This is a repository copy of *Heuristic and computer calculations for the magnitude of metric spaces*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/127459/

Version: Published Version

Article:

Willerton, S. (Submitted: 2009) Heuristic and computer calculations for the magnitude of metric spaces. arXiv. (Submitted)

HEURISTIC AND COMPUTER CALCULATIONS
FOR THE MAGNITUDE OF METRIC SPACES

SIMON WILLERTON

ABSTRACT. The notion of the magnitude of a compact metric space was considered in a previous paper with Tom Leinster, where the magnitude was calculated for line segments, circles and Cantor sets. In this paper more evidence is presented for a conjectured relationship with a geometric measure theoretic valuation. Firstly, a heuristic is given for deriving this valuation by considering 'large' subspaces of Euclidean space and, secondly, numerical approximations to the magnitude are calculated for squares, disks, cubes, annuli, tori and Sierpinski gaskets. The valuation is seen to be very close to the magnitude for the convex spaces considered and is seen to be 'asymptotically' close for some other spaces.

INTRODUCTION

This paper is concerned with an empirical look at the magnitude of subspaces of Euclidean space. Magnitude was introduced under the name 'cardinality' by Leinster in [4] as a partially defined invariant of finite metric spaces; his motivation came from enriched category theory, but the invariant had already been considered in the biological diversity literature [6]. In [5] we started to extend the definition to compact, non-finite metric spaces by approximating these with finite metric spaces; we considered, in particular, the magnitude of certain subsets of Euclidean space, such as line segments and circles and we showed that in these cases, when the metric space is scaled up in size, the magnitude asymptotically satisfies the inclusion-exclusion principle. This leads to some natural conjectures and the current paper provides numerical and heuristic evidence for these conjectures. The current paper, however, can be read independently of [5].

As hinted by the title, there are two main ingredients to this paper. The first ingredient is a heuristic argument to calculate the contribution to the magnitude from the 'bulk' of points in the closure of a 'large' open subset X of Euclidean space $\mathbb{R}^n$; this contribution is shown to be roughly $\mathrm{vol}(X)/n!\,\omega_n$ where ω_n is the volume of the unit n-ball. From here one is led to consider, for naturality reasons, the so-called penguin valuation, defined on certain compact subsets of $\mathbb{R}^n$ by

$$P(X) := \sum_{i=0}^{n} \frac{\mu_i(X)}{i!\,\omega_i}$$

where $\mu_i(X)$ is the ith intrinsic volume of X (see [3, 2]). The top intrinsic volume $\mu_n(X)$ is just the usual n-dimensional volume of X, whereas $\mu_{n-1}(X)$ is half the 'surface area', meaning half the $n-1$-dimensional volume of the boundary of X, and $\mu_0(X)$ is the Euler characteristic of X. These intrinsic volumes have the scaling property that for $t > 0$ if tX denotes X with all the distances scaled up by t then $\mu_i(tX) = t^i\mu_i(X)$. Furthermore they are normalized on unit cubes, so that $\mu_i([0,1]^i) = 1$. I am not sure the exact class of subsets of Euclidean space on which the intrinsic volumes are defined, but it includes polyconvex sets, meaning finite unions of convex sets, [3] and smooth submanifolds [2].

The second ingredient consists of computer calculations of finite spaces approximating squares, cubes, discs, annuli, tori and Sierpinski gaskets. The computer algebra `maple` was used to calculate the magnitude of finite spaces with around 25,000 points.

The two ingredients are combined when the calculated magnitudes for the first five types of spaces mentioned above are plotted together with the penguin valuation. In the convex examples — squares, cubes and discs — there is a surprisingly good fit, leading to the conjecture that for K a convex subset of Euclidean space the magnitude and penguin valuation are equal: $|K| = P(K)$. In the case that K is a straight line segment this is proved rigourously in [5]. In the non-convex examples there is a good fit for larger spaces, leading to the conjecture that in those cases the magnitude is asymptotically the same as the penguin functional in the sense that

$$|tX| - P(tX) \to 0 \qquad \text{as } t \to \infty,$$

although it is not clear what spaces X this would be expected to hold for in general, other than for the closure of open sets. For instance, for the 'bent line', in which two straight line segments are stuck together at an angle, numerical calculations of the magnitude by David Speyer seem to show a small, constant discrepancy with the penguin valuation; also, asymptotic calculations seem to show that the lower order terms are not right in the case of the 3-sphere, although I do not have much confidence in these calculations at the moment.

In the fractal case it is not possible to make sense of the penguin functional. However, in [5] we were able to calculate precisely the magnitude of the Cantor set and observed that the asymptotic growth rate was the Hausdorff dimension of the Cantor set. In Section 4 we see that the empirical data for the Sierpinski gasket is consistent with its magnitude of the growing like its side length to the power of its Hausdorff dimension.

The rest of this introduction consists of a reminder of the definition of magnitude and an informal analogy which might provide some intuition, followed by a more detailed synopsis of the paper.

Definition of magnitude and a useful analogy. We first define the magnitude on finite metric spaces, then after some comments on the definition we will say how we wish to extend it to non-finite metric spaces.

If X is a finite metric space then a weighting on X is an assignment of a real number $w_x \in \mathbb{R}$ to each point $x \in X$ such that for each point $x' \in X$ the *weight equation* for x' is satisfied:

$$\sum_{x \in X} e^{-d(x,x')} w_x = 1.$$

If a weighting exists then the *magnitude* $|X|$ is defined to be the sum of the weights: $|X| := \sum_{x \in X} w_x$.

Two standard observations to make here are, firstly, that the weights are not necessarily positive, as will be seen in examples below, and, secondly, that if there is more than one weighting then the magnitude is independent of the weighting.

One informal analogy that might be useful to keep in mind is the following. Imagine that the points are certain organisms that wish to maintain a certain body temperature which we normalize to be a temperature of 1 unit. To achieve this body temperature each organism can generate heat or cold and experiences the heat or cold from other organisms in a way that falls off exponentially with the distance. If the amount of heat that organism x is generating is given by w_x then the fact that an organism x' maintains unit body temperature is expressed in the weight equation for x'.

