}$ is a set $\mathrm{ob}\,\mathbf{G}$ of objects together with, for each $a, b \in \mathrm{ob}\,\mathbf{G}$, an object $\mathrm{Hom}(a, b)$ of $\mathscr{V}$.) However, it is not clear that it is fruitful to do so. Two theorems on the magnitude or Euler characteristic of ordinary categories illuminate the general situation.

The first, Proposition 2.10 of [21], concerns directed graphs. The Euler characteristic of a category $\mathbf{A}$ is *not* in general equal to the Euler characteristic of its underlying graph $U(\mathbf{A})$. But the functor U has a left adjoint F, assigning to a graph $\mathbf{G}$ the category $F(\mathbf{G})$ whose objects are the vertices and whose maps are the paths in $\mathbf{G}$. If $\mathbf{G}$ is finite and circuit-free then $F(\mathbf{G})$ is finite, and the theorem is that $\chi(F(\mathbf{G})) = \chi(\mathbf{G})$. So the Euler characteristics of categories and graphs are closely related, but not in the most obvious way.

The second theorem, Proposition 1.5 of [25], concerns the classifying space $|N\mathbf{A}|$ of a category $\mathbf{A}$ (the geometric realization of its simplicial nerve $N\mathbf{A}$). Under a suitable finiteness condition, the topological space $|N\mathbf{A}|$ has a well-defined Euler characteristic, and it is a theorem that $\chi(|N\mathbf{A}|) = \chi(\mathbf{A})$. It follows that if two categories have the same underlying graph but different compositions then their classifying spaces, although not usually homotopy equivalent, have the same Euler characteristic. So if we wish the Euler characteristic of a category to be defined in such a way that it is equal to the Euler characteristic of its classifying space, it is destined to be independent of composition.

(At least, this is the case under the finiteness condition concerned. This states that the simplicial nerve of $\mathbf{A}$ contains only finitely many nondegenerate simplices, as in Proposition 2.11 of [21]. In a different setting, with a different definition of Euler characteristic, Fiore, Lück and Sauer have found a pair of finite categories whose underlying graphs are the same but whose Euler characteristics are different: see [7], after Theorem 1.14.)

1.4 Properties

Much of ordinary category theory generalizes smoothly to enriched categories. This includes many of the properties of the Euler characteristic of categories [21]. We list some of those properties now, using the symbols $\mathscr{V}$, k and $|\cdot|$ as in the previous section.

There are notions of *adjunction* and *equivalence* between $\mathscr{V}$-categories [16], generalizing the case $\mathscr{V} = \mathbf{Set}$ of ordinary categories. We write $\simeq$ for equivalence of $\mathscr{V}$-categories.

Proposition 1.4.1 *Let $\mathbf{A}, \mathbf{B} \in \mathscr{V}$-$\mathbf{cat}$.*

i. If there exist adjoint $\mathscr{V}$-functors $\mathbf{A} \rightleftarrows \mathbf{B}$, and $\mathbf{A}$ and $\mathbf{B}$ have magnitude, then $|\mathbf{A}| = |\mathbf{B}|$.

ii. If $\mathbf{A} \simeq \mathbf{B}$, and $\mathbf{A}$ and $\mathbf{B}$ have magnitude, then $|\mathbf{A}| = |\mathbf{B}|$.

iii. If $\mathbf{A} \simeq \mathbf{B}$ and $n \cdot 1 \in k$ has a multiplicative inverse for all positive integers n, then $\mathbf{A}$ has magnitude if and only if $\mathbf{B}$ does.

Proof Part (i) has the same proof as Proposition 2.4(a) of [21], and part (ii) follows immediately. Part (iii) has the same proof as Lemma 1.12 of [21]. □

For example, take a generalized metric space A and adjoin a new point at distance zero from some existing point. Then the new space A' is equivalent to A. By Proposition 1.4.1, if A has magnitude then A' does too, and $|A| = |A'|$. On the other hand, the proposition is trivial for classical metric spaces A, B: if there is an adjunction between A and B (and in particular if $A \simeq B$) then in fact A and B are isometric.

So far, we have not used the multiplicativity of the function $|\cdot|$ on objects of $\mathscr{V}$. We now show that it implies a multiplicativity property of the function $|\cdot|$ on $\mathscr{V}$-categories.

Assume that the monoidal category $\mathscr{V}$ is *symmetric*, that is, equipped with an isomorphism $X \otimes Y \to Y \otimes X$ for each pair X, Y of objects, satisfying axioms [28]. There is a product on $\mathscr{V}$-$\mathbf{Cat}$, also denoted by $\otimes$, defined as follows. Let $\mathbf{A}, \mathbf{B} \in \mathscr{V}$-$\mathbf{Cat}$. Then $\mathbf{A} \otimes \mathbf{B}$ is the $\mathscr{V}$-category whose object-set is $\operatorname{ob} \mathbf{A} \times \operatorname{ob} \mathbf{B}$ and whose hom-objects are given by

$$\operatorname{Hom}((a,b),(a',b')) = \operatorname{Hom}(a,a') \otimes \operatorname{Hom}(b,b').$$

Composition is defined with the aid of the symmetry [16]. The unit for this product is the one-object $\mathscr{V}$-category $\mathbf{I}$ whose single hom-object is $\mathbb{1} \in \mathscr{V}$.

EXAMPLES 1.4.2 i. When $\mathscr{V} = \mathbf{Set}$, this is the ordinary product $\times$ of categories.

ii. There is a one-parameter family of products on metric spaces. For $1 \leq p \leq \infty$ and metric spaces A and B, let $A \otimes_p B$ be the metric space whose point-set is the product of the point-sets of A and B, and whose distances are given by

$$d((a,b),(a',b')) = \begin{cases} \bigl(d(a,a')^p + d(b,b')^p\bigr)^{1/p} & \text{if } p < \infty \\ \max\{d(a,a'), d(b,b')\} & \text{if } p = \infty. \end{cases}$$

Then the tensor product $\otimes$ defined above is $\otimes_1$.

PROPOSITION 1.4.3 *Let $\mathbf{A}, \mathbf{B} \in \mathscr{V}$-$\mathbf{cat}$. If $\mathbf{A}$ and $\mathbf{B}$ have magnitude then so does $\mathbf{A} \otimes \mathbf{B}$, with*

$$|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A}||\mathbf{B}|.$$

Furthermore, the unit $\mathscr{V}$-category $\mathbf{I}$ has magnitude 1.

PROOF As for Proposition 2.6 of [21]. □

Magnitude is therefore a partially-defined monoid homomorphism

$$|\cdot|\colon (\mathscr{V}\text{-}\mathbf{cat}/\simeq, \otimes, \mathbf{I}) \dashrightarrow (k, \cdot, 1).$$

Under mild assumptions, coproducts of $\mathscr{V}$-categories exist and interact well with magnitude. Indeed, assume that $\mathscr{V}$ has an initial object 0, with $X \otimes 0 \cong 0 \cong 0 \otimes X$ for all $X \in \mathscr{V}$. Then for any two $\mathscr{V}$-categories $\mathbf{A}$ and $\mathbf{B}$, the coproduct $\mathbf{A} + \mathbf{B}$ in $\mathscr{V}$-$\mathbf{Cat}$ exists. It is constructed by taking the disjoint union of $\mathbf{A}$ and $\mathbf{B}$ and setting $\mathrm{Hom}(a,b) = \mathrm{Hom}(b,a) = 0$ whenever $a \in \mathbf{A}$ and $b \in \mathbf{B}$. There is also an initial $\mathscr{V}$-category $\emptyset$, with no objects.

When $\mathscr{V} = [0,\infty]$, the coproduct of metric spaces A and B is their *distant union*, the disjoint union of A and B with $d(a,b) = d(b,a) = \infty$ whenever $a \in A$ and $b \in B$.

Assume also that $|0| = 0$, where the 0 on the left-hand side is the initial object of $\mathscr{V}$. This assumption and the previous ones hold in all of our examples.

PROPOSITION 1.4.4 *Let $\mathbf{A}, \mathbf{B} \in \mathscr{V}$-$\mathbf{cat}$. If $\mathbf{A}$ and $\mathbf{B}$ have magnitude then so does $\mathbf{A} + \mathbf{B}$, with*

$$|\mathbf{A} + \mathbf{B}| = |\mathbf{A}| + |\mathbf{B}|.$$

Furthermore, the initial $\mathscr{V}$-category $\emptyset$ has magnitude 0.

PROOF As for Proposition 2.6 of [21]. □

It might seem unsatisfactory that not every $\mathscr{V}$-category with finite object-set has magnitude. This can be resolved as follows.

Given $\mathbf{A} \in \mathscr{V}$-$\mathbf{cat}$, there is an evident notion of (co)weighting on $\mathbf{A}$ with values in a prescribed k-algebra. As in Lemma 1.1.2, the total weight is always equal to the total coweight. Let $R(\mathbf{A})$ be the free k-algebra containing a weighting w and a coweighting v for $\mathbf{A}$. Then $\sum_a w(a) = \sum_a v(a) = [\mathbf{A}]$, say. This is *always* defined, and we may call $[\mathbf{A}] \in R(\mathbf{A})$ the *formal magnitude* of $\mathbf{A}$.

A homomorphism ϕ from $R(\mathbf{A})$ to another k-algebra S amounts to a weighting and a coweighting for $\mathbf{A}$ in S; moreover, $\phi([\mathbf{A}]) \in S$ is independent of the homomorphism ϕ chosen. In particular, $\mathbf{A}$ has magnitude in the original sense if and only if there exists a k-algebra homomorphism $\phi\colon R(\mathbf{A}) \to k$. In that case, $|\mathbf{A}| = \phi([\mathbf{A}])$ for any such ϕ.

This may lead to a more conceptually satisfactory theory, but at a price: the magnitudes of different categories lie in different rigs, complicating results such as those of the present section. In any case, we say no more about this approach.

2 Finite metric spaces

The definition of the magnitude of a finite metric space is a special case of the definition for enriched categories. Its most basic properties are special cases of general results. But metric spaces have many features not possessed by enriched categories in general. By exploiting them, we uncover a geometrically rich theory.

A crucial feature of metric spaces is that they can be rescaled. When handed a space, we gain more information about it by considering the magnitudes of its rescaled brothers and sisters than by taking it in isolation. This information is encapsulated in the so-called magnitude function of the space.

For some spaces, the magnitude function exhibits wild behaviour: singularities, negative magnitude, and so on. But for geometrically orthodox spaces such as subsets of Euclidean space, it turns out to be rather tame. This is because they belong to the important class of 'positive definite' spaces. Positive definiteness will play a central role when we come to extend the definition of magnitude from finite to infinite spaces. It is explored thoroughly in the paper of Meckes [31], who also describes its relationship with the classical notion of negative type.

The term *metric space* will be used in its standard sense, except that ∞ is permitted as a distance. Many of our theorems do hold for the generalized metric spaces of Example 1.2.2(iii), with the same proofs; but to avoid cluttering the exposition, we leave it to the reader to discern which.

Throughout, we use matrices whose rows and columns are indexed by abstract finite sets (as in Section 1.1). The identity matrix is denoted by δ.

2.1 The magnitude of a finite metric space

We begin by restating the definitions from Section 1, without reference to enriched categories. Let A be a finite metric space. Its *similarity matrix* $\zeta_A \in \mathbb{R}^{A\times A}$ is defined by $\zeta_A(a,b) = e^{-d(a,b)}$ $(a, b \in A)$. A *weighting* on A is a function $w\colon A \to \mathbb{R}$ such that $\sum_b \zeta_A(a,b)w(b) = 1$ for all $a \in A$. The space A *has magnitude* if it admits at least one weighting; its *magnitude* is then $|A| = \sum_a w(a)$ for any weighting w, and is independent of the weighting chosen.

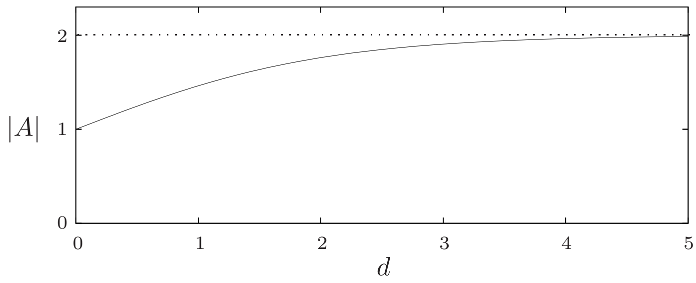

Figure 1: The magnitude of a two-point space

A finite metric space A *has Möbius inversion* if ζ_A is invertible. Its *Möbius matrix* is then $\mu_A = \zeta_A^{-1}$. In that case, there is a unique weighting w given by $w(a) = \sum_b \mu(a, b)$, and $|A| = \sum_{a,b} \mu_A(a, b)$ (Lemma 1.1.4). A generic real square matrix is invertible, which suggests that most finite metric spaces should have Möbius inversion. Proposition 2.2.6(i) makes this precise.
Here are some elementary examples.

Examples 2.1.1 i. The empty space has magnitude 0, and the one-point space has magnitude 1.

ii. Let A be the space consisting of two points distance d apart. Then

$$\zeta_A = \begin{pmatrix} 1 & e^{-d} \\ e^{-d} & 1 \end{pmatrix}.$$

This is invertible, so A has Möbius inversion and its magnitude is the sum of all four entries of $\mu_A = \zeta_A^{-1}$:

$$|A| = 1 + \tanh(d/2)$$

(Fig. 1). This can be interpreted as follows. When d is small, A closely resembles a 1-point space; correspondingly, the magnitude is little more than 1. As d grows, the points acquire increasingly separate identities and the magnitude increases. In the extreme, when $d = \infty$, the two points are entirely separate and the magnitude is 2.

iii. A metric space A is *discrete* [20] if $d(a, b) = \infty$ for all $a \neq b$ in A. Let A be a finite discrete space. Then ζ_A is the identity matrix δ, each point has weight 1, and $|A| = \#A$.

