

Cut-out sets and the Zipf law for fractal voids

José Gaite*

Instituto de Matemáticas y Física Fundamental, CSIC, Serrano 113bis, 28006 Madrid, Spain

Received 19 December 2005; received in revised form 28 August 2006; accepted 25 September 2006

Available online 25 October 2006

Communicated by C.K.R.T. Jones

Abstract

“Cut-out sets” are fractals that can be obtained by removing a sequence of disjoint regions from an initial region of d -dimensional euclidean space. Conversely, a description of some fractals in terms of their void complementary set is possible. The essential property of a sequence of fractal voids is that their sizes decrease as a power law, that is, they follow Zipf’s law. We prove the relation between the box dimension of the fractal set (for $d \leq 3$) and the exponent of the Zipf law for *convex* voids; namely, if the Zipf law exponent e is such that $1 < e < d/(d-1)$ and, in addition, we forbid the appearance of *degenerate* void shapes, we prove that the corresponding cut-out set has box dimension d/e (such that $d-1 < d/e < d$). We explore various physical applications of this result, in particular, the application to the description of the cosmic structure using “cosmic foam” models.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Fractals; Zipf’s law

1. Introduction

Since the pioneering work of Mandelbrot on fractal geometry and its applications [1], scale invariance has been observed in a broad range of physical structures. Some of the classical applications are described in Mandelbrot’s book [1] or in Feder’s book [2]. In general terms, scale invariance is related to the appearance of power laws. However, the relation of a power law to an underlying fractal geometry may not be direct. In particular, it is common that power laws appear in rank-ordering statistics (Zipf’s law) [8,9], but there may be no connection with an underlying fractal geometry, in the strict spatial sense of this notion. Here we shall analyse fractal holes or *voids*, which constitute a realization of the Zipf law that arises directly from fractal geometry; namely, it arises from the concept of “cut-out sets”.

Some fractals can be obtained by removing an infinite sequence of disjoint regions from an initial set. Mandelbrot [1] introduced the concept of fractal holes (using the Greek word *tremas*) and showed that the distribution of one-dimensional holes (gaps) follows a simple power law; namely, the number

of gaps of length U greater than u is $N(U > u) \propto u^{-D}$, where D is the fractal dimension. Falconer [3] has studied the fractal properties of “cut-out sets” in terms of box dimensions. In particular, he has proved that fractal sets in one dimension, that is, fractal subsets of \mathbb{R} , have box dimensions that depend on the size of the complementary intervals and not on their arrangement. To be precise, the fractal E with a sequence of void intervals a_k ($k = 1, 2, \dots$) has

$$\dim_B E = -1 / \lim_{k \rightarrow \infty} (\log a_k / \log k)$$

if and only if the limit exists. Roughly speaking, this theorem says that fractality in \mathbb{R} is related to Zipf’s power law [8] for the rank-ordering of void intervals (with exponent $1/\dim_B E$). In turn, Zipf’s law is equivalent to the law $N(U > u) \propto u^{-D}$ if we identify D with $\dim_B E$.

Mandelbrot and Falconer have also considered the generalization to higher dimension, that is, to fractal subsets of \mathbb{R}^d . However, a void in a subset of \mathbb{R}^d is not only defined by its size, so the problem is more complex. One can prescribe a definite shape of voids, for example, discs for $d = 2$, so that the size determines the void (up to its position). Then it still holds that the box dimension of a cut-out set depends only on the sequence of sizes of the complementary intervals and does not depend on their arrangement. However, whereas void regions can be freely

* Tel.: +34 915 616800/3104.

E-mail address: jgaite@imaff.cfmac.csic.es.

rearranged in $d = 1$, in $d > 1$ there are constraints. Examples of two-dimensional cut-out sets are the Apollonian packings of discs [1,3].

In addition, while in $d = 1$ there is no restriction on the fractal (Hausdorff–Besicovitch or box) dimension of cut-out sets, except that it be between zero and one, in $d > 1$, cut-out sets must have *topological codimension* 1, so they are formed by curves in two dimensions or surfaces in three dimensions (etc.) which enclose the voids. Therefore, their fractal dimension must be larger than $d - 1$.