Weightings of some finite subsets of $\mathbb{R}^2$ (with the subspace metric) are given in Figures 1 and 2: the area of each disk is proportional to the weight of the point at its centre and the colour of the disk represents the sign of the weight — red is positive and blue is negative. Note how the weights are greater around the edge of the metric space. In line with the intuition described above, one is led to think about emperor penguins which huddle together in large groups on the Antarctic ice during the winter in order to conserve heat (see [1]); the penguins take turns to be on the outside of the group where it is coldest and they need to use up more energy. Despite being cute and whimsical, this imagery does seem to give helpful intuition.

We are interested here in extending the notion of magnitude to compact subsets of Euclidean space. If X is such a compact subset then we approximate X by a sequence $\{X_k\}_{k=0}^{\infty}$ of finite subsets of X converging (in the Hausdorff topology) to X. If $\lim_{k \to \infty} |X_k|$ exists then we define $|X|$ the magnitude of X to be this limit. In general we do not know that this is a well defined process — maybe the limit does not exist, or the limit depends on the choice of approximating sequence, or the approximating subsets do not themselves have a magnitude — but it gives sensible answers in the examples that we can calculate. In this paper typically we start with a compact subset X of $\mathbb{R}^n$ and intersect it with some lattice $\mathscr{L}$ to give a finite set $\ddot{X} := X \cap \mathscr{L}$. (This is pronounced "$X$ dots" and is supposed to indicate the fact that it is a discrete set.) We assume that for a sufficiently fine lattice this gives rise to a good approximation for the magnitude we are interested in: $|\ddot{X}| \simeq |X|$.

What is in this paper. In Section 1 we have a look at some features of weighting numerically calculated on grid approximations to squares. In Section 2 part of these features are fed into a back-of-the-envelope calculation for the contribution to the magnitude from the bulk of point in a large subset of $\mathbb{R}^n$; this calculation then leads to the definition of the penguin valuation. Section 3 contains the main results of the paper with the penguin valuation being graphed together with numerically calculated approximations to the magnitude for some subsets of Euclidean space. In the convex examples — squares, discs and cubes — the penguin valuation and the magnitude are seen to be very close; in the other examples — annuli and tori — they are seen to be very close when the spaces are large. In Section 4 the Sierpinski gasket is considered and it is seen that the emprirical data is compatible with the growth rate of the magnitude being the Hausdorff dimension. The appendix contains some comments on the computer calculations.

Acknowledgements. I would like to thank Tom Leinster for many stimulating conversations and the information about emperor penguins, David Speyer for useful comments on an earlier draft, Sam Marsh for encouraging me to calculate some examples and also Emma McCabe and Richard Thomas for helping me develop the heat analogy.

1. Numerically approximating the magnitude of a square

In this section we have our first look at the results of calculations of the weights of points on a square grid and make some observations about these which will feed into the Bulk Approximation in the next section.

Having looked at line segments, circles and Cantor sets in [5], the next most obvious class of spaces to look at was squares. Unlike the simpler examples, we are not yet able to find an exact formula for weighting in this case, but we can look at the results of computer calculations. For $t > 0$, let tQ be a $t \times t$ square,

thought of as a subset of $\mathbb{R}^2$, this will be approximated by $t\ddot{Q}$ which is a square grid of $m \times m$ points with the distance between adjacent points being $(m-1)/t$. Remember that the magnitude $|t\ddot{Q}|$ is supposed to give an approximation to the magnitude $|tQ|$ of the continuous square—however, it isn't yet proved that the continuous square has a well-defined magnitude, although the data in this paper provides evidence for the existence of such a magnitude.

For specific values of t and m, provided that m is not too large, we can use a computer algebra package such as `maple` to numerically calculate the weights on the finite metric space $t\ddot{Q}$, and hence calculate the magnitude $|t\ddot{Q}|$, which is just the sum of the weights; details of the computer implementation are given in the appendix.

The results of some computer calculations of the weights of $t\ddot{Q}$ for some values of side length t and m points per side are represented graphically in Figures 1 and 2. In these pictures the weight at each point is represented by a disc centred at that point with the area of the disc equal to four times the weight at the point; the colour of the disc represents the sign of the weight, with red meaning a positive (or 'hot') weight and blue meaning a negative (or 'cold') weight.
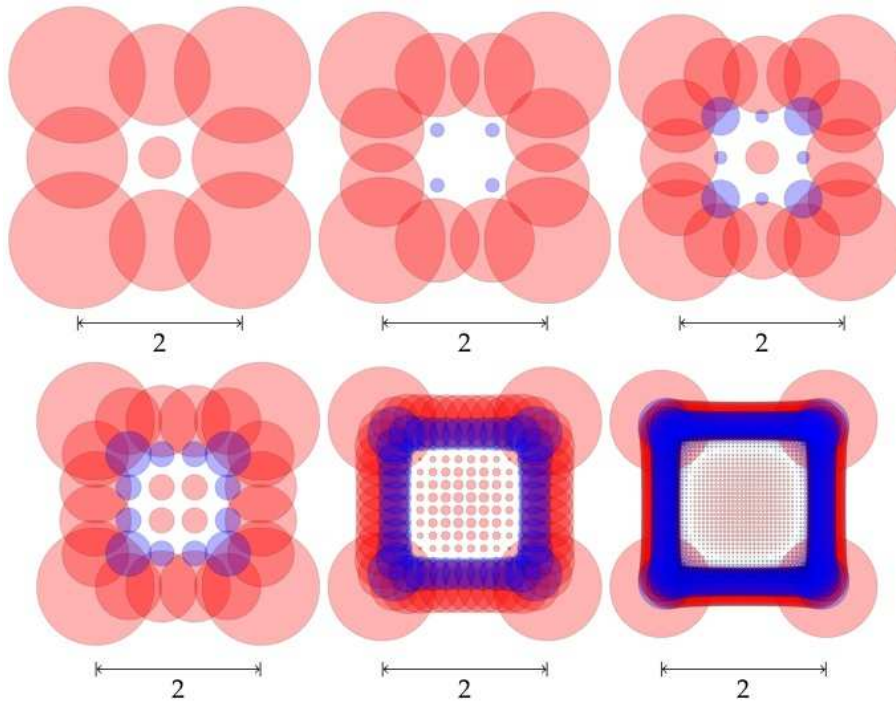

FIGURE 1. Weightings on square grids of side length 2 with respectively 3, 4, 5, 6, 14 and 40 points per side.