The definition of the magnitude of a metric space first appeared in a paper of Solow and Polasky [44], although with almost no mathematical development. They called it the 'effective number of species', since the points of their spaces represented biological species and the distances represented inter-species differences (e.g. genetic). We can view the magnitude of a metric space as the 'effective number of points'. Solow and Polasky also considered the magnitude of correlation matrices, making connections with the statistical concept of effective sample size.

Three-point spaces have magnitude; the formula follows from the proof of Proposition 2.4.15. Meckes [31, Theorem 3.6] has shown that four-point spaces have magnitude. But spaces with five or more points need not have magnitude (Example 2.2.7).

We now describe two classes of space for which the magnitude exists and is given by an explicit formula.

Definition 2.1.2 A finite metric space A is *scattered* if $d(a,b) > \log((\#A)-1)$ for all distinct points a and b. (Vacuously, the empty space and one-point space are scattered.)

Proposition 2.1.3 *A scattered space has magnitude. Indeed, any scattered space A has Möbius inversion, with Möbius matrix given by the infinite sum*

$$\mu_A(a,b) = \sum_{k=0}^{\infty} \sum_{a=a_0\neq\cdots\neq a_k=b} (-1)^k \zeta_A(a_0,a_1)\cdots\zeta_A(a_{k-1},a_k).$$

The inner sum is over all $a_0,\ldots,a_k \in A$ such that $a_0=a$, $a_k=b$, and $a_{j-1}\neq a_j$ whenever $1\leq j\leq k$. That a scattered space has magnitude was also proved in [27, Theorem 2], by a different method that does not produce a formula for the Möbius matrix.

Proof Write $n=\#A$. For $a,b\in A$ and $k\geq 0$, put

$$\mu_{A,k}(a,b) = \sum_{a=a_0\neq\cdots\neq a_k=b} \zeta_A(a_0,a_1)\cdots\zeta_A(a_{k-1},a_k).$$

(In particular, $\mu_{A,0}$ is the identity matrix.) Write $\varepsilon = \min_{a\neq b} d(a,b)$. Then

$$\begin{aligned}
\mu_{A,k+1}(a,b) &= \sum_{b' \colon b'\neq b} \sum_{a=a_0\neq\cdots\neq a_k=b'} \zeta_A(a_0,a_1)\cdots\zeta_A(a_{k-1},b')\zeta_A(b',b) \\
&\leq \sum_{b' \colon b'\neq b} \sum_{a=a_0\neq\cdots\neq a_k=b'} \zeta_A(a_0,a_1)\cdots\zeta_A(a_{k-1},b')e^{-\varepsilon} \\
&= e^{-\varepsilon} \sum_{b' \colon b'\neq b} \mu_{A,k}(a,b').
\end{aligned}$$

The last sum is over $(n-1)$ terms, so by induction, $\mu_{A,k}(a,b) \leq \big((n-1)e^{-\varepsilon}\big)^k$ for all $a,b\in A$ and $k\geq 0$. But A is scattered, so $(n-1)e^{-\varepsilon}<1$, so the sum $\sum_{k=0}^{\infty}(-1)^k\mu_{A,k}(a,b)$ converges for all $a,b\in A$. A telescoping sum argument finishes the proof. $\square$

Definition 2.1.4 A metric space is *homogeneous* if its isometry group acts transitively on points.

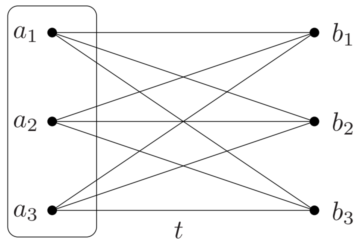

Figure 2: $tK_{n,n}$ and its subspace $2tK_n$, shown for $n = 3$.

Proposition 2.1.5 (Speyer [45]) *Every homogeneous finite metric space has magnitude. Indeed, if A is a homogeneous space with $n \geq 1$ points then*

$$|A| = \frac{n^2}{\sum_{a,b} e^{-d(a,b)}} = \frac{n}{\sum_a e^{-d(x,a)}}$$

for any $x \in A$. There is a weighting w on A given by $w(a) = |A|/n$ for all $a \in A$.

Proof By homogeneity, the sum $S = \sum_a \zeta_A(x,a)$ is independent of $x \in A$. Hence there is a weighting w given by $w(a) = 1/S$ for all $a \in A$. □

Example 2.1.6 For any (undirected) graph G and $t \in (0, \infty]$, there is a metric space tG whose points are the vertices and whose distances are minimal path-lengths, a single edge having length t. Write K_n for the complete graph on n vertices. Then

$$|tK_n| = \frac{n}{1 + (n-1)e^{-t}}.$$

In general, $e^{-d(a,b)}$ can be interpreted as the similarity or closeness of the points $a, b \in A$ [26, 44]. Proposition 2.1.5 states that the magnitude of a homogeneous space is the reciprocal mean similarity.

Example 2.1.7 A subspace can have greater magnitude than the whole space. Let $K_{n,m}$ be the graph with vertices $a_1, \ldots, a_n, b_1, \ldots, b_m$ and one edge between a_i and b_j for each i and j. If n is large then the mean similarity between two points of $tK_{n,n}$ is approximately $\frac{1}{2}(e^{-t} + e^{-2t})$ (Fig. 2). On the other hand, $tK_{n,n}$ has a subspace $2tK_n = \{a_1, \ldots, a_n\}$ in which the mean similarity is approximately e^{-2t}. Since $e^{-t} > e^{-2t}$, the mean similarity between points of $tK_{n,n}$ is greater than that of its subspace $2tK_n$; hence $|tK_{n,n}| < |2tK_n|$. In fact, it can be shown using Proposition 2.1.5 that $|tK_{n,n}| < |2tK_n|$ whenever $n > e^t + 1$.

2.2 MAGNITUDE FUNCTIONS

In physical situations, distance depends on the choice of unit of length; making a different choice rescales the metric by a constant factor. In the definition of $|x|$ as e^{-x} (Example 1.3.1(iii)), the constant e^{-1} was chosen without justification; choosing a different constant between 0 and 1 also amounts to rescaling the metric. For both these reasons, every metric space should be seen as a member of the one-parameter family of spaces obtained by rescaling it.

DEFINITION 2.2.1 Let A be a metric space and $t \in (0, \infty)$. Then tA denotes the metric space with the same points as A and $d_{tA}(a, b) = td_A(a, b)$ $(a, b \in A)$.

Most familiar invariants of metric spaces behave in a predictable way when the space is rescaled. This is true, for example, of topological invariants, diameter, and Hausdorff measure of any dimension. But magnitude does not behave predictably under rescaling. Graphing $|tA|$ against t therefore gives more information about A than is given by $|A|$ alone.

DEFINITION 2.2.2 Let A be a finite metric space. The *magnitude function* of A is the partially-defined function $t \mapsto |tA|$, defined for all $t \in (0, \infty)$ such that tA has magnitude.

EXAMPLES 2.2.3 i. Let A be the space consisting of two points distance d apart. By Example 2.1.1(ii), the magnitude function of A is defined everywhere and given by $t \mapsto 1 + \tanh(dt/2)$.

ii. Let $A = \{a_1, \ldots, a_n\}$ be a nonempty homogeneous space, and write $E_i = d(a_1, a_i)$. By Proposition 2.1.5, the magnitude function of A is

$$t \mapsto n \Big/ \sum_{i=1}^{n} e^{-E_i t}.$$

In the terminology of statistical mechanics, the denominator is the partition function for the energies E_i at inverse temperature t.[4]

iii. Let R be a finite commutative ring. For $a \in R$, write

$$\nu(a) = \min\{k \in \mathbb{N} : a^{k+1} = 0\} \in \mathbb{N} \cup \{\infty\}.$$

There is a metric d on R given by $d(a, b) = \nu(b - a)$, and the resulting metric space A_R is homogeneous. Write $q = e^{-t}$, and $\mathrm{Nil}(R)$ for the ideal of nilpotent elements. By Proposition 2.1.5, A_R has magnitude function

$$t \mapsto |tA_R| = \#R \Big/ \sum_{a \in \mathrm{Nil}(R)} q^{\nu(a)} = \#R \Big/ (1-q) \sum_{k=0}^{\infty} \#\{a \in R : a^{k+1} = 0\} \cdot q^k$$

where the last expression is an element of the field $\mathbb{Q}((q))$ of formal Laurent series.

[4] I thank Simon Willerton for suggesting that some such relationship should exist.

To establish the basic properties of magnitude functions, we need some auxiliary definitions and a lemma. A vector $v \in \mathbb{R}^I$ is *positive* if $v(i) > 0$ for all $i \in I$, and *nonnegative* if $v(i) \geq 0$ for all $i \in I$. Recall the definition of distance-decreasing map from Example 1.2.2(iii).

DEFINITION 2.2.4 A metric space A is an *expansion* of a metric space B if there exists a distance-decreasing surjection $A \to B$.

LEMMA 2.2.5 *Let A and B be finite metric spaces, each admitting a nonnegative weighting. If A is an expansion of B then $|A| \geq |B|$.*

PROOF Take a distance-decreasing surjection $f\colon A \to B$. Choose a right inverse function $g\colon B \to A$ (not necessarily distance-decreasing). Then $\zeta_B(f(a), b) \geq \zeta_A(a, g(b))$ for all $a \in A$ and $b \in B$. Let w_A and w_B be nonnegative weightings on A and B respectively. Then

$$|A| = \sum_{a,b} w_A(a)\zeta_B(f(a), b)w_B(b) \geq \sum_{a,b} w_A(a)\zeta_A(a, g(b))w_B(b) = |B|,$$

as required. $\square$

PROPOSITION 2.2.6 *Let A be a finite metric space. Then:*

i. tA has Möbius inversion (and therefore magnitude) for all but finitely many $t > 0$.

ii. The magnitude function of A is analytic at all $t > 0$ such that tA has Möbius inversion.

iii. For $t \gg 0$, there is a unique, positive, weighting on tA.

iv. For $t \gg 0$, the magnitude function of A is increasing.

v. $|tA| \to \#A$ as $t \to \infty$.

PROOF We use the space $\mathbb{R}^{A\times A}$ of real $A \times A$ matrices, and its open subset $\mathrm{GL}(A)$ of invertible matrices. We also use the notions of weighting on, and magnitude of, a matrix (Section 1.1). For $\zeta \in \mathrm{GL}(A)$, the unique weighting w_ζ on ζ and the magnitude of ζ are given by

$$w_\zeta(a) = \sum_{b\in A} \zeta^{-1}(a, b) = \sum_{b\in A} (\operatorname{adj}\zeta)(a, b)/\det\zeta, \qquad |\zeta| = \sum_{a\in A} w_\zeta(a) \qquad (1)$$

($a \in A$), where adj denotes the adjugate.

For (i), first note that $\zeta_{tA} \to \delta \in \mathrm{GL}(A)$ as $t \to \infty$; hence ζ_{tA} is invertible for $t \gg 0$. The matrix $\zeta_{tA} = (e^{-td(a,b)})$ is defined for all $t \in \mathbb{C}$, and $\det\zeta_{tA}$ is analytic in t. But $\det\zeta_{tA} \neq 0$ for real $t \gg 0$, so by analyticity, $\det\zeta_{tA}$ has only finitely many zeros in $(0, \infty)$.

Part (ii) follows from equations (1).

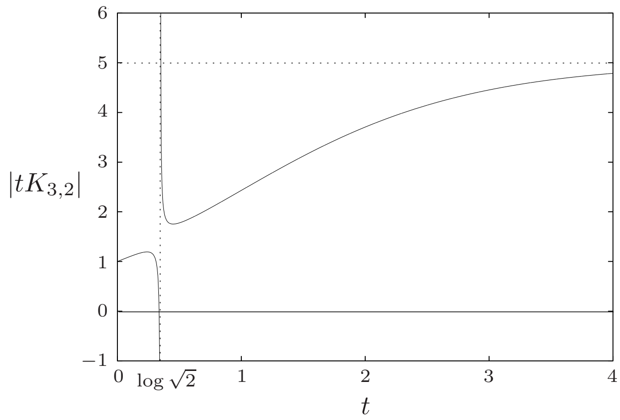

Figure 3: The magnitude function of the bipartite graph $K_{3,2}$

For (iii), each of the functions $\zeta \mapsto w_\zeta(a)$ $(a \in A)$ is continuous on $\mathrm{GL}(A)$ by (1). But $w_\delta(a) = 1$ for all $a \in A$, so there is a neighbourhood U of δ in $\mathrm{GL}(A)$ such that $w_\zeta(a) > 0$ for all $\zeta \in U$ and $a \in A$. Since $\zeta_{tA} \to \delta$ as $t \to \infty$, we have $\zeta_{tA} \in U$ for all $t \gg 0$.

Part (iv) follows from part (iii) and Lemma 2.2.5.

For (v), $\lim_{t\to\infty} |tA| = |\lim_{t\to\infty} \zeta_{tA}| = |\delta| = \#A$. □

Part (i) implies that magnitude functions have only finitely many singularities. Proposition 2.4.17 will provide an explicit lower bound for parts (iii) and (iv). Part (v) also appeared as Theorem 3 of [27].