Apart from the notion of cut-out sets, it is possible to define a sequence of void regions in a finite approximation to a fractal set by algorithmic methods that exploit the recursive structure of fractals [4,5]. These algorithms are called *void-finders*. Thus, the question is how to determine the relation between the voids found in this way and the actual fractal set. This question includes, for example, discerning in what sense (if any) a general fractal can be considered a cut-out set of the voids thus found.

The motivation for constructing void-finders has arisen in cosmology, since the observation of large voids in the distribution of galaxies, as a counterpart of galaxy clustering [6]. The fractal features of the distribution of galaxies [7] motivated us to study the scaling of voids in that distribution and, in general, the properties of fractal voids [4,5]. In particular, the already compiled catalogues of voids in the distribution of galaxies are suitable for treatment by standard rank-ordering techniques [8,9] in which voids are ranked by size. In Ref. [4] we defined voids as having constant shape: discs or squares, in two dimensions (the disc case is connected with Apollonian packings). This definition is adequate for mathematical treatment, but we found that it is not satisfactory for analysing a general fractal: often equal-shape voids of similar size touch each other and visually it seems that they should be merged into a unique void of *irregular* shape. This suggested improving the definition of void by considering voids of arbitrary shape. In particular, in Ref. [5], we devised a void-finding algorithm based on discrete geometry methods, namely, Delaunay and Voronoi tessellations. It produces a sequence of voids of polygonal but otherwise arbitrary shape (for $d = 2$). The rank-ordering of the voids of suitable examples of fractal point sets in one and two dimensions showed that Zipf's power law holds; but we did not attempt to provide any mathematical proof of this.

So the purpose of this work is: (i) to extend the notion of cut-out set for $d > 1$ to void regions of non-constant shape and to establish the relation of the box dimension of the set to the Zipf law for voids, following Falconer's methods [3]; (ii) to see in what sense and to what extent a general fractal can be considered a cut-out set of the voids found by algorithmic methods; (iii) finally, to indicate how to apply our conclusions to physical fractal structures.

To achieve these goals, we will need to introduce concepts of integral geometry [10]. These concepts, in particular, the Minkowski functionals, have already been applied to fractal distributions by Mecke [11]. His point of view, based on

clusters, can be considered dual to the point of view that we adopt here, based on voids.

2. Cut-out sets in more than one dimension

Cut-out sets are obtained by removing an infinite sequence of disjoint regions from an initial set. In one dimension, a cut-out set is a subset of \mathbb{R} obtained by removing an infinite sequence of disjoint open intervals from an initial closed interval such that the sum of the lengths of the removed intervals converges to the total length of the initial interval. A trivial example is the Cantor set. Let the fractal $E \in \mathbb{R}$ be the result of removing from the closed interval A a sequence of open disjoint intervals $\{A_k\}_{k=1}^{\infty}$ such that the length of A_k is a_k and $\sum_{k=1}^{\infty} a_k$ equals the length of A . It is not difficult to see that the set obtained by removing a finite number of intervals of decreasing length is related to a neighbourhood of E and so the pattern of sizes of removed intervals is related to the behaviour of the r -neighbourhood of E as $r \rightarrow 0$. If this behaviour is a power law, it defines the Minkowski–Bouligand dimension, which is equivalent to the box dimension [1,12]. A careful analysis then shows that the corresponding condition on the lengths of removed intervals is that they follow a power law in their rank, that is, that they follow Zipf's law [8]. To be precise, the fractal $E = A - \bigcup_{k=1}^{\infty} A_k$, with the decreasing sequence of lengths $\{a_k\}_{k=1}^{\infty}$, has box dimension

$$\dim_B E = -1 / \lim_{k \rightarrow \infty} (\log a_k / \log k)$$

if and only if the limit exists.

In more than one dimension, a cut-out set is a subset of \mathbb{R}^d obtained by removing an infinite sequence of disjoint connected open regions $\{A_k\}_{k=1}^{\infty}$ from an initial compact region A , which it is natural to choose convex [3]. We need to restrict the possible shapes of the A_k . The simplest option is to prescribe a definite shape for them. Before we analyse the problem in detail, let us consider a particular type of fractal with voids of constant shape.