This graphical representation of the weights is not perfect: for instance where the discs overlap near the edges it is not entirely clear what is happening (see Figure 3 for a better picture of the behaviour near the edge). However, these pictures make the following features apparent.

(1) In the 'bulk' of the square the weights are reasonably constant.
(2) On the boundary the weights are bigger, but again are reasonably constant.
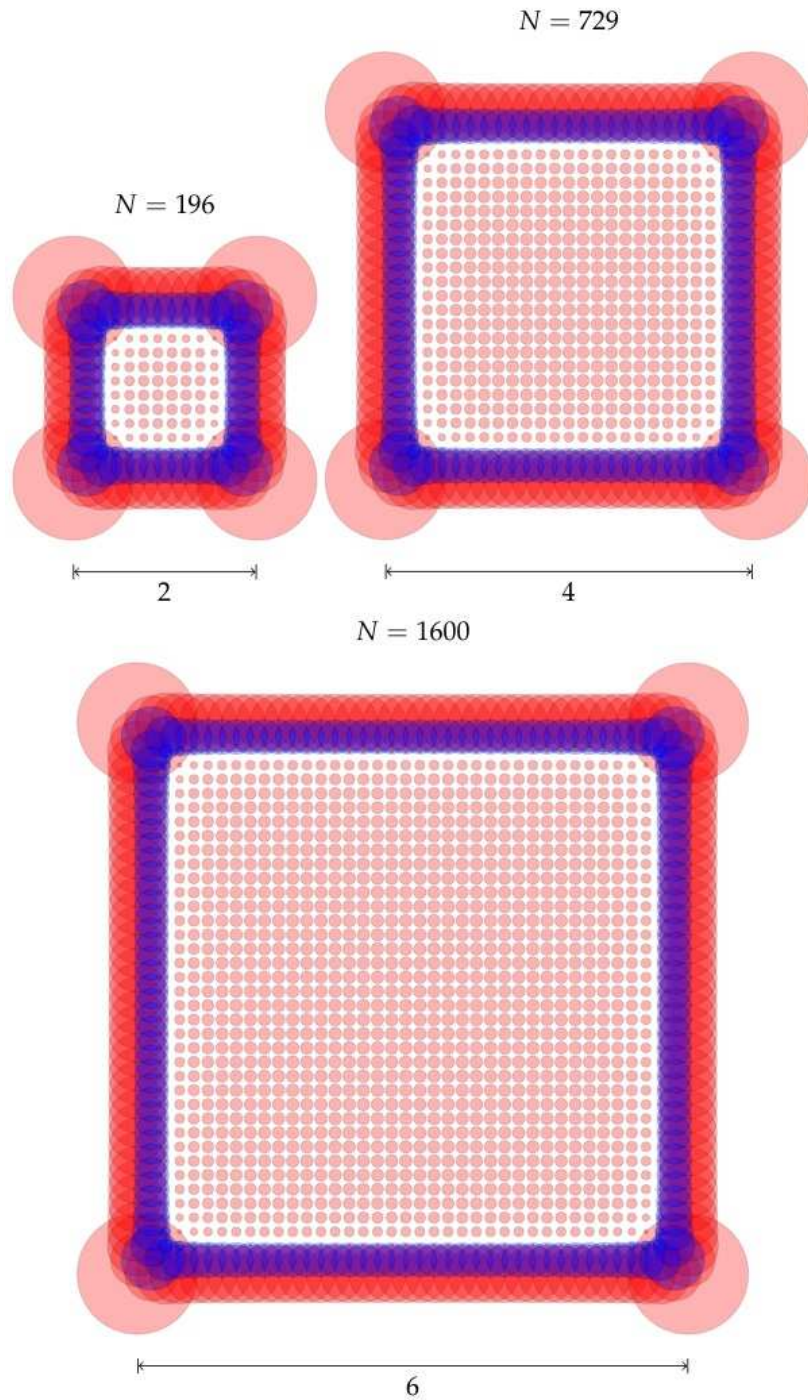(3) Near to the boundary the weights are negative.

FIGURE 2. Weightings on the square grid of side length 2, 4 and 6. The grid spacing is the same, approximately 0.15, in each case. The area of a disc is four times the weight of the point at its centre; red denotes a positive weight and blue a negative weight.

In the next section we give a heuristic argument for the first feature and this argument implies that the constant should be $1/2\pi$ times the area of the little squares in the grid. A vague heuristic for the second and third features was given in the introduction — you should think of penguins huddling together to keep warm, those on the outside of the huddle need to generate more heat to keep warm whilst those next to them are overheated and so need to cool themselves down.

We finish this section by looking a little closer at the behaviour near the edge. We consider a square with $(2s+1) \times (2s+1)$ points then we look at the weights near the edge along the middle row — that is the sth row. If a point x has weight w_x then, based on the calculations of the next section, we define the *bulk-normalized weight* of a point x in the square to be $2\pi w_x / \operatorname{vol}(V)$ where $\operatorname{vol}(V)$ is the volume of a little square in the grid — that is the volume of a Voronoi cell in the language of the next section. According to my rather limited calculations, if the side length t of the square is sufficiently large, say at least 10, then the bulk-normalized weight of a point on the middle row near the edge depends only on the distance from the edge and not on the side length or the spacing between points, with the singular exception of the weight of the point *nearest* to the edge which is usually more negative than might be expected. A sample profile of bulk-normalized weights near the edge is given in Figure 3.
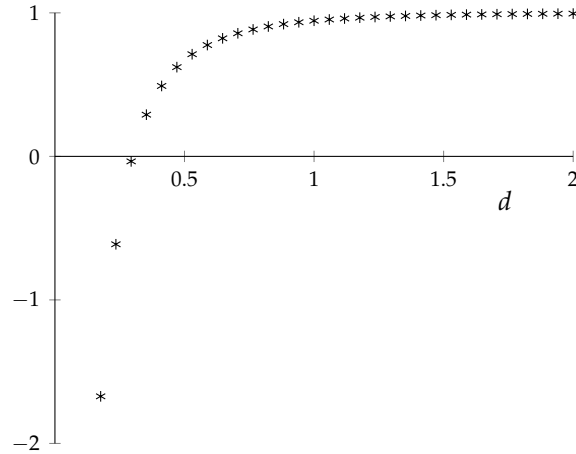

FIGURE 3. The bulk-normalized weights of the middle row of points at a distance d from the edge on a 10×10 square with 171×171 grid of points.