Many natural conjectures about magnitude are disproved by the following example. Later we will see that subspaces of Euclidean space are less prone to surprising behaviour.[5]

Example 2.2.7 Fig. 3 shows the magnitude function of the space $K_{3,2}$ defined in Example 2.1.7. It is given by

$$|tK_{3,2}| = \frac{5 - 7e^{-t}}{(1 + e^{-t})(1 - 2e^{-2t})}$$

$(t \neq \log\sqrt{2})$; the magnitude of $(\log\sqrt{2})K_{3,2}$ is undefined. (One can compute this directly or use Proposition 2.3.13.) Several features of the graph are apparent. At some scales, the magnitude is negative; at others, it is greater than the number of points. There are also intervals on which the magnitude function is strictly decreasing. Furthermore, this example shows that a space with magnitude can have a subspace without magnitude: for $(\log\sqrt{2})K_{3,2}$ is a subspace of $(\log\sqrt{2})K_{3,3}$, which, being homogeneous, has magnitude (Proposition 2.1.5).

[5]'Our approach to general metric spaces bears the undeniable imprint of early exposure to Euclidean geometry. We just love spaces sharing a common feature with $\mathbb{R}^n$.' (Gromov [11], page xvi.)

(The graph $K_{3,2}$ is also a well-known counterexample in the theory of spaces of negative type [10]. The connection is explained, in broad terms, by the remarks in Section 2.4.)

The first example of a finite metric space with undefined magnitude was found by Tao [48], and had 6 points. The first examples of n-point spaces with magnitude outside the interval $[0, n]$ were found by the author and Simon Willerton, and were again 6-point spaces.

EXAMPLE 2.2.8 This is an example of a space A for which $\lim_{t\to 0} |tA| \neq 1$, due to Willerton (personal communication, 2009). Let A be the graph $K_{3,3}$ (Fig. 2) with three new edges adjoined: one from b_i to b_j whenever $1 \leq i < j \leq 3$. Then $|tA| = 6/(1 + 4e^{-t}) \to 6/5$ as $t \to 0$.

2.3 New spaces from old

For each way of constructing a new metric space from old, we can ask: is the magnitude of the new space determined by the magnitudes of the old ones? Here we give a positive answer for four constructions: unions of a particular type, tensor products, fibrations, and constant-distance gluing.

Unions

Let X be a metric space with subspaces A and B. The magnitude of $A \cup B$ is not in general determined by the magnitudes of A, B and $A \cap B$: consider one-point spaces. In this respect, magnitude of metric spaces is unlike cardinality of sets, for which there is the inclusion-exclusion formula. We do, however, have an inclusion-exclusion formula for magnitude when the union is of a special type.

DEFINITION 2.3.1 Let X be a metric space and $A, B \subseteq X$. Then A *projects to* B if for all $a \in A$ there exists $\pi(a) \in A \cap B$ such that for all $b \in B$,

$$d(a, b) = d(a, \pi(a)) + d(\pi(a), b).$$

In this situation, $d(a, \pi(a)) = \inf_{b\in B} d(a, b)$. If all distances in X are finite then $\pi(a)$ is unique for a.

PROPOSITION 2.3.2 *Let X be a finite metric space and $A, B \subseteq X$. Suppose that A projects to B and B projects to A. If A, B and $A \cap B$ have magnitude then so does $A \cup B$, with*

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Indeed, if w_A, w_B and $w_{A\cap B}$ are weightings on A, B and $A \cap B$ respectively then there is a weighting w on $A \cup B$ defined by

$$w(x) = \begin{cases} w_A(x) & \textit{if } x \in A \setminus B \\ w_B(x) & \textit{if } x \in B \setminus A \\ w_A(x) + w_B(x) - w_{A\cap B}(x) & \textit{if } x \in A \cap B. \end{cases}$$

Proof Let $a \in A \setminus B$. Choose a point $\pi(a)$ as in Definition 2.3.1. Then

$$\begin{aligned}
&\sum_{x \in A \cup B} \zeta(a,x) w(x) \\
&= \sum_{a' \in A} \zeta(a,a') w_A(a') + \sum_{b \in B} \zeta(a,b) w_B(b) - \sum_{c \in A \cap B} \zeta(a,c) w_{A \cap B}(c) \\
&= 1 + \zeta(a,\pi(a)) \left\{ \sum_{b \in B} \zeta(\pi(a),b) w_B(b) - \sum_{c \in A \cap B} \zeta(\pi(a),c) w_{A \cap B}(c) \right\} = 1.
\end{aligned}$$

Similar arguments apply when we start with a point of $B \setminus A$ or $A \cap B$. This proves that w is a weighting, and the result follows. □

It can similarly be shown that if A, B and $A \cap B$ all have Möbius inversion then so does $A \cup B$. The proof is left to the reader; we just need the following special case.

Corollary 2.3.3 *Let X be a finite metric space and $A, B \subseteq X$. Suppose that $A \cap B$ is a singleton $\{c\}$, that for all $a \in A$ and $b \in B$,*

$$d(a,b) = d(a,c) + d(c,b),$$

and that A and B have magnitude. Then $A \cup B$ has magnitude $|A| + |B| - 1$. Moreover, if A and B have Möbius inversion then so does $A \cup B$, with

$$\mu_{A \cup B}(x,y) = \begin{cases} \mu_A(x,y) & \text{if } x,y \in A \text{ and } (x,y) \neq (c,c) \\ \mu_B(x,y) & \text{if } x,y \in B \text{ and } (x,y) \neq (c,c) \\ \mu_A(c,c) + \mu_B(c,c) - 1 & \text{if } (x,y) = (c,c) \\ 0 & \text{otherwise.} \end{cases}$$

Proof The first statement follows from Proposition 2.3.2, and the second is easily checked. □

Corollary 2.3.4 *Every finite subspace of $\mathbb{R}$ has Möbius inversion. If $A = \{a_0 < \cdots < a_n\} \subseteq \mathbb{R}$ then, writing $d_i = a_i - a_{i-1}$,*

$$|A| = 1 + \sum_{i=1}^{n} \tanh \frac{d_i}{2}.$$

The weighting w on A is given by

$$w(a_i) = \frac{1}{2} \left(\tanh \frac{d_i}{2} + \tanh \frac{d_{i+1}}{2} \right)$$

($0 \leq i \leq n$), where by convention $d_0 = d_{n+1} = \infty$ and $\tanh \infty = 1$.

Proof This follows by induction from Example 2.1.1(ii), Proposition 2.3.2 and Corollary 2.3.3. (An alternative proof is given in [27, Theorem 4].) □

Thus, in a finite subspace of $\mathbb{R}$, the weight of a point depends only on the distances to its neighbours. This is reminiscent of the Ising model in statistical mechanics [5], although whether there is a substantial connection remains to be seen.

EXAMPLE 2.3.5 The magnitude function is not a complete invariant of finite metric spaces. Indeed, let $X = \{0, 1, 2, 3\} \subseteq \mathbb{R}$. Let Y be the four-vertex Y-shaped graph, viewed as a metric space as in Example 2.1.6. I claim that X and Y have the same magnitude function, even though they are not isometric. For put $A = \{0, 1, 2\} \subseteq \mathbb{R}$ and $B = \{0, 1\} \subseteq \mathbb{R}$. Both tX and tY can be expressed as unions, satisfying the hypotheses of Corollary 2.3.3, of isometric copies of tA and tB. Hence $|tX| = |tA| + |tB| - 1 = |tY|$ for all $t > 0$.

TENSOR PRODUCTS

Recall from Example 1.4.2(ii) the definition of the tensor product of metric spaces. Proposition 1.4.3 implies (and it is easy to prove directly):

PROPOSITION 2.3.6 *If A and B are finite metric spaces with magnitude then $A \otimes B$ has magnitude, given by $|A \otimes B| = |A||B|$.* $\square$

EXAMPLE 2.3.7 Let q be a prime power, and denote by $\mathbb{F}_q$ the field of q elements metrized by $d(a, b) = 1$ whenever $a \neq b$. Then for $N \in \mathbb{N}$, the metric tensor product $\mathbb{F}_q^{\otimes N}$ is the set $\mathbb{F}_q^N$ with the Hamming metric. Its magnitude function is

$$t \mapsto |t\mathbb{F}_q|^N = \left(\frac{q}{1 + (q-1)e^{-t}} \right)^N$$

by Example 2.1.6 and Proposition 2.3.6.
More generally, a *linear code* is a vector subspace C of $\mathbb{F}_q^N$ [29]. Its (single-variable) *weight enumerator* is the polynomial $W_C(x) = \sum_{i=0}^N A_i(C)x^i \in \mathbb{Z}[x]$, where $A_i(C)$ is the number of elements of C whose Hamming distance from 0 is i. Since C is homogeneous, Proposition 2.1.5 implies that its magnitude function is

$$t \mapsto (\#C)/W_C(e^{-t}).$$

The magnitude function of a linear code therefore carries the same, important, information as its weight enumerator.

Similarly, if A and B are finite metric spaces with magnitude then their coproduct or distant union $A+B$ (Section 1.4) has magnitude $|A+B| = |A|+|B|$.

FIBRATIONS

A fundamental property of the Euler characteristic of topological spaces is its behaviour with respect to fibrations. If a space A is fibred over a connected base B, with fibre F, then under suitable hypotheses, $\chi(A) = \chi(B)\chi(F)$.

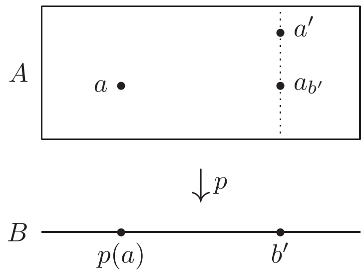

Figure 4: Metric fibration

An analogous formula holds for the Euler characteristic of a fibred category (Proposition 2.8 of [21] and, in a different context, Theorem 7.7 of [7]).
Apparently no general notion of fibration of enriched categories has yet been formulated. Nevertheless, we define here a notion of fibration of metric spaces sharing common features with the categorical and topological notions, and we prove an analogous theorem on magnitude.

DEFINITION 2.3.8 Let A and B be metric spaces. A *(metric) fibration* from A to B is a distance-decreasing map $p\colon A \to B$ with the following property (Fig. 4): for all $a \in A$ and $b' \in B$ with $d(p(a), b') < \infty$, there exists $a_{b'} \in p^{-1}(b')$ such that for all $a' \in p^{-1}(b')$,

$$d(a, a') = d(p(a), b') + d(a_{b'}, a'). \tag{2}$$

EXAMPLE 2.3.9 Let C_t be the circle of circumference t, metrized non-symmetrically by taking $d(a, b)$ to be the length of the anticlockwise arc from a to b. (This is a generalized metric space in the sense of Example 1.2.2(iii).) Let k be a positive integer. Then the k-fold covering $C_{kt} \to C_t$, locally an isometry, is a fibration.

LEMMA 2.3.10 *Let $p\colon A \to B$ be a fibration of metric spaces. Let $b, b' \in B$ with $d(b, b') < \infty$. Then the fibres $p^{-1}(b)$ and $p^{-1}(b')$ are isometric.*

PROOF Equation (2) and finiteness of $d(b, b')$ imply that $a_{b'}$ is unique for $a \in p^{-1}(b)$, so we may define a function $\gamma_{b,b'}\colon p^{-1}(b) \to p^{-1}(b')$ by $\gamma_{b,b'}(a) = a_{b'}$. It is distance-decreasing: for if $a, c \in p^{-1}(b)$ then

$$\begin{aligned} d(b, b') + d(\gamma_{b,b'}(a), \gamma_{b,b'}(c)) &= d(a, \gamma_{b,b'}(c)) \\ &\leq d(a, c) + d(c, \gamma_{b,b'}(c)) = d(a, c) + d(b, b'), \end{aligned}$$

giving $d(\gamma_{b,b'}(a), \gamma_{b,b'}(c)) \leq d(a, c)$ by finiteness of $d(b, b')$.
There is a distance-decreasing map $\gamma_{b',b}\colon p^{-1}(b') \to p^{-1}(b)$ defined in the same way. It is readily shown that $\gamma_{b,b'}$ and $\gamma_{b',b}$ are mutually inverse; hence they are isometries. □

Let B be a nonempty metric space all of whose distances are finite, and let $p\colon A \to B$ be a fibration. The *fibre* of p is any of the spaces $p^{-1}(b)$ $(b \in B)$; it is well-defined up to isometry.

Theorem 2.3.11 *Let $p\colon A \to B$ be a fibration of finite metric spaces. Suppose that B is nonempty with $d(b, b') < \infty$ for all $b, b' \in B$, and that B and the fibre F of p both have magnitude. Then A has magnitude, given by $|A| = |B||F|$.*

Proof Choose a weighting w_B on B. Choose, for each $b \in B$, a weighting w_b on the space $p^{-1}(b)$. For $a \in A$, put $w_A(a) = w_{p(a)}(a)w_B(p(a))$. It is straightforward to check that w_A is a weighting, and the theorem follows. □

Examples 2.3.12 i. A trivial example of a fibration is a product-projection $B \otimes F \to B$. In that case, Theorem 2.3.11 reduces to Proposition 2.3.6.

ii. Let B be a finite metric space in which the triangle inequality holds strictly for every triple of distinct points. Let F be a finite metric space of small diameter:

$$\operatorname{diam}(F) \leq \min\bigl\{d(b, b') + d(b', b'') - d(b, b'') : b, b', b'' \in B,\ b \neq b' \neq b''\bigr\}.$$

Choose for each $b, b' \in B$ an isometry $\gamma_{b,b'}\colon F \to F$, in such a way that $\gamma_{b,b}$ is the identity and $\gamma_{b',b} = \gamma_{b,b'}^{-1}$. Then the set $A = B \times F$ can be metrized by putting

$$d((b, c), (b', c')) = d(b, b') + d(\gamma_{b,b'}(c), c')$$

$(b, b' \in B,\ c, c' \in F)$. The projection $A \to B$ is a fibration (but not a product-projection unless $\gamma_{b',b''} \circ \gamma_{b,b'} = \gamma_{b,b''}$ for all b, b', b''). So if B and F have magnitude, $|A| = |B||F|$.