2.1. Cantor-like fractals and merging criteria

For Cantor-like fractals, which are strictly self-similar, Zipf's law for voids is a consequence of their construction [4]. A (deterministic) Cantor set is constructed with a generator characterized by two numbers, namely, the scaling factor r and the number of pieces N to remain, and by the arrangement of these remaining pieces. In one dimension, we convene that if the arrangement of the removed $1/r - N$ open intervals is such that two or more are adjacent then one also removes the isolated points between them (*merging* criterion). This criterion leads to characterizing the fractal generator by one more number: its number of voids (gaps) $m \leq 1/r - N$ [4]. If we do not apply the merging criterion, the resulting countable set of isolated points does not contribute to the Hausdorff–Besicovitch dimension of the Cantor-like set but may contribute to its box dimension [12].

In d dimensions, the merging criterion generalizes to the removal of the $d - 1$ -dimensional boundaries between adjacent open d -cubes. However, a complication arises: it is

also possible that some open d -cubes touch the boundary of the initial closed d -cube. If we also convene to remove the corresponding $d - 1$ -dimensional boundaries, then, when we iterate the generator, we merge voids of different levels. Actually, all the void regions merge and form one connected void region. We may not allow this by preserving in the generator the boundary of the initial closed d -cube. This boundary has topological codimension 1, so the box and Hausdorff–Besicovitch dimensions of the resulting Cantor-like set are larger than $d - 1$. This holds even when the expected value of the box and Hausdorff–Besicovitch dimensions, $-\log N / \log r$, is smaller than $d - 1$, so then the boundary preserving criterion alters them significantly. Even if the expected Hausdorff–Besicovitch dimension is larger than $d - 1$, so that it may not be altered by this criterion, the structure of the resulting Cantor-like set may change significantly.

For example, consider the two-dimensional middle-third Cantor set, constructed as the cartesian product of two middle-third Cantor sets and with dimension $2 \log 2 / \log 3 > 1$; it has zero topological dimension, that is, it is a set of points rather than lines. Its construction with the two-dimensional generator consisting of removing the appropriate pattern of five open squares from the unit closed square with the boundary preserving criterion makes it a set of lines (with the same box and Hausdorff–Besicovitch dimensions as the point set). This happens irrespective of any criterion for merging across inter-cube boundaries: if no merging is prescribed, then there are additional lines.

Of course, if the generator does not include any open d -cube that touches the boundary of the initial closed d -cube, cubes removed in different iterations cannot merge. This condition ensures that the fractal set has topological codimension 1. Typical fractals generated in this form are the Sierpinski carpets [1]. Assuming inter-cube merging, it is convenient to introduce the number m of voids of the generator, such that $1 \leq m \leq 1/r^d - N$. The Hausdorff–Besicovitch dimension $D = -\log N / \log r > d - 1$ is independent of m and the sizes a_k follow a power law with exponent $-d/D$ (on average) [4,5].

Note that the inter-cube merging criterion implies that voids in the generator may not have a common shape but a finite number of shapes instead. However, this finite number of void shapes is conserved in the iterations, which only change their size.

2.2. Voids in general fractals and cut-out sets

We have seen in the preceding examples of Cantor-like fractals that a precise definition of voids demands them to have a precise boundary and, therefore, the corresponding fractal set has topological codimension 1, since it contains the boundaries which have themselves codimension 1. Fractals of topological codimension 1 are curves in $d = 2$, surfaces in $d = 3$, etc. One may wonder whether a codimension-1 fractal set has an associated set of voids that make it equivalent to a cut-out set, and whether those voids fulfill Zipf's law. Since we take the initial compact region A of a cut-out set to be convex, we may consider the convex hull of the fractal and the complementary

set of the fractal with respect to it. If this complementary set is formed by an infinite set of connected regions, this set is the natural set of voids. For example, this approach works for the von Koch curve, whose set of voids can be shown to fulfill Zipf's law by relying on its self-similarity. However, these voids have themselves fractal boundary and, therefore, are not sufficiently simple for our purposes.

A different approach is algorithmic, namely, to define fractal voids as the regions found by void-finders. These regions may have constant shape, the obvious shapes being a square or a disc [4]. Or we may allow arbitrary shapes; in particular, we have devised a new void-finder, based on discrete geometry methods (Delaunay and Voronoi tessellations), which produces a sequence of voids of arbitrary polygonal or polyhedral shape, respectively, in $d = 2$ or in $d = 3$ [5].