It seems from Figure 3 that the weight of a point at a distance d from the edge is of the form

$$(1 + h(d)) \frac{\operatorname{vol}(V)}{2\pi}$$

where $h \colon (0, \infty) \to (-\infty, 0)$ is a function such that

$$h(d) \to -\infty \quad \text{as } d \to 0 \qquad \text{and} \qquad h(d) \to 0 \quad \text{exponentially as } d \to \infty.$$

This behaviour is reminiscent of acutance in photographic images and the Gibbs phenomenon in signal processing. However, I have not been able to figure out any heuristic giving a quantitative explanation of this.

2. The bulk approximation and the penguin valuation

In this section we give a heuristic approximation to the contribution to the magnitude by the 'bulk' of points in a 'large' compact subset of $\mathbb{R}^n$. Continuing in the heuristic vein and assuming some connection between magnitude and invariant valuations we are led to the definition of the penguin valuation P at the end of this section. In the next section we will see numerical evidence that the penguin valuation is often close to the magnitude.

In the previous section it was observed that in the 'bulk' of a sufficiently large square all of the weights appeared to be constant. Here we use that observation as the seed for a heuristic argument describing, for X the closure of an open subset of $\mathbb{R}^n$, the contribution of the bulk of X to the magnitude of X. By bulk(X), the *bulk* of X, we will mean all of the points 'reasonably far from the boundary'. Remember that the influence of the weight of a point decreases exponentially with distance, so you might take 'reasonably far' to mean 'at least ten units away': this is just a heuristic argument, after all. The following is not a precise statement, but will be supported by numerical evidence in the next section.

Bulk Approximation. *Let X be the closure of a large open set in $\mathbb{R}^n$, let $\ddot{X}$ be the intersection of X with a sufficiently fine lattice and let ω_n be the volume of the unit n-ball in $\mathbb{R}^n$. Then the contribution to the magnitude of $\ddot{X}$ by the bulk of points in $\ddot{X}$ is roughly* $\mathrm{vol}(X)/n!\,\omega_n$.

Heuristic justification. The idea of the heuristic is that because the space is large and the effect of a point on other points decays exponentially with distance, the points in the bulk of $\ddot{X}$ only really know what is happening locally and so, as far as the weight equation is concerned, do not know that they are not part of an infinite lattice rather than a finite subset of one. We can solve the weight equation for the infinite lattice and get a simple approximation to the resulting weight. It is this approximation that we use to approximate the weight of points in the bulk of $\ddot{X}$.

So denote the lattice by $\mathscr{L}$. We can solve the weight equations for $\mathscr{L}$, essentially using Speyer's method for homogeneous metric spaces (see [5, Theorem 1]). We will assume without loss of generality that the origin 0 is in the lattice, so we can consider the weight equation for 0:

$$\sum_{x \in \mathscr{L}} e^{-d(x,0)} w_x = 1.$$

We now assume, motivated by the observation in Section 1 about the points in the bulk of the square, that all the points in $\mathscr{L}$ have the same weight w. Writing, as usual, $|x| = d(0, x)$ we get

$$w = \frac{1}{\sum_{x \in \mathscr{L}} e^{-|x|}}.$$

It is not immediate that the sum in the denominator converges, but we can approximate it by Riemann sums in the following way.

First pick a fundamental domain V for $\mathscr{L}$; for concreteness we can take this to be the Voronoi cell for the origin, which means all of the points in $\mathbb{R}^n$ which are at least as close to the origin as to any other lattice point. The translates by $\mathscr{L}$ of the fundamental domain V give a tessellation of $\mathbb{R}^n$, and each of the cells of the tessellation contains one point of the lattice. The volume of the fundamental domain will be denoted by $\mathrm{vol}(V)$ and is sometimes called the *covolume* of the lattice.

We can use this tessellation of the plane to calculate a Riemann sum approximation of an integral. For the function $e^{-|x|}$, with the lattice sufficiently fine, we

find

$$\sum_{x \in \mathscr{L}} e^{-|x|} \operatorname{vol}(V) \simeq \int_{x \in \mathbb{R}^n} e^{-|x|} \, dx.$$

The integral can be calculated by a change of variable to 'polar coordinates':

$$\int_{x \in \mathbb{R}^n} e^{-|x|} \, dx = \int_{\theta \in S^{n-1}} \int_{r=0}^{\infty} e^{-r} r^{n-1} \, dr \, d\theta = (n-1)! \operatorname{vol}(S^{n-1}),$$

where the equality $\int_{r=0}^{\infty} e^{-r} r^{n-1} \, dr = (n-1)!$ can be shown using integration by parts and induction — it is just the evaluation of a Γ function. As the the volume of the $(n-1)$-sphere and the n-ball are related by $\operatorname{vol}(S^{n-1}) = n \operatorname{vol}(B^n) =: n\omega_n$ we see

$$\sum_{x \in \mathscr{L}} e^{-|x|} \operatorname{vol}(V) \simeq n! \, \omega_n$$

and thus the weight of every point in the lattice $\mathscr{L}$ is given by

$$w = \frac{\operatorname{vol}(V)}{\sum_{x \in \mathscr{L}} e^{-|x|} \operatorname{vol}(V)} \simeq \frac{\operatorname{vol}(V)}{n! \, \omega_n}.$$

Now return to the case of the finite metric space $\ddot{X}$ which is the intersection of the large space X with a lattice $\mathscr{L}$. As X was taken to be the closure of an open set in $\mathbb{R}^n$, if we consider a point x in the bulk, then as far x is concerned it might as well be in the lattice $\mathscr{L}$ rather than $X \cap \mathscr{L}$, because the weight equation says that its interactions with points at a distance r is e^{-r}, so it has exponentially small interaction with far-away points. So assuming local homogeneity, we take the weight of x to be the same as if it were considered as a point in $\mathscr{L}$ rather than $\ddot{X}$. So the points in the bulk of $\ddot{X}$ all have roughly the same weight which is

$$w = \frac{\operatorname{vol}(V)}{n! \, \omega_n}.$$

The contribution of the points in the bulk to the magnitude is given by summing these weights, so we obtain roughly

$$\frac{\operatorname{vol}(\operatorname{bulk} X)}{n! \, \omega_n} \simeq \frac{\operatorname{vol}(X)}{n! \, \omega_n},$$

as required.   □

There are lots of approximations in the above argument, so it is not clear that it gives a sensible answer. However, the answer is supported by the empirical data given in the next section; moreover, if we continue in a heuristic manner, as follows, we obtain an answer which is an even better fit with the empirical data.