Arguments similar to Lemma 2.3.10 show that a fibration over B amounts to a family $(A_b)_{b \in B}$ of metric spaces together with a distance-decreasing map $\gamma_{b,b'}\colon A_b \to A_{b'}$ for each $b, b' \in B$ such that $d(b, b') < \infty$, satisfying the following three conditions. First, $\gamma_{b,b}$ is the identity for all $b \in B$. Second, $\gamma_{b',b} = \gamma_{b,b'}^{-1}$. Third,

$$\sup_{a \in A_b} d\bigl(\gamma_{b',b''}\gamma_{b,b'}(a), \gamma_{b,b''}(a)\bigr) \leq d(b, b') + d(b', b'') - d(b, b'')$$

for all $b, b', b'' \in B$ such that $d(b, b'), d(b', b'') < \infty$.

Constant-distance gluing

Given metric spaces A and B and a real number $D \geq \max\{\operatorname{diam} A, \operatorname{diam} B\}/2$, there is a metric space $A +_D B$ defined as follows. As a set, it is the disjoint union of A and B. The metric restricted to A is the original metric on A; similarly for B; and $d(a, b) = d(b, a) = D$ for all $a \in A$ and $b \in B$.

Proposition 2.3.13 *Let A and B be finite metric spaces, and take D as above. Suppose that A and B have magnitude, with $|A||B| \neq e^{2D}$. Then $A +_D B$ has magnitude*

$$\frac{|A| + |B| - 2e^{-D}|A||B|}{1 - e^{-2D}|A||B|}.$$

Proof Given weightings w_A on A and w_B on B, there is a weighting w on $A +_D B$ defined by

$$w(a) = \frac{1 - e^{-D}|B|}{1 - e^{-2D}|A||B|} w_A(a), \qquad w(b) = \frac{1 - e^{-D}|A|}{1 - e^{-2D}|A||B|} w_B(b)$$

$(a \in A,\, b \in B)$. The result follows. □

This provides an easy way to compute the magnitude functions in Examples 2.2.7 and 2.2.8.

2.4 Positive definite spaces

We saw in Example 2.2.7 that the magnitude of a finite metric space may be undefined, or smaller than the magnitude of one of its subspaces, or even negative. We now introduce a class of spaces for which no such behaviour occurs. Very many spaces of interest—including all subsets of Euclidean space—belong to this class. It has been studied in greater depth by Meckes [31].

Definition 2.4.1 A finite metric space A is *positive definite* if the matrix ζ_A is positive definite.

We emphasize that positive definiteness of a matrix is meant in the *strict* sense.

Lemma 2.4.2 *i. A positive definite space has Möbius inversion.*

ii. The tensor product of positive definite spaces is positive definite.

iii. A subspace of a positive definite space is positive definite.

Proof Parts (i) and (iii) are elementary. For (ii), $\zeta_{A \otimes B}$ is the Kronecker product $\zeta_A \otimes \zeta_B$, and the Kronecker product of positive definite matrices is positive definite. □

In particular, a positive definite space has magnitude and a unique weighting.

Proposition 2.4.3 *Let A be a positive definite finite metric space. Then*

$$|A| = \sup_{v \neq 0} \frac{\left(\sum_{a \in A} v(a)\right)^2}{v^* \zeta_A v}$$

where the supremum is over $v \in \mathbb{R}^A \setminus \{0\}$ and v^* denotes the transpose of v. A vector v attains the supremum if and only if it is a nonzero scalar multiple of the unique weighting on A.

PROOF Since ζ_A is positive definite, we have the Cauchy–Schwarz inequality:

$$(v^*\zeta_A v)\cdot(w^*\zeta_A w) \geq (v^*\zeta_A w)^2$$

for all $v, w \in \mathbb{R}^A$, with equality if and only if one of v and w is a scalar multiple of the other. Taking w to be the unique weighting on A gives the result. $\square$

COROLLARY 2.4.4 *If A is a positive definite finite metric space and $B \subseteq A$, then $|B| \leq |A|$.* $\square$

COROLLARY 2.4.5 *A nonempty positive definite finite metric space has magnitude ≥ 1.* $\square$

For any finite metric space A, the set $\mathrm{Sing}(A) = \{t \in (0,\infty) : \zeta_{tA} \text{ is singular}\}$ is finite (Proposition 2.2.6(i)). When $\mathrm{Sing}(A) = \emptyset$, put $\sup(\mathrm{Sing}(A)) = 0$.

PROPOSITION 2.4.6 *Let A be a finite metric space. Then tA is positive definite for all $t > \sup(\mathrm{Sing}(A))$. In particular, tA is positive definite for all $t \gg 0$.*

PROOF Write $\lambda_{\min}(\xi)$ for the minimum eigenvalue of a real symmetric $A \times A$ matrix ξ. Then $\lambda_{\min}(\xi)$ is continuous in ξ. Also $\lambda_{\min}(\xi) > 0$ if and only if ξ is positive definite, and if $\lambda_{\min}(\xi) = 0$ then ξ is singular.
Now $\zeta_{tA} \to \delta$ as $t \to \infty$, and $\lambda_{\min}(\delta) = 1$, so $\lambda_{\min}(\zeta_{tA}) > 0$ for all $t \gg 0$. On the other hand, $\lambda_{\min}(\zeta_{tA})$ is continuous and nonzero for $t > \sup(\mathrm{Sing}(A))$. Hence $\lambda_{\min}(\zeta_{tA}) > 0$ for all $t > \sup(\mathrm{Sing}(A))$. $\square$

It follows that a space with Möbius inversion at all scales also satisfies an apparently stronger condition.

DEFINITION 2.4.7 A finite metric space A is *stably positive definite* if tA is positive definite for all $t > 0$.

COROLLARY 2.4.8 *Let A be a finite metric space. Then tA has Möbius inversion for all $t > 0$ if and only if A is stably positive definite.* $\square$

EXAMPLE 2.4.9 Let A be the space of Example 2.2.8. It is readily shown that tA has a unique weighting for all $t > 0$. By the remarks after Definition 1.1.3, tA has Möbius inversion for all $t > 0$, so A is stably positive definite. Hence magnitude is not continuous with respect to the Gromov–Hausdorff metric even when restricted to stably positive definite finite spaces. (Theorem 2.6 of [31] implies that it is, however, lower semicontinuous.)

Meckes [31, Theorem 3.3] has shown that a finite metric space is stably positive definite if and only if it is of negative type. By definition, a finite metric space A is of *negative type* if $\sum_{a,b} v(a)d(a,b)v(b) \leq 0$ for all $v \in \mathbb{R}^A$ such that $\sum_a v(a) = 0$. A general metric space A is of negative type if every finite subspace is of negative type, or equivalently if $(A, \sqrt{d_A})$ embeds isometrically into some Hilbert space [43]. Many of the most commonly encountered spaces are of negative type, including those that we prove below to be stably positive definite; see [31, Theorem 3.6] for a list. But whereas the classical results on negative type typically rely on embedding theorems, we are able to prove our results directly.

Lemma 2.2.5 gave additional hypotheses on finite metric spaces A and B guaranteeing that if A is an expansion of B then $|A| \geq |B|$. *Some* additional hypotheses are needed, since not every magnitude function is increasing (Example 2.2.7). The following will also do.

Lemma 2.4.10 *Let A and B be finite metric spaces. Suppose that A is positive definite and B admits a nonnegative weighting. If A is an expansion of B then $|A| \geq |B|$.*

Proof First consider a distance-decreasing *bi*jection $f \colon A \to B$. Choose a nonnegative weighting w_B on B. Without loss of generality, f is the identity as a map of sets; thus, $\zeta_A(a, a') \leq \zeta_B(a, a')$ for all points a, a'. Hence

$$|A| \geq \frac{(\sum w_B(a))^2}{w_B^* \zeta_A w_B} \geq \frac{(\sum w_B(a))^2}{w_B^* \zeta_B w_B} = |B|,$$

by Proposition 2.4.3.

Now consider the general case of a distance-decreasing surjection from A to B. We may choose a subspace $A' \subseteq A$ and a distance-decreasing bijection $A' \to B$. The space A' is positive definite, so $|A'| \geq |B|$ by the previous argument; but also $|A| \geq |A'|$ by Corollary 2.4.4. □

A positive definite space cannot have negative magnitude, but the following example shows that it can have magnitude greater than the number of points.

Example 2.4.11 Take the space $K_{3,2}$ of Example 2.2.7. It is easily shown that $\mathrm{Sing}(K_{3,2}) = \{\log \sqrt{2}\}$. Choose $u > \log \sqrt{2}$ such that $|uK_{3,2}| > 5$ (say, $u = 0.35$): then $A = uK_{3,2}$ is positive definite by Proposition 2.4.6, and $|A| > \#A$. This example also shows that a positive definite expansion of a positive definite space may have smaller magnitude: for if $s > 1$ then sA is an expansion of A, but $|sA| < |A|$ (Fig. 3).

A different positivity condition is sometimes useful: the existence of a nonnegative weighting.

Lemma 2.4.12 *Let A be a finite metric space admitting a nonnegative weighting. Then $0 \leq |A| \leq \#A$.*

Proof Choose a nonnegative weighting w on A. For all $a \in A$ we have $0 \leq w(a) \leq (\zeta_A w)(a) = 1$, so $0 \leq w(a) \leq 1$. Summing, $0 \leq |A| \leq \#A$. □

We now list some sufficient conditions for a space to be positive definite, or have a positive weighting, or both.

Proposition 2.4.13 *Every finite subspace of* $\mathbb{R}$ *is positive definite with positive weighting.*

Proof Let us temporarily say that a finite metric space A is *good* if it has Möbius inversion and for all $v \in \mathbb{R}^A$,

$$v^* \mu_A v \geq \max_{a \in A} v(a)^2.$$

I claim that if $A \cup B$ is a union of the type in Corollary 2.3.3 and A and B are both good, then $A \cup B$ is good. Indeed, let $v \in \mathbb{R}^{A \cup B}$. By Corollary 2.3.3,

$$v^* \mu_{A \cup B} v = v|_A^* \mu_A v|_A + v|_B^* \mu_B v|_B - v(c)^2$$

where $v|_A$ is the restriction of v to A. Now let $x \in A \cup B$. Without loss of generality, $x \in A$. Since A is good, $v|_A^* \mu_A v|_A \geq v(x)^2$. Since B is good, $v|_B^* \mu_B v|_B \geq v(c)^2$. Hence $v^* \mu_{A \cup B} v \geq v(x)^2$, proving the claim.
Every metric space with 0, 1 or 2 points is good. Every finite subset of $\mathbb{R}$ with 3 or more points can be expressed nontrivially as a union of the type in Corollary 2.3.3. It follows by induction that every finite subset of $\mathbb{R}$ is good and therefore positive definite.
Positivity of the weighting is immediate from Corollary 2.3.4. □

For $N \in \mathbb{N}$ and $1 \leq p \leq \infty$, write $\ell_p^N = \mathbb{R}^{\otimes_p N}$, where $\otimes_p$ is as defined in Example 1.4.2(ii). Thus, ℓ_p^N is $\mathbb{R}^N$ with the metric induced by the p-norm, $\|x\|_p = (\sum_r |x_r|^p)^{1/p}$.

Theorem 2.4.14 *Every finite subspace of* ℓ_1^N *is positive definite.*

Proof Let A be a finite subspace of ℓ_1^N. Write $\mathrm{pr}_1, \ldots, \mathrm{pr}_N \colon \ell_1^N \to \mathbb{R}$ for the projections. Each space $\mathrm{pr}_r A$ is positive definite by Proposition 2.4.13, so $\prod_{r=1}^N \mathrm{pr}_r A \subseteq \ell_1^N$ is positive definite by Lemma 2.4.2(ii), so A is positive definite by Lemma 2.4.2(iii). □

We prove the same result for Euclidean space in the next section.
In the category of metric spaces and distance-decreasing maps (Example 1.2.2(iii)), the categorical product $\times$ is $\otimes_\infty$. The class of positive definite spaces is not closed under $\times$. For if it were then, by an argument similar to the proof of Theorem 2.4.14, every finite subspace of ℓ_∞^N would be positive definite. But in fact, *every* finite metric space embeds isometrically into ℓ_∞^N for some N ([43], p.535), whereas not every finite metric space is positive definite. Comprehensive results on (non-)preservation of positive definiteness by the products $\otimes_p$ have been proved by Meckes [31, Section 3.2].

Proposition 2.4.15 *Every space with* 3 *or fewer points is positive definite with positive weighting.*

PROOF The proposition is trivial for spaces with 2 or fewer points. Now take a 3-point space $A = \{a_1, a_2, a_3\}$, writing $Z_{ij} = \zeta(a_i, a_j)$. We use Sylvester's criterion: a symmetric real $n \times n$ matrix is positive definite if and only if the upper-left $m \times m$ submatrix has positive determinant whenever $1 \leq m \leq n$. This holds for Z when $m = 1$ or $m = 2$, and

$$\begin{aligned}\det Z = (1 - Z_{12})(1 - Z_{23})(1 - Z_{31}) + (1 - Z_{12})(Z_{12} - Z_{13}Z_{32}) \\ + (1 - Z_{23})(Z_{23} - Z_{21}Z_{13}) + (1 - Z_{31})(Z_{31} - Z_{32}Z_{21})\end{aligned}$$

which is positive by the triangle inequality. The unique weighting is $v/\det Z$, where

$$v_1 = (1 - Z_{12})(1 - Z_{23})(1 - Z_{31}) + (1 - Z_{23})(Z_{23} - Z_{21}Z_{13}) > 0$$

and similarly v_2 and v_3. □

Meckes [31, Theorem 3.6] has shown that 4-point spaces are also positive definite. By Example 2.2.7, his result is optimal.