The problem with the constant-shape voids found in an arbitrary fractal is that they may not fill its real voids of variable shape, if they exist. For example, let us assume that a fractal for $d = 2$ has a square void that we are trying to fill with discs. This filling will form the Apollonian packing of the square, which leaves out a fractal region of dimension about 1.31 [3]. Then it may happen that this dimension is larger than the dimension of the original fractal with the square void. In contrast, arbitrary polygonal (or polyhedral) shapes are adaptable. Therefore, if a fractal has topological codimension 1 and so has well-defined voids, appropriate polygonal (or polyhedral) shapes will reproduce them with sufficient approximation. However, this approach still allows for too general void shapes because the voids may have fractal boundary and diverging perimeter length (in $d = 2$, or diverging surface area in $d = 3$).

Arbitrary polygonal (or polyhedral) void shapes seem a suitable starting point but they are still too general. We would like them to have a bounded perimeter length (or surface area) for a given *diameter* (we define diameter as the greatest distance apart of pairs of points [12]). A nice way of implementing this condition without being too restrictive, that is, of preventing the boundary of voids from becoming wrinkled, is to require them to be convex. So we define henceforth a cut-out set as a subset of \mathbb{R}^d obtained by removing an infinite sequence of disjoint open convex d -polyhedral regions $\{A_k\}_{k=1}^\infty$ from an initial compact convex region A . In the end, the restriction to d -polyhedra proves to be superfluous and we may consider general convex voids. However, from a computational point of view, that is, when finding voids in a finite set of points, these voids must be polyhedra. In the next section, we generalize the results of Falconer regarding the box dimension of fractals resulting from cutting out discs [3] to these more general cut-out sets, restricting ourselves to $d = 2$ as well. Further generalization to $d = 3$ is achieved in the following section.

2.3. Two-dimensional cut-out sets with convex polygonal voids

Let A be a plane compact convex region of perimeter p and let $\{A_k\}_{k=1}^\infty$ be a sequence of disjoint open convex polygonal regions, such that A_k has diameter δ_k , perimeter p_k , and area a_k . Contrary to the case of discs (or other constant shapes), these three quantities are independent. So there is no obvious order

of the A_k . Let $\{A_k\}_{k=1}^\infty$ be such that the total area $\sum_{k=1}^\infty a_k$ is equal to the area of A . Then the set $E = A - \bigcup_{k=1}^\infty A_k$ has zero area. To calculate its box dimension, we can use the Minkowski–Bouligand dimension [1,12], which is equivalent to the box dimension [12].

The Minkowski–Bouligand dimension expresses the power-law behaviour of the r -neighbourhood of E as $r \rightarrow 0$. To apply it to the cut-out set E , one needs to distinguish voids that are included in the r -neighbourhood of E from voids that are not included. In any one of the latter, the r -neighbourhood of E is a band of width r that leaves an empty part. We can try to estimate the areas of these bands, in order to obtain the area of the r -neighbourhood of E . Before doing so, let us recall Falconer's formula for the area of the r -neighbourhood of E in the case of disc-shaped voids.

2.3.1. Area of the r -neighbourhood of E with disc-shaped voids

Falconer's formula [3] for the area of the r -neighbourhood of E in the case of disc-shaped voids reads:

$$V(r) = (pr + \pi r^2) + \sum_{i=1}^k \pi(r_i^2 - (r_i - r)^2) + \sum_{i=k+1}^\infty \pi r_i^2 \quad (1)$$

$$= pr + 2\pi r \sum_{i=1}^k r_i + \pi \sum_{i=k+1}^\infty r_i^2 + \pi r^2(1 - k), \quad (2)$$

where k is such that r is between the disc radii of indices k and $k+1$ ($r_{k+1} \leq r \leq r_k$). The first term of Eq. (1) is the area of a band surrounding A , the second term is the area of the annuli of width r inside the discs A_i for $1 \leq i \leq k$, and the third term the area of the discs A_i for $i \geq k+1$. In the second expression, Eq. (2), the second term, which is the area of the annuli of width r except for $-k\pi r^2$ (added to the last term), represents the sum of the perimeters of the discs A_i for $1 \leq i \leq k$ multiplied by r .