The above heuristic indicates that for X the closure of a large open subset of $\mathbb{R}^n$ then

$$|X| \simeq \frac{\operatorname{vol}(X)}{n! \, \omega_n} + \text{contributions from near the boundary.}$$

Based on the rather meagre examples from [5], we now assume that for large X the magnitude $|X|$ is close to some valuation $P(X)$. Recall from [3] and [5] that the intrinsic volumes $\{\mu_i\}$ form a basis for the space of valuations. As the μ_n is the n-dimensional volume of a subset of $\mathbb{R}^n$ we know that in dimension n we must have $P(X) = \frac{\mu_n(X)}{n! \, \omega_n} + \text{lower order terms}$. Now if $\iota \colon \mathbb{R}^i \hookrightarrow \mathbb{R}^n$ is an inclusion of Euclidean spaces and $Y \subset \mathbb{R}^i$ is some polyconvex subset then by the naturality

of the intrinsic volumes with respect to inclusion of Euclidean spaces

$$P(\iota Y) \simeq |\iota Y| = |Y| \simeq \frac{\mu_i(Y)}{i!\,\omega_i} + \text{lower order terms}$$

$$= \frac{\mu_i(\iota Y)}{i!\,\omega_i} + \text{lower order terms}.$$

So the coefficient of μ_i in P is of the same form for all i. We therefore define, for want of a better name, the *penguin valuation P* by

$$P(X) := \sum_i \frac{\mu_i(X)}{i!\,\omega_i}$$

and speculate whether or not $|X| \simeq P(X)$ for large spaces X.

Again, this looks like rather wild speculation; however, when, as in the next section, for certain spaces X we plot $P(tX)$ against the magnitude of approximations to tX an impressively close fit is seen across a wide range of values of t. This is especially true for the convex sets considered, namely the square, the disc and the cube, where $P(tX)$ is seen to be close to the approximation to $|tX|$ for all t, not just for large t.

From this, one is lead to ask if there is an interesting class of spaces X for which the magnitude is asymptotically just the valuation P, i.e., $|tX| - P(tX) \to 0$ as $t \to \infty$, and is it true that for X convex the magnitude is just the valuation, i.e., $|X| = P(X)$.

3. Numerically comparing the magnitude with the penguin valuation

In this section we will plot the magnitude of various subsets of Euclidean spaces — squares, discs, cubes, annuli and tori — together with the penguin valuation of these subsets. We can exactly calculate the penguin valuation for these and for the magnitude we numerically evaluate the magnitude of finite approximations to these subsets with around $25,000$ points. We see a good correlation between the magnitude and the penguin valuation. However, as pointed out in the introduction, this correlation seems unlikely to hold for every subset of Euclidean space.

In each of the following cases we take an interesting subset X of Euclidean space and approximate it by $\ddot{X}$ a finite subset of X consisting of N points. We pick N as large as possible subject to constraints of available computing power; these constraints are discussed in the appendix. We then use `maple` to calculate the magnitude $|t\ddot{X}|$ for various scale factors t; again, see the appendix for details of the `maple` code. We know from [5] that for a finite metric space in which the points are far apart from one another, the magnitude will essentially be the number of points, this means that $|t\ddot{X}| \to N$ as $t \to \infty$. However, for t not so large, $|t\ddot{X}|$ should give a reasonable approximation to the mythical magnitude $|tX|$ of the space we are actually interested in. Here, empirically, "not so large" seems to mean that the points are at most of order 0.1 units from their nearest neighbour.

Based on the ideas of the previous section we want to compare the magnitude $|tX|$ with the penguin valuation $P(tX)$ as we suspect that they might be the same asymptotically, at least for when X is the closure of an open subset. So for t large, but not too large, we suspect $|t\ddot{X}| \simeq |tX|$ and $|tX| \simeq P(tX)$, thus we would expect $|t\ddot{X}| \simeq P(tX)$ and it is this which is observed in the graphs below. The graphs are plotted on log-log axes so that wide range of orders of magnitude can be observed. It is striking that for the *convex* examples we see that $|t\ddot{X}| \simeq P(tX)$

even for *small t*, providing the motivation for the conjecture that $|X| = P(X)$ whenever X is convex.

3.1. **Squares.** Here we consider a 1×1 square Q which we approximate by a 150×150 grid of points $\ddot{Q}$, this has total number of points, N, of $22,500$. Of course the square Q can be approximated by say a hexagonal grid or a large random set of points inside the square, but these appear to give similar results. Weightings for squares with a square grid approximation were seen in Figures 1 and 2. We scale the finite square $\ddot{Q}$ by a factor of t for various values of t from 0.1 to 1000, numerically calculate the magnitude $|t\ddot{Q}|$ and plot the magnitudes against t in Figure 4. We also plot $P(tQ)$ on the graph, which is given by

$$P(tQ) = \frac{t^2}{2\pi} + \frac{2t}{2} + 1,$$

as the linear intrinsic volume, μ_1, of a planar region is given by half the perimeter, which in this case is $2t$. The curve gives a remarkable fit to the data points and fuels the suspicion that $|X| = P(X)$ when X is a convex set.
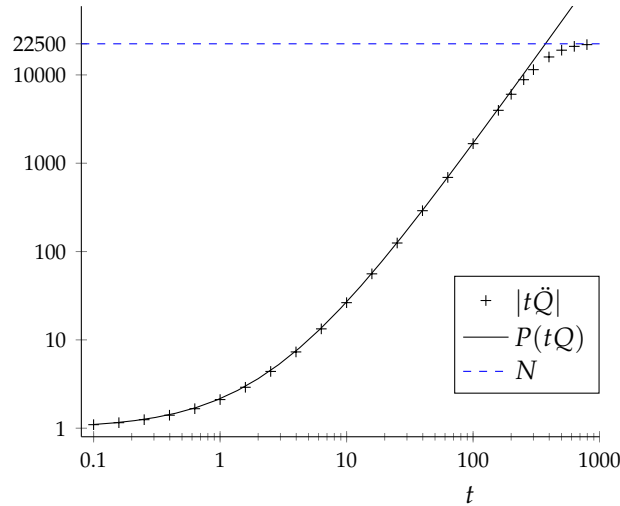

FIGURE 4. Squares: comparison of the penguin valuation $P(tQ) = t^2/2\pi + t + 1$ with $|t\ddot{Q}|$, an approximation to the magnitude.