EXAMPLE 2.4.16 The weighting on a 4-point space may have negative components, as may the weighting on a finite subspace of ℓ_1^N. Indeed, using Proposition 2.3.2 one can show that in the space $\{(0,0), (t,0), (0,t), (-t,0)\} \subseteq \ell_1^2$, the weight at $(0,0)$ is negative whenever $t < \log 2$.

Every finite metric space, when scaled up sufficiently, becomes positive definite with positive weighting (Propositions 2.2.6 and 2.4.6). The following result provides an alternative, quantitative proof, using the notion of scattered space (Definition 2.1.2).

PROPOSITION 2.4.17 *Every scattered space is positive definite with positive weighting.*

PROOF Let A be a scattered space with $n \geq 2$ points. Positive definiteness follows from a version of the Levy–Desplanques theorem (Theorem 6.1.10 of [13]), but since the argument is simple, we repeat it here. Let $v \in \mathbb{R}^A$. Then

$$\begin{aligned}v^*\zeta_A v &= \sum_a v(a)^2 + \sum_{a \neq b} v(a)\zeta_A(a,b)v(b) \geq \sum_a v(a)^2 - \frac{1}{n-1}\sum_{a\neq b}|v(a)||v(b)| \\ &= \frac{1}{2(n-1)}\sum_{a\neq b}\big(|v(a)| - |v(b)|\big)^2 \geq 0.\end{aligned}$$

The inequality $\zeta_A(a,b) < 1/(n-1)$ $(a \neq b)$ is strict, so if $v^*\zeta_A v = 0$ then $v = 0$. To show that the unique weighting w_A on A is positive, we use the proof of Proposition 2.1.3. There we showed that A has Möbius inversion and that the Möbius matrix is a sum $\mu_A = \sum_{k=0}^{\infty}(-1)^k \mu_{A,k}$, where the matrices $\mu_{A,k}$ satisfy

$$\mu_{A,k+1}(a,b) < \frac{1}{n-1}\sum_{b'\,:\,b'\neq b}\mu_{A,k}(a,b') \tag{3}$$

for all a, b. Hence $w_A = \sum_{k=0}^{\infty}(-1)^k w_{A,k}$, where $w_{A,k}(a) = \sum_b \mu_{A,k}(a,b)$. Summing (3) over all $b \in A$ gives

$$w_{A,k+1}(a) < \frac{1}{n-1} \sum_{b,b' \colon b' \neq b} \mu_{A,k}(a,b') = w_{A,k}(a)$$

($a \in A$). Hence $w_A(a) = \sum_{k=0}^{\infty}(-1)^k w_{A,k}(a) > 0$ for all $a \in A$. □

A metric space A is *ultrametric* if $\max\{d(a,b), d(b,c)\} \geq d(a,c)$ for all $a, b, c \in A$.

Proposition 2.4.18 *Every finite ultrametric space is positive definite with positive weighting.*

Positive definiteness was proved by Varga and Nabben [49], and positivity of the weighting (rather indirectly) by Pavoine, Ollier and Pontier [38]. Another proof of positive definiteness is given by Meckes [31, Theorem 3.6]. Both parts of the following proof are different from those cited.

Proof Let Ω be the set of symmetric matrices Z over $[0,\infty)$ such that $Z_{ik} \geq \min\{Z_{ij}, Z_{jk}\}$ for all i, j, k and $Z_{ii} > \max_{j \neq k} Z_{jk}$ for all i. (For a 1×1 matrix, this maximum is to be interpreted as 0.) We show by induction that every matrix in Ω is positive definite and that its unique weighting (Definition 1.1.1) is positive. The proposition will follow immediately.

The result is trivial for 0×0 and 1×1 matrices. Now let $Z \in \Omega$ be an $n \times n$ matrix with $n \geq 2$. Put $z = \min_{i,j} Z_{ij}$. There is an equivalence relation $\sim$ on $\{1, \ldots, n\}$ defined by $i \sim j$ if and only if $Z_{ij} > z$.

It is not the case that $i \sim j$ for all i, j. Hence we may partition $\{1, \ldots, n\}$ into two nonempty subsets that are each a union of equivalence classes: say $\{1, \ldots, m\}$ and $\{m+1, \ldots n\}$. We have $Z_{ij} = z$ whenever $i \leq m < j$, so Z is a block sum

$$Z = \begin{pmatrix} Z' & zU_m^{n-m} \\ zU_{n-m}^m & Z'' \end{pmatrix}$$

where U_k^ℓ denotes the $k \times \ell$ matrix all of whose entries are 1. Since $Z' \in \Omega$ and $Z'_{ij} = Z_{ij} \geq z$ for all $i, j \leq m$, we have $Y' = Z' - zU_m^m \in \Omega$. Similarly, $Y'' = Z'' - zU_{n-m}^{n-m} \in \Omega$, and

$$Z = zU_n^n + \begin{pmatrix} Y' & 0 \\ 0 & Y'' \end{pmatrix}.$$

The first summand is positive semidefinite. By inductive hypothesis, Y' and Y'' are positive definite, so the second summand is positive definite. Hence Z is positive definite.

Also by inductive hypothesis, Y' and Y'' have positive weightings v' and v'' respectively. Let v be the concatenation of v' and v''. It is straightforward to verify that

$$\frac{v}{z(|Y'| + |Y''|) + 1}$$

is a weighting on Z, and it is positive since v' and v'' are positive and $z, |Y'|, |Y''| \geq 0$. □

Corollary 2.4.19 *If A is a finite ultrametric space then $|A| \leq e^{\operatorname{diam} A}$.*

Proof Let Δ be the metric space with the same point-set as A and $d(a, b) = \operatorname{diam} A$ for all distinct points a, b. By Proposition 2.1.5, $|\Delta| \leq e^{\operatorname{diam} A}$ and Δ has a positive weighting. But Δ is an expansion of A, so $|A| \leq |\Delta|$ by Lemma 2.2.5. □

A homogeneous space always has a positive weighting, by Proposition 2.1.5. However, Example 2.1.7 and Corollary 2.4.4 together show that a homogeneous space need not be positive definite. A homogeneous space need not even have Möbius inversion: $(\log 2)K_{3,3}$ is an example. In particular, a finite metric space may have magnitude but not Möbius inversion.

Magnitude can be understood in terms of entropy or diversity. For every finite metric space A and $q \in [0, \infty]$, there is a function ${}^qD^A$ assigning to each probability distribution $\mathbf{p}$ on A a real number ${}^qD^A(\mathbf{p})$, the *diversity of order q* of the distribution [26]. An ecological community can be modelled as a finite metric space A (as explained in Section 2.1) together with a probability distribution $\mathbf{p}$ on A (representing the relative abundances of the species). Then ${}^qD^A(\mathbf{p})$ is a measure of the biodiversity of the community. In the special case that A is discrete, the diversities are the exponentials of the Rényi entropies [39], and in particular, the diversity of order 1 is the exponential of the Shannon entropy. It is a theorem [23] that for each finite metric space A, there is some probability distribution $\mathbf{p}$ maximizing ${}^qD^A(\mathbf{p})$ for all $q \in [0, \infty]$ simultaneously. Moreover, the maximal value of ${}^qD^A(\mathbf{p})$ is independent of q; call it $D_{\max}(A)$. If A is positive definite with nonnegative weighting then, in fact, $|A| = D_{\max}(A)$: magnitude is maximum diversity.

2.5 Subsets of Euclidean space

Here we show that every finite subspace of Euclidean space ℓ_2^N is positive definite. In particular, every such space has well-defined magnitude.

Write $L_1(\mathbb{R}^N)$ for the space of Lebesgue-integrable complex-valued functions on $\mathbb{R}^N$. Define the Fourier transform $\hat{f}$ of $f \in L_1(\mathbb{R}^N)$ by

$$\hat{f}(\xi) = \int_{\mathbb{R}^N} e^{-2\pi i \langle \xi, x \rangle} f(x)\, dx$$

($\xi \in \mathbb{R}^N$). Define functions g and ψ on $\mathbb{R}^N$ by

$$g(x) = e^{-\|x\|_2}, \quad \psi(\xi) = C_N / (1 + 4\pi^2 \|\xi\|_2^2)^{(N+1)/2}$$

where C_N is the constant $2^N \pi^{(N-1)/2} \Gamma((N+1)/2) > 0$.

Lemma 2.5.1 *$\psi \in L_1(\mathbb{R}^N)$ and $\hat{\psi} = g$.*

Proof The first statement is straightforward. Theorem 1.14 of [47] states that $\hat{g} = \psi$; but g is continuous and even, so the second statement follows by Fourier inversion. □

The next lemma is elementary and standard (e.g. [15], Theorem VI.2.8).

Lemma 2.5.2 *Let $\phi \in L_1(\mathbb{R}^N)$, let A be a finite subset of $\mathbb{R}^N$, and let $v \in \mathbb{R}^A$. Then*

$$\sum_{a,b\in A} v(a)\hat{\phi}(a-b)v(b) = \int_{\mathbb{R}^N} \left| \sum_{a\in A} v(a) e^{-2\pi i\langle \xi, a\rangle} \right|^2 \phi(\xi)\, d\xi.$$

□

In analytic language, our task is to show that the function g is strictly positive definite. This would follow from the easy half of Bochner's theorem [15], except that Bochner's theorem concerns *non-strict* positive definiteness. We therefore need to refine the argument slightly.

Theorem 2.5.3 *Every finite subspace of Euclidean space is positive definite.*

Proof Let A be a finite subspace of ℓ_2^N. Let $v \in \mathbb{R}^A$. Then

$$v^* \zeta_A v = \sum_{a,b\in A} v(a) g(a-b) v(b) = \int_{\mathbb{R}^N} \left| \sum_{a\in A} v(a) e^{-2\pi i\langle \xi, a\rangle} \right|^2 \psi(\xi)\, d\xi \geq 0$$

by Lemmas 2.5.1 and 2.5.2. Suppose that $v \neq 0$. The characters $e^{-2\pi i\langle \cdot, a\rangle}$ ($a \in A$) are linearly independent, so the squared term is positive (that is, strictly positive) for some $\xi \in \mathbb{R}^N$. By continuity, the squared term is positive for all ξ in some nonempty open subset of $\mathbb{R}^N$. Moreover, ψ is continuous and everywhere positive. So the integral is positive, as required. □

On the other hand, some of the weights on a finite subspace of Euclidean space can be negative; see Willerton [50] for examples.

Corollary 2.5.4 *Every finite subspace of Euclidean space has magnitude.* □

A similar argument gives an alternative proof of Theorem 2.4.14, that finite subspaces of ℓ_1^N are positive definite. For this we use the explicit formula for the Fourier transform of $x \mapsto e^{-\|x\|_1}$. For $p \neq 1, 2$ there is no known formula for the Fourier transform of $e^{-\|x\|_p}$, so matters become more difficult. Nevertheless, Meckes [31, Section 3] has shown that every finite subspace of ℓ_p^N is positive definite whenever $0 < p \leq 2$, and that this is false for $p > 2$.

3 Compact metric spaces

To extend the notion of magnitude from finite to infinite spaces, there are broadly speaking two strategies.

In the first, we approximate an infinite space by finite spaces. As an initial attempt, given a compact metric space A, we might take a sequence (A_k) of finite metric spaces converging to A in the Gromov–Hausdorff metric, and try to define $|A|$ as the limit of the sequence $(|A_k|)$. However, this 'definition' is inconsistent; recall Example 2.2.8. We might respond by constraining the sequence (A_k): for example, by taking (A_k) to be a sequence of subsets of A converging to A in the Hausdorff metric.

The second strategy is to work directly with the infinite space, replacing finite sums by integrals. Weightings are now measures, or perhaps distributions. For example, a *weight measure* on a metric space A is a finite signed Borel measure w such that $\int_A e^{-d(a,b)} dw(b) = 1$ for all $a \in A$. If A admits a weight measure w then an argument similar to Lemma 1.1.2 shows that $w(A)$ is independent of the choice of w, and we may define the magnitude of A to be $w(A)$. This was the definition used by Willerton in [51].

Meckes [31] has shown that to a large extent, these different approaches produce the same result. Here we implement the first strategy, defining the magnitude of a space to be the supremum of the magnitudes of its finite subspaces. This works well when the space is compact and its finite subspaces are positive definite.

3.1 The magnitude of a positive definite compact metric space

Definition 3.1.1 A metric space is *positive definite* if every finite subspace is positive definite. The *magnitude* of a compact positive definite space A is

$$|A| = \sup\{|B| : B \text{ is a finite subspace of } A\} \in [0, \infty].$$

These definitions are consistent with the definitions for finite metric spaces, by Lemma 2.4.2(iii) and Corollary 2.4.4.

There may even be non-compact spaces for which this definition of magnitude is sensible. For example, let $t > 0$, and let A be a space with infinitely many points and $d(a, b) = t$ for all $a \neq b$; then every finite subspace of A is positive definite, and the supremum of their magnitudes is $e^t < \infty$. In any case, we confine ourselves to compact spaces.