Falconer assumes that the disc radii and, therefore, areas fulfill Zipf's law, with a concrete formulation, namely, $r_k \asymp k^{-\alpha}$, in terms of the relation \asymp . The relation $a_k \asymp b_k$ between two sequences $\{a_k\}$ and $\{b_k\}$ means that there are two constants $c_1, c_2 > 0$ such that $c_1 \leq a_k/b_k \leq c_2$ for all k (this concept also applies to functions having the relation $f(x) \asymp g(x)$ with analogous meaning). Falconer further assumes that $1/2 < \alpha < 1$. The inequality $\alpha > 1/2$ ensures that the series of disc areas converges. The role of the inequality $\alpha < 1$ is more subtle: it implies that the series of disc perimeters diverges. So, for $r_{k+1} \leq r \leq r_k$, we have

$$\begin{aligned} V(r) &\asymp r + r \sum_{i=1}^k i^{-\alpha} + \sum_{i=k+1}^\infty i^{-2\alpha} - r^2 k \\ &\asymp r k^{1-\alpha} + k^{1-2\alpha} \asymp r^{2-1/\alpha}, \end{aligned} \quad (3)$$

and

$$\dim_B E = 2 - \lim_{r \rightarrow 0} \frac{\log V(r)}{\log r} = \frac{1}{\alpha} > 1.$$

If $\alpha > 1$ the series of perimeters converges and the second term in Eq. (2) behaves like r , dominating over the terms that

behave like $r^{2-1/\alpha}$. This implies that $\dim_B E = 1$, instead of $\dim_B E = 1/\alpha < 1$. We encounter again that a cut-out set must have dimension larger than $d - 1$, since this is the topological dimension of the boundaries of voids.

2.3.2. Area of the r -neighbourhood of E with triangular voids

Before proceeding to analysing the case of general convex polygonal voids, let us analyse the simpler case of triangles (which are necessarily convex).

Let us consider voids that are not included in the r -neighbourhood of E . In any of these voids, the r -neighbourhood of E forms a band of width r that leaves an empty part, which we can calculate. To be precise, the r -neighbourhood of the boundary of a triangle is a band of width r so the region uncovered is another triangle with the same angles and, therefore, similar. The similarity ratio can be determined employing some elementary geometry. The vertices of the similar triangles for various r are placed along the angle bisectors and so the triangles themselves are homothetical with respect to the common centre of their inscribed circles. The area of any of these triangles is given in terms of its perimeter p and the radius of its inscribed circle ρ by $a = p\rho/2$. If ρ refers to the original triangle and ρ' to another, the similarity ratio is ρ'/ρ . So the area of the latter is $a' = a(\rho'/\rho)^2$ and $a - a' = a[1 - (\rho'/\rho)^2]$. Given that $r = \rho - \rho'$, we obtain for the area of the band of width r

$$a - a' = pr - \frac{p r^2}{2\rho} = pr - \frac{p^2 r^2}{4a}. \quad (4)$$

In this formula, the factor $p^2/(4a)$ of r^2 depends only on the angles and is a measure of shape. The value of a' decreases as r increases and vanishes for the maximal $r = \rho$.

The preceding results show us the appropriate version of formulas (1) and (2) for the area of the r -neighbourhood of E :

$$V(r) = (pr + \pi r^2) + \sum_{i=1}^k \left(p_i r - \frac{p_i^2 r^2}{4a_i} \right) + \sum_{i=k+1}^\infty a_i \quad (5)$$

$$= pr + r \sum_{i=1}^k p_i + \sum_{i=k+1}^\infty a_i + r^2 \left(\pi - \sum_{i=1}^k \frac{p_i^2}{4a_i} \right), \quad (6)$$

where k is such that $\rho_{k+1} \leq r \leq \rho_k$. To proceed like in the case of discs, we would like to have

$$\rho_k \asymp k^{-\alpha}, \quad p_k \asymp \rho_k, \quad a_k \asymp \rho_k^2.$$

So we must have that $a_k \asymp p_k^2$, that is, that the perimeter to area ratio $p^2/(4a)$ is bounded above and below. According to the basic *isoperimetric inequality*, this quantity has an absolute lower bound of π , reached by a disc (for triangles the lower bound is larger, namely, $3\sqrt{3}$, reached by the equilateral triangle) [10]. The upper bound must be explicitly imposed. Then, from $\rho_k \asymp k^{-\alpha}$ and $a_k \asymp p_k^2$ we deduce $p_k \asymp k^{-\alpha}$. Therefore, a proof analogous to Falconer's proof for disc voids leads to the same result, namely, $\dim_B E = 1/\alpha$ ($1/2 < \alpha < 1$).