3.2. **Discs.** Similar to the last example, we consider D to be a circle of radius 1 and take $\ddot{D}$ to be a finite, square grid approximation to D; in this calculation we choose the mesh size of the grid to be such that we have approximately $25,000$ points in D. We numerically calculate $|t\ddot{D}|$ the magnitude of the finite set $\ddot{D}$ scaled by a factor of t, taking various values of t from 0.1 to $1,000$, and in Figure 5 we plot these values along with the graph of $P(tD)$, the suspected asymptotic magnitude of the disc D, where

$$P(tD) = \frac{\pi t^2}{2\pi} + \frac{\pi t}{2} + 1.$$

Again, we get a remarkably good fit with the data points up until the data points start flattening out.
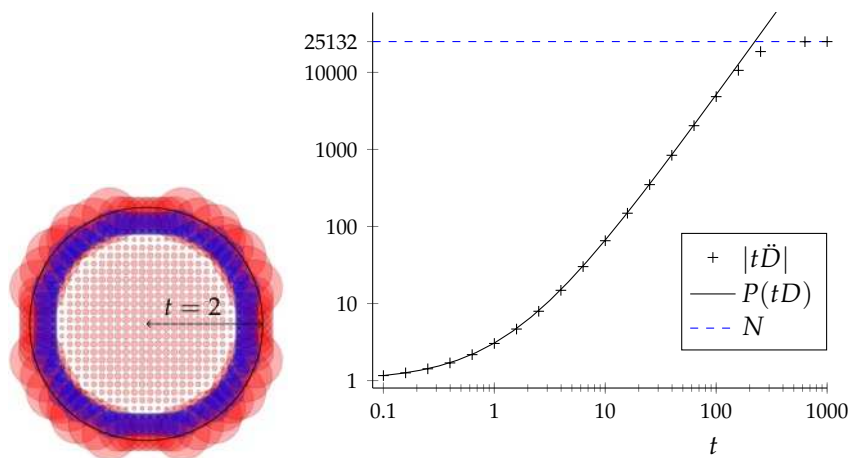
FIGURE 5. Discs: comparison of the penguin valuation $P(tD) = t^2/2 + \pi t/2 + 1$ with $|t\ddot{D}|$, an approximation to the magnitude.
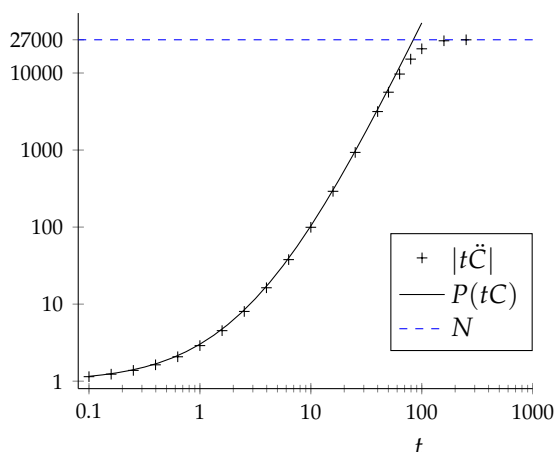

FIGURE 6. Cubes: comparison of the penguin valuation $P(tC) = t^3/8\pi + 3t^2/2\pi + 3t/2 + 1$ with $|t\ddot{C}|$, an approximation to the magnitude.

3.3. **Cubes.** Again we consider a convex set, the cube, but this is one dimension higher than the above two examples of the square and the disc. Let C be the cube of side length 1. Now we approximate it by a $30 \times 30 \times 30$ cubic lattice to get the total number of points, $N = 27,000$, close to our computational maximum; this means that we get a less good approximation to the cube than we had for the square, as the points are five times further away from each other and, consequently, we see that the data points flatten out at smaller values of t than for the square. This is a problem of looking at higher dimensional sets. However, there is still a very good fit with the graph of the penguin valuation

$$P(tC) = \frac{t^3}{8\pi} + \frac{3t^2}{2\pi} + \frac{3t}{2} + 1,$$

for t from 0.1 to 40.

3.4. **Annuli.** We now look at our first non-convex example. We consider A, the annulus with outer radius 1 and inner radius 1/2. This time we approximate not with a square grid but with $\ddot{A}$ a grid in polar coordinates. In Figure 7 we compare the calculated the calculated magnitudes with

$$P(tA) = \frac{\frac{3}{4}\pi t^2}{2\pi} + \frac{\frac{3}{2}\pi t}{2}.$$

In this case we don't see a good fit at small t, which we didn't necessarily expect anyway, but we certainly see a good fit for medium-sized t.
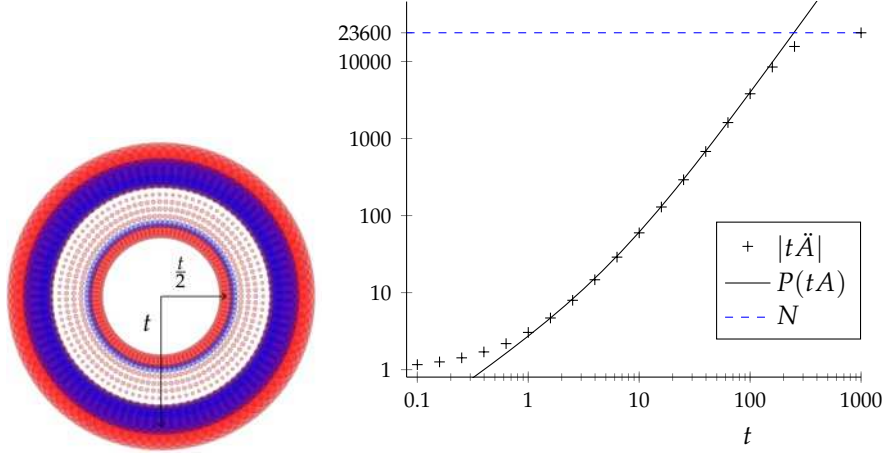

FIGURE 7. Annuli: comparison of the penguin valuation $P(tA) = 3t^2/8 + 3\pi t/4$ with $|t\ddot{A}|$, an approximation to the magnitude.