A metric space A is *stably positive definite* if tA is positive definite for all $t > 0$, or equivalently if every finite subspace of A is stably positive definite. (A further equivalent condition, due to Meckes, is that A is of negative type [31, Theorem 3.3].) We already know that ℓ_1^N and ℓ_2^N are stably positive definite; much of the rest of this paper concerns the magnitudes of their compact subspaces. Ultrametric spaces are also stably positive definite (Proposition 2.4.18), and, if compact, have finite magnitude (Corollary 2.4.19). Many other commonly occurring spaces, such as hyperbolic space, are stably positive definite too; see [31, Theorem 3.6].

Definition 3.1.2 Let A be a stably positive definite compact metric space. The *magnitude function* of A is the function

$$\begin{array}{ccc}(0,\infty) & \to & [0,\infty] \\ t & \mapsto & |tA|.\end{array}$$

Lemma 3.1.3 *Let A be a positive definite compact metric space. Then:*

i. *Every closed subspace B of A is positive definite, and $|B| \leq |A|$.*

ii. *If A is nonempty then $|A| \geq 1$.* □

Proposition 3.1.4 *Let A and B be positive definite compact spaces. Then $A \otimes B$ is positive definite and compact, and $|A \otimes B| = |A||B|$.*

In the case $A = \emptyset$ and $|B| = \infty$, we interpret $0 \cdot \infty$ as 0.

Proof Let C be a finite subspace of $A \otimes B$. Then $C \subseteq A' \otimes B'$ for some finite subspaces $A' \subseteq A$ and $B' \subseteq B$. Since A and B are positive definite, so are A' and B'. By Lemma 2.4.2, $A' \otimes B'$ is positive definite, so C is positive definite. Hence $A \otimes B$ is positive definite. A similar argument shows that $|A \otimes B| = |A||B|$, using Proposition 2.3.6 and Corollary 2.4.4. □

Similarly, Proposition 2.3.2 on unions extends to the compact setting.

Proposition 3.1.5 *Let X be a metric space and $A, B \subseteq X$, with A and B compact and $A \cup B$ positive definite. Suppose that A projects to B and B projects to A. Then*

$$|A \cup B| + |A \cap B| = |A| + |B|.$$

Proof Let $\varepsilon > 0$. Choose finite sets $E \subseteq A \cup B$ and $H \subseteq A \cap B$ such that $|A \cup B| \leq |E| + \varepsilon$ and $|A \cap B| \leq |H| + \varepsilon$. For each $a \in E \cap A$, choose $\pi_A(a) \in A \cap B$ satisfying the condition of Definition 2.3.1, and similarly $\pi_B(b)$ for $b \in E \cap B$. Put

$$H' = H \cup \pi_A(E \cap A) \cup \pi_B(E \cap B), \quad F = (E \cap A) \cup H', \quad G = (E \cap B) \cup H'.$$

Then F and G are finite subsets of X, each projecting to the other. Also $E \subseteq F \cup G$ and $H \subseteq F \cap G$. Applying Proposition 2.3.2 to F and G gives $|A \cup B| + |A \cap B| \leq |A| + |B| + 2\varepsilon$. Since ε was arbitrary, $|A \cup B| + |A \cap B| \leq |A| + |B|$.

For the opposite inequality, again let $\varepsilon > 0$, and choose finite sets $F \subseteq A$ and $G \subseteq B$ such that $|A| \leq |F| + \varepsilon$ and $|B| \leq |G| + \varepsilon$. For each $a \in F$, choose

$\pi_A(a) \in A \cap B$ satisfying the condition of Definition 2.3.1, and similarly $\pi_B(b)$ for $b \in G$. Put

$$F' = F \cup \pi_A F \cup \pi_B G, \qquad G' = G \cup \pi_A F \cup \pi_B G.$$

Then F' and G' are finite subsets of X, each projecting to the other; also $F \subseteq F' \subseteq A$ and $G \subseteq G' \subseteq B$. A similar argument proves that $|A| + |B| \leq |A \cup B| + |A \cap B| + 2\varepsilon$. □

3.2 Subsets of the real line

As soon as we ask about the magnitude of real intervals, connections with geometric measure begin to appear.

Proposition 3.2.1 *Let $t \geq 0$ and let (A_k) be a sequence of finite subsets of $\mathbb{R}$ converging to $[0,t]$ in the Hausdorff metric. Then $(|A_k|)$ converges to $1 + t/2$.*

This result was announced in [22], and also appears, with a different proof, as Proposition 6 of [27].

Proof Given $A = \{a_0 < \cdots < a_n\} \subseteq \mathbb{R}$, we have

$$(1 + t/2) - |A| = \sum_{i=1}^{n} \Big\{ \frac{a_i - a_{i-1}}{2} - \tanh\Big(\frac{a_i - a_{i-1}}{2}\Big) \Big\} + \frac{t - (a_n - a_0)}{2}$$

by Corollary 2.3.4. The result will follow from the facts that $\tanh(0) = 0$ and $\tanh'(0) = 1$. Indeed, write $f(x) = (x - \tanh(x))/x$, so that $f(x) \to 0$ as $x \to 0$. Then

$$\big|(1 + t/2) - |A|\big| \leq \Big(\frac{a_n - a_0}{2}\Big) \max_{1 \leq i \leq n} \Big| f\Big(\frac{a_i - a_{i-1}}{2}\Big)\Big| + \Big|\frac{t - (a_n - a_0)}{2}\Big|.$$

But $\max_i (a_i - a_{i-1}) \to 0$ and $a_n - a_0 \to t$ as $A \to [0,t]$, proving the proposition. □

Theorem 3.2.2 *The magnitude of a closed interval $[0,t]$ is $1 + t/2$.*

Proof Proposition 3.2.1 immediately implies that $|[0,t]| \geq 1 + t/2$. Now let A be a finite subset of $[0,t]$. We may choose a sequence (A_k) of finite subsets of $\mathbb{R}$ such that $\lim_{k\to\infty} A_k = [0,t]$ and $A \subseteq A_k$ for all k. Then $|A| \leq |A_k| \to 1 + t/2$ as $k \to \infty$, so $|A| \leq 1 + t/2$. □

Schanuel [41] argued from basic geometric intuition that the 'size' of a closed interval of length t inches ought to be (t inches $+ 1$). Ignoring the factor of $1/2$ (which is purely a product of convention), Theorem 3.2.2 makes his idea rigorous.

As noted by Willerton [51], there is a weight measure on $[0,t]$. It is $w = (\delta_0 + \lambda + \delta_t)/2$, where δ_x is the Dirac measure at x and λ is Lebesgue measure on $[0,t]$. Then $w([0,t]) = 1 + t/2$. Hence $w([0,t]) = |[0,t]|$, as guaranteed by Theorems 2.3 and 2.4 of Meckes [31].

The magnitude of subsets of $\mathbb{R}$ is also described by the following formula, which has no known analogue in higher dimensions.

Proposition 3.2.3 *Let A be a compact subspace of $\mathbb{R}$. Then*

$$|A| = \frac{1}{2}\int_{\mathbb{R}} \operatorname{sech}^2 d(x,A)\,dx$$

where $d(x,A) = \inf_{a\in A} d(x,a)$.

Proof First we prove the identity for finite spaces $A \subseteq \mathbb{R}$, by induction on $n = \#A$. It is elementary when $n \leq 2$. Now suppose that $n \geq 3$, writing the points of A as $a_1 < \cdots < a_n$. Put $B = \{a_1, \ldots, a_{n-1}\}$ and $C = \{a_{n-1}, a_n\}$. Then

$$\frac{1}{2}\int_{\mathbb{R}} \operatorname{sech}^2 d(x,A)\,dx = \frac{1}{2}\int_{-\infty}^{a_{n-1}} \operatorname{sech}^2 d(x,B)\,dx + \frac{1}{2}\int_{a_{n-1}}^{\infty} \operatorname{sech}^2 d(x,C)\,dx.$$

Since $\int_0^\infty \operatorname{sech}^2 u\,du = 1$, this in turn is equal to

$$\frac{1}{2}\Bigl(\int_{\mathbb{R}} \operatorname{sech}^2 d(x,B)\,dx - 1\Bigr) + \frac{1}{2}\Bigl(\int_{\mathbb{R}} \operatorname{sech}^2 d(x,C)\,dx - 1\Bigr)$$

which by inductive hypothesis is $|B| + |C| - 1$. On the other hand, $|A| = |B| + |C| - 1$ by Corollary 2.3.3. This completes the induction.

Now take a compact space $A \subseteq \mathbb{R}$. We know that

$$|A| = \sup\Bigl\{\frac{1}{2}\int_{\mathbb{R}} \operatorname{sech}^2 d(x,B)\,dx : B \text{ is a finite subset of } A\Bigr\}.$$

Since sech^2 is decreasing on $[0,\infty)$, this implies that

$$|A| \leq \frac{1}{2}\int_{\mathbb{R}} \operatorname{sech}^2 d(x,A)\,dx.$$

To prove the opposite inequality, choose a sequence (B_k) of finite subsets of A converging to A in the Hausdorff metric. We have $0 \leq \operatorname{sech}^2 d(x,B_k) \leq \operatorname{sech}^2 d(x,A)$ for all x and k, so

$$\lim_{k\to\infty}\int_{\mathbb{R}} \operatorname{sech}^2 d(x,B_k)\,dx = \int_{\mathbb{R}} \operatorname{sech}^2 d(x,A)\,dx$$

by the dominated convergence theorem. The result follows. □

3.3 Background on integral geometry

To go further, we need some concepts and results from integral geometry. Those concerning ℓ_2^N can be found in standard texts such as [17]. Those concerning ℓ_1^N can be found in [24].

Write $\mathscr{K}_N$ for the set of compact convex subsets of $\mathbb{R}^N$. A *valuation* on $\mathscr{K}_N$ is a function $\phi\colon \mathscr{K}_N \to \mathbb{R}$ such that

$$\phi(\emptyset) = 0, \qquad \phi(A \cup B) = \phi(A) + \phi(B) - \phi(A \cap B) \tag{4}$$

whenever $A, B, A \cup B \in \mathscr{K}_N$. It is *continuous* if continuous with respect to the Hausdorff metric on $\mathscr{K}_N$, and *invariant* if $\phi(gA) = \phi(A)$ for all $A \in \mathscr{K}_N$ and isometries $g\colon \ell_2^N \to \ell_2^N$ (not necessarily fixing the origin).

Examples 3.3.1 i. N-dimensional Lebesgue measure is a continuous invariant valuation on $\mathscr{K}_N$, denoted by Vol.

ii. Euler characteristic χ is a continuous invariant valuation on $\mathscr{K}_N$. Since the sets are convex, $\chi(A)$ is 0 or 1 according as A is empty or not.

The continuous invariant valuations on $\mathscr{K}_N$ form a real vector space, Val_N.

When $A \subseteq \ell_p^N$ and $t > 0$, the abstract metric space tA can be interpreted as the subspace $\{ta : a \in A\}$ of ℓ_p^N. A valuation ϕ is *homogeneous* of degree i if $\phi(tA) = t^i \phi(A)$ for all $A \in \mathscr{K}_N$ and $t > 0$.

Theorem 3.3.2 (Hadwiger [12]) *The vector space* Val_N *has dimension* $N+1$ *and a basis* $V_0, \ldots, V_N$ *where* V_i *is homogeneous of degree* i. □

This description determines the valuations V_i uniquely up to scale factor. They can be uniquely normalized to satisfy two conditions. First, $V_N(A) = \mathrm{Vol}(A)$ for $A \in \mathscr{K}_N$. Second, whenever ℓ_2^N is embedded isometrically into ℓ_2^{N+1} and $0 \leq i \leq N$, the value $V_i(A)$ is the same whether A is regarded as a subset of ℓ_2^N or of ℓ_2^{N+1}. With this normalization, V_i is called the ith *intrinsic volume*.

For example, $V_0 = \chi$. When $A \in \mathscr{K}_2$, $V_1(A)$ is half of the perimeter of A; when $A \in \mathscr{K}_3$, $V_2(A)$ is half of the surface area.

Here is a general formula for the intrinsic volumes. For each $0 \leq i \leq N$, there is an $\mathrm{O}(N)$-invariant measure $\nu_{N,i}$ on the Grassmannian $\mathrm{Gr}_{N,i}$, unique up to scale factor. Given $P \in \mathrm{Gr}_{N,i}$, write $\pi_P\colon \mathbb{R}^N \to P$ for orthogonal projection. Then for $A \in \mathscr{K}_N$,

$$V_i(A) = c_{N,i} \int_{\mathrm{Gr}_{N,i}} \mathrm{Vol}(\pi_P A)\, d\nu_{N,i}(P)$$

where $c_{N,i}$ is a positive constant chosen so that the normalizing conditions are satisfied.

Hadwiger's theorem solves the classification problem for valuations on ℓ_2^N. More generally, we can try to classify the valuations on any metric space, in the following sense.

A metric space A is *geodesic* [37] if for all $a, b \in A$ there exists an isometry $\gamma\colon [0, d(a,b)] \to A$ with $\gamma(0) = a$ and $\gamma(d(a,b)) = b$. Given a metric space X, write $\mathscr{K}(X)$ for the set of compact subsets of X that are geodesic with respect to the subspace metric. For example, $\mathscr{K}(\ell_2^N) = \mathscr{K}_N$.

A *valuation* on $\mathscr{K}(X)$ is a function $\phi\colon \mathscr{K}(X) \to \mathbb{R}$ satisfying equations (4) whenever $A, B, A \cup B, A \cap B \in \mathscr{K}(X)$. It is *continuous* if continuous with respect to the Hausdorff metric, and *invariant* if $\phi(gA) = \phi(A)$ for all isometries g of X. Write $\mathrm{Val}(X)$ for the vector space of continuous invariant valuations on $\mathscr{K}(X)$. For example, $\mathrm{Val}(\ell_2^N) = \mathrm{Val}_N$.