An upper bound to the perimeter to area ratio $p^2/(4a)$ is an intuitively reasonable requirement. This ratio only depends on the angles of the triangle and an upper bound to it is equivalent

to a lower bound to them. In other words, we are excluding “spiky” triangles, which are nearly one-dimensional and can be packed in a small area without reducing their diameter.

2.3.3. Area of the r -neighbourhood of E with convex polygonal voids

In this case, in a void convex polygon not included in the r -neighbourhood of E , the r -neighbourhood also forms a band of width r that leaves an empty part, but its area is harder to calculate. Therefore, instead of attempting to derive an equation analogous to Eq. (6), we look for independent bounds to $V(r)$. A lower bound is certainly provided by the sum of the areas of polygons fully included in the r -neighbourhood. Finding an upper bound to $V(r)$ requires us to estimate the area of bands of width r inside larger polygons, and the bound will be given by an expression simpler than the right-hand side of Eq. (6).

To find a precise expression for the lower bound to $V(r)$, we need a criterion to determine when a void polygon is fully included in the r -neighbourhood of E . A simple criterion is given by the diameter δ of the polygons: if $r \geq \delta$, then the void polygons of diameter equal to or smaller than δ are certainly covered. So, if we order the void polygons by their diameter and k is such that $\delta_{k+1} \leq r \leq \delta_k$,

$$\sum_{i=k+1}^{\infty} a_i \leq V(r). \quad (7)$$

To find a precise expression of the upper bound to $V(r)$, we need a criterion to determine when a void polygon is *not* fully included in the r -neighbourhood of E . Take a particular polygon A_i , with area a_i and perimeter p_i . Given that the area of a band of width r inside A_i and around its boundary is smaller than $p_i r$, a sufficient condition is that $r \leq a/p$. Therefore, an upper bound to $V(r)$ is

$$V(r) \leq pr + r \sum_{i=1}^k p_i + \sum_{i=k+1}^{\infty} a_i + \pi r^2, \quad (8)$$

where the void polygons are ordered by their value of a/p and k is such that

$$\frac{a_{k+1}}{p_{k+1}} \leq r \leq \frac{a_k}{p_k}.$$

We again need the relation $a_k \asymp p_k^2$ to relate the second and third terms of inequality (8). This condition implies a lower bound to the angles of polygons, like for triangles. However, the condition is now stronger: it is possible to have “spiky” convex polygons (with large diameter to area ratio) with non-small angles; a simple example is the rectangle, with no upper bound to the ratio $p^2/(4a)$.

Then, let us impose the conditions

$$p_k \asymp k^{-\alpha}, \quad a_k \asymp k^{-2\alpha},$$

like we did for triangles. We have, for $a_{k+1}/p_{k+1} \leq r \leq a_k/p_k$,

$$\begin{aligned} V(r) &\leq c \left(r \sum_{i=1}^k i^{-\alpha} + \sum_{i=k+1}^{\infty} i^{-2\alpha} \right) + \pi r^2 \\ &\leq c' (r k^{1-\alpha} + k^{1-2\alpha}) + \pi r^2 \leq c'' r^{2-1/\alpha}, \end{aligned} \quad (9)$$

for some positive numbers c , c' , c'' , and $r < 1$. On the other hand, we can apply a similar procedure to obtain the lower bound. To do so, we first relate the diameter of a convex polygon with its perimeter. We note that there are both lower and upper bounds to their ratio: $2\delta < p$ and $p \leq \pi\delta$ (which actually hold for any *convex* figure) [10]. So $\delta_k \asymp p_k$. Therefore, for $\delta_{k+1} \leq r \leq \delta_k$,

$$c''' r^{2-1/\alpha} \leq V(r),$$

for some positive number c''' . Both bounds are equivalent to

$$V(r) \asymp r^{2-1/\alpha},$$

implying that $\dim_B E = 1/\alpha$ ($1/2 < \alpha < 1$).