3.5. **Tori.** In this example we consider a subset of $\mathbb{R}^3$ which does not fit the hypotheses of the Bulk Approximation: the torus is a closed submanifold and is not the closure of an open subset of $\mathbb{R}^3$. Nonetheless, the torus does seem to have the asymptotic behaviour predicted by the Bulk Approximation. We consider U to be the standardly embedded torus with major radius 1 and minor radius 1/5, in other words the locus of points

$$\left((1 + \tfrac{1}{5}\cos\varphi)\cos\theta, (1 + \tfrac{1}{5}\cos\varphi)\sin\theta, \tfrac{1}{5}\sin\varphi\right) \quad \text{for } (\theta, \varphi) \in [0, 2\pi) \times [0, 2\pi).$$

We get the finite approximation $\ddot{U}$ by taking (θ, φ) to lie on a $m \times m$ grid of points in $[0, 2\pi) \times [0, 2\pi)$. In Figure 8 we compare the resulting magnitude data with the graph of the function

$$P(tU) = \frac{\pi t^2/5}{2\pi}$$

and again see a good fit before the magnitudes of the finite space flatten out.

4. SIERPINSKI GASKET

This example is of a different nature to the ones in Section 3 so is given a section of its own. In [5] we considered the fractal example of the ternary Cantor set. It is not a polyconvex set and we can not make sensible interpretation of the intrinsic volumes $\{\mu_i\}_{i=0}^{\infty}$. However we observed that the magnitude of Cantor sets asymptotically satisfies an inclusion-exclusion principle in the following sense. Suppose P is some function on a set of subsets of Euclidean space which satisfies the inclusion-exclusion principle, so that $P(A) + P(B) - P(A \cap B) = P(A \cup B)$. If
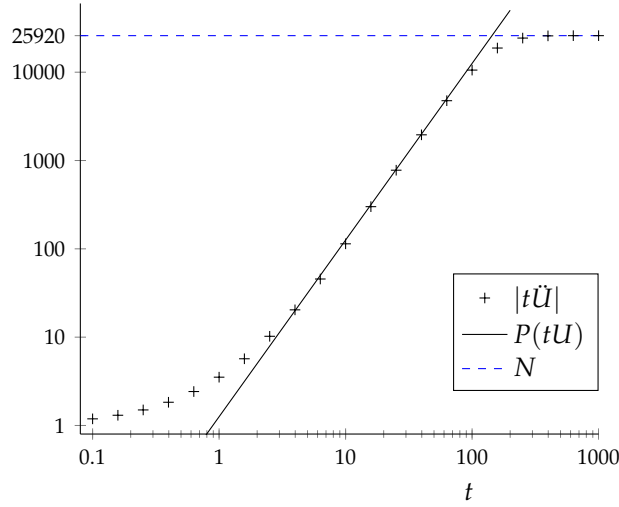
FIGURE 8. Tori: comparison of the penguin valuation $P(tU) = t^2/10$ with $|t\ddot{U}|$, an approximation to the magnitude.

T_t denotes the Cantor set of length t, then we have the self-similarity homeomorphism for all $t > 0$:

$$T_{3t} \cong T_t \sqcup T_t.$$

Thus if the Cantor sets are in the domain of P and we set $p(t) := P(T_t)$ then

$$p(3t) = 2p(t).$$

The general solution to this functional equation is of the form $p(t) = f(t)t^{\log_3 2}$, where f is a multiplicatively periodic function with $f(t) = f(3t)$ for all $t > 0$. In [5] we gave the explicit formula for such a function p so that, as $t \to \infty$, asymptotically, $|T_t|$ is $p(t)$; the corresponding function f in that case is almost-constant, $f(t) \simeq 1.3$ for all $t > 0$.

Now we can consider the Sierpinski gasket. Unlike the case of the Cantor set we cannot calculate the magnitude exactly so we will have to numerically approximate it. We can, however, carry out a similar analysis involving the inclusion-exclusion principle. If S_t denotes the Sierpinski gasket of side length t then for $t > 0$ the self-similarity can be represented by the homeomorphism

$$S_{2t} \cong S_t \sqcup S_t \sqcup S_t \smallsetminus \{3 \text{ points}\},$$

so if P a function satisfying the inclusion-exclusion principle, with the Sierpinski gaskets within its domain and with $P(\text{point}) = 1$, then defining $\widehat{p}(t) := P(S_t)$ we obtain the functional equation for all $t > 0$:

$$\widehat{p}(2t) = 2\widehat{p}(t) - 3.$$

The general solution to this is straightforwardly seen to be be $\widehat{p}(t) = \widehat{f}(t)t^{\log_2 3} + 3/2$, where $\widehat{f}$ is a multiplicatively periodic function with $\widehat{f}(2t) = \widehat{f}(t)$ for all $t > 0$. Based on the behaviour of the Cantor sets, we might expect then that the magnitude $|S_t|$ of the Sierpinski gaskets asymptotically takes this form as $t \to \infty$. We might also expect that the corresponding $\widehat{f}$ might is almost-constant.