Given any metric space X, one can attempt to describe the vector space $\mathrm{Val}(X)$. Here we will need to know the answer for ℓ_1^N, as well as ℓ_2^N. To state it, we write $\mathscr{K}'_N = \mathscr{K}(\ell_1^N)$ and call its elements compact ℓ_1-*convex* sets; similarly, we write $\mathrm{Val}'_N = \mathrm{Val}(\ell_1^N)$.

There are far more ℓ_1-convex sets than convex sets. On the other hand, there are far fewer isometries of ℓ_1^N than of ℓ_2^N; they are generated by translations, coordinate permutations, and reflections in coordinate hyperplanes. The following Hadwiger-type theorem is proved in [24].

Theorem 3.3.3 *The vector space* Val'_N *has dimension* $N + 1$ *and a basis* $V'_0, \ldots, V'_N$ *where* V'_i *is homogeneous of degree* i. □

Again, this determines the valuations V'_i uniquely up to scaling. They can be described as follows. For $0 \leq i \leq N$, let $\mathrm{Gr}'_{N,i}$ be the set of i-dimensional vector subspaces of $\mathbb{R}^N$ spanned by some subset of the standard basis. For $A \in \mathscr{K}'_N$, put

$$V'_i(A) = \sum_{P \in \mathrm{Gr}'_{N,i}} \mathrm{Vol}(\pi_P A).$$

These valuations $V'_0, \ldots, V'_N$ are called the ℓ_1-*intrinsic volumes*, and satisfy two normalization conditions analogous to those in the Euclidean case.

The intrinsic volumes of a product space are given by the following formula, proved in [24, Proposition 8.1] and precisely analogous to the classical Euclidean formula [17, Theorem 9.7.1].

Proposition 3.3.4 *Let* $A \in \mathscr{K}'_M$ *and* $B \in \mathscr{K}'_N$. *Then* $A \times B \in \mathscr{K}'_{M+N}$, *and*

$$V'_k(A \times B) = \sum_{i+j=k} V'_i(A) V'_j(B)$$

whenever $0 \leq k \leq M + N$. □

3.4 Subsets of ℓ_1^N

Our investigation of the magnitude of subsets of ℓ_1^N begins with sets of a particularly amenable type.

DEFINITION 3.4.1 A *cuboid* in ℓ_1^N is a subspace of the form $[x_1, y_1] \times \cdots \times [x_N, y_N]$, where $x_r, y_r \in \mathbb{R}$ with $x_r \leq y_r$.

As an abstract metric space, a cuboid is a tensor product $[x_1, y_1] \otimes \cdots \otimes [x_N, y_N]$.

THEOREM 3.4.2 *For cuboids $A \subseteq \ell_1^N$,*

$$|A| = \sum_{i=0}^{N} 2^{-i} V_i'(A). \tag{5}$$

PROOF First let $I = [x, y] \subseteq \mathbb{R}$ be a nonempty interval. By Theorem 3.2.2,

$$|I| = 1 + (y - x)/2 = \chi(I) + \mathrm{Vol}(I)/2 = V_0'(I) + 2^{-1} V_1'(I).$$

This proves the theorem for $N = 1$. The theorem also holds for $N = 0$.
It now suffices to show that if $A \in \mathscr{K}_M'$ and $B \in \mathscr{K}_N'$ satisfy (5) then so does $A \times B \in \mathscr{K}_{M+N}'$. Indeed, as a metric space, $A \times B \subseteq \ell_1^{M+N}$ is $A \otimes B$, and the result follows from Propositions 3.1.4 and 3.3.4. $\square$

In fact, $V_i'(\prod [x_r, y_r])$ is the ith elementary symmetric polynomial in $(y_r - x_r)_{r=1}^N$, again by Proposition 3.3.4. It is also equal to $V_i(\prod [x_r, y_r])$, the Euclidean intrinsic volume. But in general, the Euclidean and ℓ_1-intrinsic volumes of a convex set are not equal.

COROLLARY 3.4.3 *The magnitude function of a cuboid $A \subseteq \ell_1^N$ is given by*

$$|tA| = \sum_{i=0}^{N} 2^{-i} V_i'(A) t^i.$$

In particular, the magnitude function of a cuboid A is a polynomial whose degree is the dimension of A, and whose coefficients are proportional to the ℓ_1-intrinsic volumes of A. $\square$

The moral is that for spaces belonging to this small class, the dimension and all of the ℓ_1-intrinsic volumes can be recovered from the magnitude function. In this sense, magnitude encodes those invariants. For the rest of this work we advance the conjectural principle—first set out in [27]—that the same is true for a much larger class of spaces, in both ℓ_1^N and ℓ_2^N.
We begin by showing that the principle holds for subspaces of ℓ_1^N when the invariant concerned is dimension.

DEFINITION 3.4.4 The *growth* of a function $f\colon (0, \infty) \to \mathbb{R}$ is

$$\inf\{\nu \in \mathbb{R} : f(t)/t^\nu \text{ is bounded for } t \gg 0\} \in [-\infty, \infty].$$

For example, the growth of a polynomial is its degree.

DEFINITION 3.4.5 The *(magnitude) dimension* $\dim A$ of a stably positive definite compact metric space A is the growth of its magnitude function.

EXAMPLES 3.4.6 i. The magnitude dimension of a cuboid in ℓ_1^N is its dimension in the usual sense, by Corollary 3.4.3.

ii. The magnitude dimension of a nonempty finite space is 0, by Proposition 2.2.6(v).

LEMMA 3.4.7 *Let A be a stably positive definite compact space. Then:*

i. Every closed subspace $B \subseteq A$ satisfies $\dim B \leq \dim A$.

ii. If $A \neq \emptyset$ then $\dim A \geq 0$.

PROOF For (i), we have $0 \leq |tB| \leq |tA|$ for all $t > 0$, so $\dim B \leq \dim A$. For (ii), take B to be a one-point subspace of A. $\square$

Recall that the magnitude of a positive definite compact space can in principle be infinite (although there are no known examples).

THEOREM 3.4.8 *Let A be a compact subset of ℓ_1^N. Then:*

i. $|A| < \infty$.

ii. $\dim A \leq N$, with equality if A has nonempty interior.

We will show in Theorem 3.5.8 that the hypothesis 'nonempty interior' can be relaxed to 'nonzero measure'.

PROOF A is a subset of some cuboid $B \subseteq \ell_1^N$, which has finite magnitude by Theorem 3.4.2, so $|A| \leq |B| < \infty$. Also $\dim A \leq \dim B \leq N$ by Lemma 3.4.7 and Example 3.4.6(i). If A has nonempty interior then it contains an N-dimensional cuboid, giving $\dim A \geq N$. $\square$

We now ask whether the ℓ_1-intrinsic volumes of an ℓ_1-convex set can be extracted from its magnitude function.

Let $\mathscr{C}_N$ be the smallest class of compact subsets of ℓ_1^N containing all cuboids and closed under unions of the type in Proposition 3.1.5. By that proposition and Theorem 3.4.2, equation (5) holds for all $A \in \mathscr{C}_N$.

EXAMPLE 3.4.9 Let T be a solid, compact triangle in ℓ_1^2 with two edges parallel to the coordinate axes. We compute $|T|$ by exhaustion. For each $k \geq 1$, let I_k be the union of k rectangles approximating T from the inside as in Fig. 5; similarly, let E_k be the exterior approximation by k rectangles. Then T, I_k

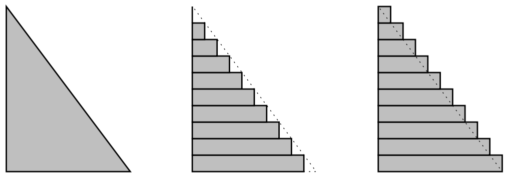

Figure 5: Triangle $T \subseteq \ell_1^2$, with interior and exterior approximations I_{10}, E_{10}

and E_k are all ℓ_1-convex with $I_k, E_k \in \mathscr{C}_2$, and $\lim_{k\to\infty} I_k = T = \lim_{k\to\infty} E_k$, so

$$\begin{aligned}\lim_{k\to\infty} |I_k| = \lim_{k\to\infty} \sum_{i=0}^{2} 2^{-i} V_i'(I_k) &= \sum_{i=0}^{2} 2^{-i} V_i'(T) \\ = \lim_{k\to\infty} \sum_{i=0}^{2} 2^{-i} V_i'(E_k) &= \lim_{k\to\infty} |E_k|\end{aligned}$$

(using continuity of V_i' in the second and third equalities). But $|I_k| \leq |T| \leq |E_k|$ for all k, so $|T| = \sum_{i=0}^{2} 2^{-i} V_i'(T)$. Similar arguments prove this identity for all compact convex polygons in ℓ_1^2.

These and other examples suggest the following conjecture.

Conjecture 3.4.10 *Let A be a compact ℓ_1-convex subspace of ℓ_1^N. Then*

$$|A| = \sum_{i=0}^{N} 2^{-i} V_i'(A).$$

If the conjecture holds then $|tA| = \sum_{i=0}^{N} 2^{-i} V_i'(A) t^i$ for all $t > 0$ and $A \in \mathscr{K}_N'$. Hence we can recover all of the ℓ_1-intrinsic volumes of an ℓ_1-convex set from its magnitude function.

3.5 Subsets of Euclidean space

We now prove results for ℓ_2^N similar to some of those for ℓ_1^N. Our first task is to prove that the magnitude of a compact subset of Euclidean space is finite. Given $A \subseteq \mathbb{R}^N$, write

$$\begin{aligned}\mathscr{S}(A) = \{&\text{Schwartz functions } \phi\colon \mathbb{R}^N \to \mathbb{R} \text{ such that } \hat{\phi}(a-b) = 1 \\ &\text{for all } a, b \in A\}.\end{aligned}$$

Lemma 3.5.1 *Let $A \subseteq \mathbb{R}^N$ be a bounded set. Then $\mathscr{S}(A) \neq \emptyset$.*

PROOF Since A is bounded, there is a real even Schwartz function f such that $f(a-b)=1$ for all $a,b\in A$; then there is a unique real Schwartz function ϕ such that $\hat{\phi}=f$. $\square$

The rest of the proof uses the function ψ from Section 2.5. For a Schwartz function ϕ on $\mathbb{R}^N$, write

$$c(\phi)=\sup_{\xi\in\mathbb{R}^N}|\phi(\xi)/\psi(\xi)|<\infty.$$

LEMMA 3.5.2 *Let A be a compact subspace of ℓ_2^N and $\phi\in\mathscr{S}(A)$. Then $|A|\leq c(\phi)$.*

PROOF Let B be a finite subset of A; thus, B is positive definite by Theorem 2.5.3. Then for all $v\in\mathbb{R}^B$, using Lemma 2.5.2 twice,

$$\begin{aligned}
c(\phi)\cdot v^*\zeta_B v &= c(\phi)\int_{\mathbb{R}^N}\left|\sum_{a\in B}v(a)e^{-2\pi i\langle\xi,a\rangle}\right|^2\psi(\xi)\,d\xi\\
&\geq \int_{\mathbb{R}^N}\left|\sum_{a\in B}v(a)e^{-2\pi i\langle\xi,a\rangle}\right|^2\phi(\xi)\,d\xi\\
&= \sum_{a,b\in B}v(a)\hat{\phi}(a-b)v(b)=\Bigl(\sum_{a\in B}v(a)\Bigr)^2.
\end{aligned}$$

Taking v to be the unique weighting on B gives $c(\phi)\geq|B|$. $\square$

PROPOSITION 3.5.3 *The magnitude of a compact subspace of ℓ_2^N is finite.* $\square$

We can extract more from the argument. For a compact set $A\subseteq\mathbb{R}^N$, write

$$\langle A\rangle=\inf\{c(\phi):\phi\in\mathscr{S}(A)\}<\infty.$$

Lemma 3.5.2 states that $|A|\leq\langle A\rangle$.

LEMMA 3.5.4 *Let A be a compact subset of $\mathbb{R}^N$ and $t\geq 1$. Then $\langle tA\rangle\leq t^N\langle A\rangle$.*

PROOF Let $\phi\in\mathscr{S}(A)$. Define $\theta\colon\mathbb{R}^N\to\mathbb{R}$ by $\theta(\xi)=t^N\phi(t\xi)$. Then θ is Schwartz, and if $a,b\in tA$ then $\hat{\theta}(a-b)=\hat{\phi}((a-b)/t)=1$. Hence $\theta\in\mathscr{S}(tA)$. I now claim that $c(\theta)\leq t^N c(\phi)$. Indeed, using the fact that $\psi(\xi)\geq\psi(t\xi)$ for all $\xi\in\mathbb{R}^N$,

$$c(\theta)=t^N\sup_{\xi\in\mathbb{R}^N}\left|\frac{\phi(t\xi)}{\psi(\xi)}\right|\leq t^N\sup_{\xi\in\mathbb{R}^N}\left|\frac{\phi(t\xi)}{\psi(t\xi)}\right|=t^N c(\phi).$$

This proves the claim, and the result follows. $\square$

THEOREM 3.5.5 *A compact subspace of ℓ_2^N has dimension at most N.*

PROOF For compact $A \subseteq \ell_2^N$ and $t \geq 1$, we have $|tA| \leq \langle tA \rangle \leq t^N \langle A \rangle$; hence $\dim A \leq N$. $\square$

The same proof can be adapted to ℓ_1^N, although we already have a much more elementary proof (Theorem 3.4.8).
Having bounded magnitude from above, we now bound it from below.