In conclusion, the conditions

$$a_k \asymp k^{-2\alpha}, \quad a_k \asymp p_k^2 \quad (10)$$

seem to be as suitable for convex polygonal voids as for triangular voids or discs. In fact, an approximation argument would show that these conditions are suitable for general convex voids. The rationale for this proof is that the quotient p^2/a is the measure of shape for convex figures and its being bounded ensures that they do not degenerate into one-dimensional figures (segments), so that perimeter and area have the natural scaling behaviour.

2.4. Generalization to three-dimensional cut-out sets

We proceed to the generalization to three-dimensional cut-out sets with convex polyhedral voids. We need to generalize the geometrical properties that we have used from convex polygons to convex polyhedra.

Let A be a compact convex region that is to become a cut-out set. Falconer’s formula (1) for the area of the r -neighbourhood of E can be generalized to three dimensions (ball-shaped voids):

$$\begin{aligned} V(r) &= \left(ar + Hr^2 + \frac{4}{3} \pi r^3 \right) \\ &+ \sum_{i=1}^k \frac{4}{3} \pi (r_i^3 - (r_i - r)^3) + \sum_{i=k+1}^{\infty} \frac{4}{3} \pi r_i^3 \end{aligned} \quad (11)$$

$$\begin{aligned} &= (ar + Hr^2) + 4\pi r \sum_{i=1}^k r_i^2 - 4\pi r^2 \sum_{i=1}^k r_i \\ &+ \frac{4}{3} \pi \sum_{i=k+1}^{\infty} r_i^3 + \frac{4}{3} \pi r^3 (1+k), \end{aligned} \quad (12)$$

where k is such that r is between the ball radii of indices k and $k+1$ ($r_{k+1} \leq r \leq r_k$). The first term of Eq. (11) is the volume of the layer surrounding A , given by Steiner’s formula, where H is A ’s *linear measure* (mean curvature) [10]. The second term is the volume of the shells of width r inside the balls A_i , for $1 \leq i \leq k$, and the third term the area of the balls A_i for $i \geq k+1$. In the second expression, Eq. (12), the second term represents the sum of the areas of the balls A_i for $1 \leq i \leq k$ multiplied by r , while the third term represents the sum of their linear measures multiplied by r^2 .

We assume that $r_k \asymp k^{-\alpha}$, $1/3 < \alpha < 1/2$, ensuring that the series of ball areas diverges while the series of their volumes converges. So, for $r_{k+1} \leq r \leq r_k$, we have

$$\begin{aligned} V(r) &\asymp r + r^2 + r \sum_{i=1}^k i^{-2\alpha} + \sum_{i=k+1}^{\infty} i^{-3\alpha} + r^3 k \\ &\asymp r k^{1-2\alpha} + k^{1-3\alpha} + r^3 k \asymp r^{3-1/\alpha}, \end{aligned} \quad (13)$$

and

$$\dim_B E = 3 - \lim_{r \rightarrow 0} \frac{\log V(r)}{\log r} = \frac{1}{\alpha} > 2.$$

Again, a cut-out set must have dimension larger than $d - 1$ (the topological dimension of the boundaries of voids).

For tetrahedral voids, the r -neighbourhood of the boundary of a tetrahedron is a layer of thickness r and its volume is (in analogy with Eq. (4))

$$v - v' = ar - \frac{a^2 r^2}{3v} + \frac{a^3 r^3}{27v^2}. \quad (14)$$

Here a and v are the area surface and volume of the tetrahedral void, respectively, and v' is the volume of the smaller homothetical tetrahedron. The factor $a^3/(27v^2)$ only depends on the angles of the tetrahedron and is a measure of its shape. The maximal r , such that v' vanishes, is $r = \rho = 3v/a$, that is, the radius of the inscribed sphere. The appropriate relations for tetrahedral voids are

$$\rho_k \asymp k^{-\alpha}, \quad a_k \asymp \rho_k^2, \quad v_k \asymp \rho_k^3.$$

So the relation between surface area and volume to be extended to general convex polyhedra is $v_k \asymp a_k^{3/2}$. The lower bound to $a^{3/2}/v$ is again universal, according to a three-dimensional isoperimetric inequality, and corresponds to a ball (for tetrahedra, to the regular tetrahedron). The upper bound forbids again small angles that give rise to flattened or spiky tetrahedra. Note that a tetrahedron can degenerate either into a two-dimensional or a one-dimensional figure (triangle or segment, respectively), but the former is more generic.