For various values of t the magnitude $|S_t|$ was approximated by numerically calculating the magnitude of $\ddot{S}_t$, a finite approximation to S_t with 9843 points. The results are plotted in Figure 9 where $\frac{1}{3}t^{\log_2 3} + \frac{3}{2}$ is also plotted. Here the

function $\widehat{f}$ was taken to be the constant $1/3$ purely because that gave a good fit by eye to the calculated data.
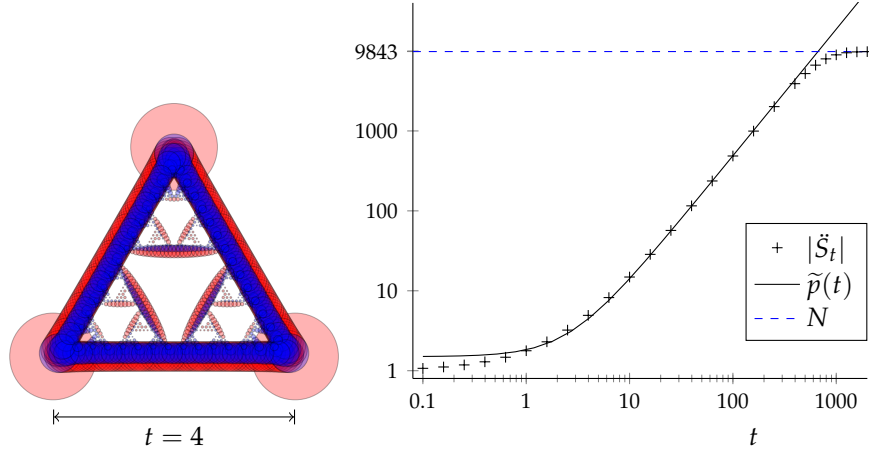

FIGURE 9. Sierpinski gasket: comparison of the calculated magnitudes $|\ddot{S}_t|$ with the function $\widetilde{p}(t) := \frac{1}{3}t^{\log_2 3} + \frac{3}{2}$.

The plot supports the guess that we have an asymptotic inclusion-exclusion principle for the magnitude, so there is an almost-constant function $\widehat{f}$ with $\widehat{f}(t) \simeq \frac{1}{3}$ for all $t > 0$, giving rise to the corresponding function $\widehat{p}$ as above with

$$|S_t| - \widehat{p}(t) \to 0 \qquad \text{as } t \to \infty.$$

One can observe that this is consistent with the more general guess that for X a compact subset of $\mathbb{R}^n$, the asymptotic growth rate of the magnitude $|tX|$, i.e.,

$$\lim_{t \to \infty} \frac{\log |tX|}{\log t},$$

is the Hausdorff dimension of X. This is in agreement with the notion that for X the closure of an open subset of $\mathbb{R}^n$ we have

$$|tX| = \frac{t^n \operatorname{vol}(X)}{n! \, \omega_n} + \text{lower order terms}.$$

The next question to ask is what is the "coefficient" of the highest power of t is in $|tX|$. Whereas in some cases it appears to be proportional to the volume of X, in the fractal case it isn't even necessarily a constant but is an almost-constant function. This is more food for thought.

REFERENCES

[1] David Attenborough, Life in the Freezer, BBC television documentary, Episode 5, The Big Freeze, 1993. See `http://www.youtube.com/watch?v=cgGL6SbMfKc`
[2] H. Federer, *Curvature measures*, Transactions of the American Mathematics Society **93** (1959), 418–491.
[3] D. A. Klain, G.-C. Rota, *Introduction to Geometric Probability*, Cambridge University Press, 1997.
[4] T. Leinster, *Metric spaces*, post at The *n*-Category Café, February 2008. `http://golem.ph.utexas.edu/category/2008/02/metric_spaces.html`
[5] T. Leinster and S. Willerton, *On the asymptotic magnitude of subsets of Euclidean space*. arXiv:0908:1582
[6] A. Solow, S. Polasky, *Measuring biological diversity*, Environmental and Ecological Statistics **1** (1994), 95–107.

Appendix A. Computer considerations

The programming code involved in finding a weighting for a finite metric space is very simple as we are just solving N linear equations in N unknowns, where N is the number of points. Computer algebras have good algorithms built-in for solving such problems, so we can just rely on this. This can be done as follows. If X is a metric space with points called $P_1, \dots, P_N$, then define Z to be the matrix of exponentiated distances, define w to be the vector of weights and define $\mathbf{1}$ to be the vector of ones, so

$$Z_{ij} := \exp\big(-d(P_i, P_j)\big); \quad w := (w_{P_1}, \dots, w_{P_N})^{\mathrm{T}} \in \mathbb{R}^N; \quad \mathbf{1} := (1, \dots, 1)^{\mathrm{T}} \in \mathbb{R}^N.$$

The weight equations are then expressed as the single vector equation

$$Zw = \mathbf{1}.$$

We need to solve this for w and add up the entries of w to obtain the magnitude of X. Here is a piece of pseudo `maple` code, so you can see how straightforward it is.

```
 # Load in a library so we can use Matrix, Vector and LinearSolve
with(LinearAlgebra):

 # Define N to be the number of points

 # Define P to be a list with N entries so that
 # P[i]:=[i-th x-coordinate, i-th y-coordinate]:
 # or possibly with a z-coordinate as well

 # Define the distance d(a,b) between
 # points a=[a[1],a[2]] and b=[b[1],b[2]],
 # adding a third term if working in 3-space
d:=(a,b)->sqrt((a[1]-b[1])^2+(a[2]-b[2])^2):

 # Define the matrix of exponentiated distances
Z:=Matrix(1..N,1..N,
          (i,j) -> exp(-d(P[i],P[j])),
          datatype=float, shape=symmetric):

 # Solve the weight equations and put the solution in the vector w
 # so that w[i] is the weight of the i-th point
w:=LinearSolve(Z,Vector(1..N,1));

 # Add all of the weights together to give the magnitude
Magnitude:=add(w[i],i=1..N):
```

To get a good approximation we want to take the number of points, N, to be as large as possible. The principal constraints on how large we can take N to be are those of computing time and available memory. I was using the `iceberg` cluster at the University of Sheffield, each processor there has 12GB RAM and that turned out to be the main constraint. Taking N to be around $25,000$ resulted in between 3GB and 12GB RAM being used, and between 5 and 12 hours of CPU time for the calculation of a single magnitude. Each data point plotted in Section 3 required that level of resources. Interestingly the metric spaces which needed the most time and memory were those near the 'shoulder' in each of the graphs of Section 3 where the data-points begin to flatten out at the top.

I should add that I have given no thought to how accurate the answer maple gives is, in terms of what sort of rounding errors would be present after row reducing a $25,000 \times 25,000$ matrix, but, nonetheless, the numerical answers obtained looked sufficiently convincing.

E-mail address: s.willerton@sheffield.ac.uk

School of Mathematics and Statistics, Hicks Building, University of Sheffield, Sheffield, S3 7RH, UK