THEOREM 3.5.6 *Let $\|\cdot\|$ be a norm on $\mathbb{R}^N$ whose induced metric is positive definite. Write $B = \{x \in \mathbb{R}^N : \|x\| \leq 1\}$. For a compact set $A \subseteq \mathbb{R}^N$, equipped with the subspace metric,*

$$|A| \geq \frac{\mathrm{Vol}(A)}{N!\,\mathrm{Vol}(B)}.$$

Before proving this, we state some consequences. Write ω_N for the volume of the unit Euclidean N-ball.

COROLLARY 3.5.7 *Let A be a compact subset of $\mathbb{R}^N$.*

i. If A is given the subspace metric from ℓ_2^N then $|A| \geq \mathrm{Vol}(A)/N!\omega_N$.

ii. If A is given the subspace metric from ℓ_1^N then $|A| \geq \mathrm{Vol}(A)/2^N$.

PROOF Part (i) is immediate. Part (ii) follows from the fact that the unit ball in ℓ_1^N has volume $2^N/N!$, or can be derived from Lemma 3.5.9 below. $\square$

THEOREM 3.5.8 *Let $p \in \{1, 2\}$ and let A be a compact subspace of ℓ_p^N. Then $\dim A \leq N$, with equality if A has nonzero Lebesgue measure.*

PROOF The inequality follows from Theorems 3.4.8 and 3.5.5. Now suppose that $\mathrm{Vol}(A) > 0$. By Corollary 3.5.7, there is a constant $K_A > 0$ such that $|tA| \geq K_A t^N$ for all $t > 0$, giving $\dim A \geq N$. $\square$

Generalizations of these theorems have been proved by Meckes, using more sophisticated methods ([31], Theorems 4.4 and 4.5). In particular, Theorem 3.5.8 is extended to ℓ_p^N for all $p \in [1, 2]$.
To prove Theorem 3.5.6, we first need a standard calculation.

LEMMA 3.5.9 *Let $\|\cdot\|$ be a norm on $\mathbb{R}^N$. Write B for the unit ball. Then*

$$\int_{\mathbb{R}^N} e^{-\|x\|}\,dx = N!\,\mathrm{Vol}(B).$$

PROOF $\displaystyle\int_{\mathbb{R}^N} e^{-\|x\|}\,dx = \int_{r=0}^{\infty} e^{-r}\,d(\mathrm{Vol}(rB)) = \int_0^{\infty} e^{-r} N r^{N-1}\,\mathrm{Vol}(B)\,dr = N!\,\mathrm{Vol}(B)$. $\square$

Proof of Theorem 3.5.6 We use the result of Meckes [31, Theorem 2.4] that for a compact positive definite space A and a finite Borel measure v on A,

$$|A| \geq v(A)^2 \Big/ \int_A \int_A e^{-d(a,b)}\, dv(a)\, dv(b).$$

Let $A \subseteq \mathbb{R}^N$ be a compact set and take v to be Lebesgue measure: then

$$\begin{aligned} |A| &\geq \frac{\mathrm{Vol}(A)^2}{\int_A \int_A e^{-\|a-b\|}\, da\, db} \geq \frac{\mathrm{Vol}(A)^2}{\int_A \int_{\mathbb{R}^N} e^{-\|a-b\|}\, da\, db} \\ &= \frac{\mathrm{Vol}(A)^2}{\int_A \int_{\mathbb{R}^N} e^{-\|x\|}\, dx\, db} = \frac{\mathrm{Vol}(A)}{\int_{\mathbb{R}^N} e^{-\|x\|}\, dx}. \end{aligned}$$

The theorem follows from Lemma 3.5.9. □

This proof is a rigorous rendition of part of Willerton's bulk approximation argument [50]. There is an alternative proof in the same spirit, not depending on the results of Meckes but instead working with finite approximations. We sketch it now.

Alternative proof of Theorem 3.5.6 For $\delta > 0$, write

$$S_\delta = \Big\{ x \in \delta\mathbb{Z}^N : A \cap \prod_{r=1}^{N} [x_r, x_r + \delta) \neq \emptyset \Big\}.$$

Define $\alpha\colon \delta\mathbb{Z}^N \to \mathbb{R}^N$ by choosing for each $x \in S_\delta$ an element $\alpha(x) \in A \cap \prod [x_r, x_r + \delta)$, and putting $\alpha(x) = x$ for $x \in \delta\mathbb{Z}^n \setminus S_\delta$.
A calculation similar to that in the first proof of Theorem 3.5.6 shows that for all $\delta > 0$,

$$|A| \geq \frac{\# S_\delta}{\sum_{x \in \delta\mathbb{Z}^N} E_\delta(x)}$$

where

$$E_\delta(x) = \frac{1}{\# S_\delta} \sum_{y \in S_\delta} e^{-\|\alpha(x+y) - \alpha(y)\|} \quad \big(\approx e^{-\|x\|}\big).$$

(Apply Proposition 2.4.3 to the finite space αS_δ.) Since Lebesgue measure is outer regular, $\lim_{\delta\to 0}(\delta^N(\# S_\delta)) = \mathrm{Vol}(A)$. From the fact that $\|\alpha(x) - x\| \leq \mathrm{diam}([0,\delta)^N)$ for all $\delta > 0$ and $x \in \delta\mathbb{Z}^N$, it also follows that

$$\lim_{\delta\to 0} \Big(\delta^N \sum_{x \in \delta\mathbb{Z}^N} E_\delta(x) \Big) = \int_{\mathbb{R}^N} e^{-\|x\|}\, dx.$$

The theorem now follows from Lemma 3.5.9. □

These results suggest the following conjecture, first stated in [27]:

Conjecture 3.5.10 *Let A be a compact convex subspace of ℓ_2^N. Then*

$$|A| = \sum_{i=0}^{N} \frac{1}{i!\omega_i} V_i(A).$$

Assuming the conjecture, the magnitude function of a compact convex set $A \subseteq \ell_2^N$ is a polynomial:

$$|tA| = \sum_{i=0}^{N} \frac{1}{i!\omega_i} V_i(A) t^i. \tag{6}$$

All of the intrinsic volumes, as well as the dimension, can therefore be recovered from the magnitude function.
The evidence for Conjecture 3.5.10 is as follows.

- The two sides of equation (6) have the same growth (by Theorem 3.5.8).
- The left-hand side of (6) is greater than or equal to the leading term of the right-hand side (by Corollary 3.5.7).
- Both sides of (6) are monotone increasing in A (by Lemma 3.1.3).
- The conjecture holds for $N = 1$ (by Theorem 3.2.2).
- It is closely analogous to Conjecture 3.4.10, which, while itself a conjecture, is known to hold for a nontrivial class of examples. (To see the analogy, note that in both cases the ith coefficient is $1/i! \operatorname{Vol}(B_i)$, where B_i is the i-dimensional unit ball.)
- There is good numerical evidence, due to Willerton [50], when A is a disk, square or cube.

One strategy for proving Conjecture 3.5.10 would be to apply Hadwiger's theorem (3.3.2). There are currently two obstacles. First, it is not known that magnitude is a valuation on compact convex sets. Certainly it is not a valuation on *all* compact subsets of ℓ_2^N: consider the union of two points.
Second, even if we knew that magnitude was a valuation on convex sets, the conjecture would still not be proved. We would know that magnitude was an invariant valuation, monotone and therefore continuous by Theorem 8 of McMullen [30]. By Hadwiger's theorem, there would be constants c_i such that $|A| = \sum c_i V_i(A)$ for all convex sets A. However, current techniques provide no way of computing those constants. Knowing the magnitude of balls or cubes would be enough; but apart from subsets of the line, there is not a single convex subset of Euclidean space whose magnitude is known.

References

[1] J. Baez and J. Dolan. From finite sets to Feynman diagrams. In *Mathematics Unlimited—2001 and Beyond*. Springer, Berlin, 2001.

[2] H.-J. Bandelt and A. W. M. Dress. A canonical decomposition theory for metrics on a finite set. *Advances in Mathematics*, 92:47–105, 1992.

[3] C. Berger and T. Leinster. The Euler characteristic of a category as the sum of a divergent series. *Homology, Homotopy and Applications*, 10:41–51, 2008.

[4] L. M. Blumenthal. *Theory and Applications of Distance Geometry*. Chelsea, New York, 2nd edition, 1970.

[5] P. Di Francesco, P. Mathieu, and D. Sénéchal. *Conformal Field Theory*. Graduate Texts in Contemporary Physics. Springer, Berlin, 1997.

[6] A. Dress, K. T. Huber, and V. Moulton. Metric spaces in pure and applied mathematics. *Documenta Mathematica*, Quadratic Forms LSU:121–139, 2001.

[7] T. M. Fiore, W. Lück, and R. Sauer. Euler characteristics of categories and homotopy colimits. *Documenta Mathematica*, 16:301–354, 2011.

[8] T. M. Fiore, W. Lück, and R. Sauer. Finiteness obstructions and Euler characteristics of categories. *Advances in Mathematics*, 226:2371–2469, 2011.

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

[10] R. L. Graham and P. M. Winkler. On isometric embeddings of graphs. *Transactions of the American Mathematical Society*, 288:527–536, 1985.

[11] M. Gromov. *Metric Structures for Riemannian and Non-Riemannian Spaces*. Birkhäuser, Boston, 2001.

[12] H. Hadwiger. *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*. Springer, Berlin, 1957.

[13] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[14] M. W. Jacobsen and J. Møller. Euler characteristics and Möbius algebras of p-subgroup categories. *Journal of Pure and Applied Algebra*, 216:2665–2696, 2012.

[15] Y. Katznelson. *An Introduction to Harmonic Analysis*. Cambridge University Press, Cambridge, 3rd edition, 2004.

[16] 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.

[17] D. A. Klain and G.-C. Rota. *Introduction to Geometric Probability*. Lezioni Lincee. Cambridge University Press, Cambridge, 1997.

[18] M. Kormes. On the functional equation $f(x+y) = f(x) + f(y)$. *Bulletin of the American Mathematical Society*, 32:689–693, 1926.

[19] 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.

[20] 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.

[21] T. Leinster. The Euler characteristic of a category. *Documenta Mathematica*, 13:21–49, 2008.

[22] T. Leinster. Metric spaces. Post at *The n-Category Café*, http://golem.ph.utexas.edu/category/2008/02/metric_spaces.html, 2008.

[23] T. Leinster. A maximum entropy theorem with applications to the measurement of biodiversity. arXiv:0910.0906, 2009.

[24] T. Leinster. Integral geometry for the 1-norm. *Advances in Applied Mathematics*, 49:81–96, 2012.

[25] T. Leinster. Notions of Möbius inversion. *Bulletin of the Belgian Mathematical Society*, 19:911–935, 2012.

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

[27] T. Leinster and S. Willerton. On the asymptotic magnitude of subsets of Euclidean space. *Geometriae Dedicata*, 164:287–310, 2013.

[28] S. Mac Lane. *Categories for the Working Mathematician*. Graduate Texts in Mathematics 5. Springer, Berlin, 1971.

[29] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. North-Holland, Amsterdam, 1977.

[30] P. McMullen. Valuations and Euler-type relations on certain classes of convex polytopes. *Proceedings of the London Mathematical Society*, 35:113–135, 1977.

[31] M. W. Meckes. Positive definite metric spaces. *Positivity*, in press, 2013.

[32] K. Menger. Die Metrik des Hilbertschen Raumes. *Anzeiger der Akademie der Wissenschaften in Wien, mathematisch-naturwissenschaftliche Klasse*, 65:159–160, 1928.

[33] K. Noguchi. The Euler characteristic of acyclic categories. *Kyushu Journal of Mathematics*, 65:85–99, 2011.

[34] K. Noguchi. The zeta function of a finite category and the series Euler characteristic. arXiv:1207.6750, 2012.

[35] K. Noguchi. The zeta function of a finite category which has Möbius inversion. arXiv:1205.4380, 2012.

[36] K. Noguchi. Euler characteristics of categories and barycentric subdivision. *Münster Journal of Mathematics*, in press, 2013.

[37] A. Papadopoulos. *Metric Spaces, Convexity, and Nonpositive Curvature*. European Mathematical Society, Zürich, 2005.

[38] S. Pavoine, S. Ollier, and D. Pontier. Measuring diversity from dissimilarities with Rao's quadratic entropy: Are any dissimilarities suitable? *Theoretical Population Biology*, 67:231–239, 2005.

[39] A. Rényi. On measures of entropy and information. In J. Neymann, editor, *4th Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 547–561. University of California Press, 1961.

[40] G.-C. Rota. On the foundations of combinatorial theory I: theory of Möbius functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 2:340–368, 1964.

[41] 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.

[42] 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.

[43] I. J. Schoenberg. Metric spaces and positive definite functions. *Transactions of the American Mathematical Society*, 44:522–536, 1938.

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

[45] D. Speyer. Re: Metric spaces. Comment at [22], 2008.

[46] R. P. Stanley. *Enumerative Combinatorics Volume 1*. Cambridge Studies in Advanced Mathematics 49. Cambridge University Press, Cambridge, 1997.

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

[48] T. Tao. Re: Metric spaces. Comments at [22], 2008.

[49] R. S. Varga and R. Nabben. On symmetric ultrametric matrices. In L. Reichel, A. Ruttan, and R. S. Varga, editors, *Numerical Linear Algebra*, pages 193–199. Walter de Gruyter, 1993.

[50] S. Willerton. Heuristic and computer calculations for the magnitude of metric spaces. arXiv:0910.5500, 2009.

[51] S. Willerton. On the magnitude of spheres, surfaces and other homogeneous spaces. *Geometriae Dedicata*, in press, 2013.

Tom Leinster
School of Mathematics
University of Edinburgh
The King's Buildings
Mayfield Road
Edinburgh EH9 3JZ
United Kingdom