In the case of convex polyhedral voids, in analogy with two dimensions, a lower bound to $V(r)$ is provided by the sum of the volumes of polyhedra fully included in the r -neighbourhood and an upper bound is given by the three-dimensional version of Eq. (8). For the lower bound to $V(r)$, to determine when a void polyhedron is fully included in the r -neighbourhood of E , we use again the criterion given by the diameter δ of the polyhedra, namely, $r \geq \delta$. So, if we order the void polyhedra by their diameter and k is such that $\delta_{k+1} \leq r \leq \delta_k$,

$$\sum_{i=k+1}^{\infty} v_i \leq V(r). \quad (15)$$

For the upper bound to $V(r)$ we need a condition that ensures that a given polyhedron is not included in the r -neighbourhood of E . The volume of the layer of thickness r inside a polyhedron A_i is smaller than $a_i r$. We have (in analogy

with inequality (8) for polygons):

$$V(r) \leq ar + Hr^2 + r \sum_{i=1}^k a_i + \sum_{i=k+1}^{\infty} v_i + \frac{4}{3}\pi r^3 \quad (16)$$

where the void polyhedra are ordered by their value of v/a and k is such that

$$\frac{v_{k+1}}{a_{k+1}} \leq r \leq \frac{v_k}{a_k}.$$

Assuming that

$$v_k \asymp k^{-3\alpha}, \quad v_k \asymp a_k^{3/2}, \quad (17)$$

it follows that

$$V(r) \leq c r^{3-1/\alpha},$$

for some positive number c , and $r < 1$.

For the lower bound to $V(r)$, we may relate the diameter of a convex polyhedron with its linear measure H . Like in the two-dimensional case, there are both lower and upper bounds to their ratio: $\delta < H/\pi$ and $H \leq 2\pi\delta$, which hold for any convex body [10]. H is independent of v and a , but the relations (17) imply $H_k \asymp k^{-\alpha}$ nonetheless. This follows from the two fundamental three-dimensional isoperimetric inequalities: $a^2 \geq 3vH$, $H^2 \geq 4\pi a$ [10]. Therefore, for $\delta_{k+1} \leq r \leq \delta_k$,

$$c' r^{3-1/\alpha} \leq V(r),$$

for some positive number c' . Both upper and lower bound inequalities are equivalent to

$$V(r) \asymp r^{3-1/\alpha},$$

so $\dim_B E = 1/\alpha$ ($1/3 < \alpha < 1/2$).

In conclusion, the conditions (17) are suitable for convex polyhedral voids and, furthermore, an approximation argument would show that they are suitable for general convex voids. Note that our proof relies on the sufficiency of the upper bound to the quotient a^3/v^2 (as a measure of shape) for preventing degeneracy into lower dimension. This holds regardless of the actual existence of two independent measures of shape of three-dimensional convex bodies.

3. Applications in physics

The type of physical fractals corresponding to “cut-out sets” consists of foam-like or web-like fractals. These fractals appear in many areas, not only in physics but also in geology, biology, etc. In physics, they appear in models of gels, porous media, isoscalar surfaces in turbulence, etc. Some of the many applications are described in the book by Feder [2]. At any rate, we note that the theory we have developed only applies to convex voids, while in some phenomena, e.g., percolation, the voids produced have non-convex, fractal boundary. A particular application that we consider in more detail below is the model of cosmic structure that has been called the “cosmic web”.

A different realization of “cut-out sets” is provided by the reverse of foams, namely, by dense packings of granular systems. In this case, the fractal is formed by the empty space

left by the grains, of vanishing volume as the size of filling grains decreases and the density of the packing increases. The sizes of grains (instead of the sizes of voids) fulfill the Zipf law. In particular, three-dimensional Apollonian packings (osculating sphere packings) have been proposed as a model of dense granular systems [13]. Soil grains are usually irregular, but their sizes fulfill the Zipf law [14].

Given the variety of the above mentioned applications, the physical processes that give rise to the corresponding fractal structures are very diverse and can be very complex. Hence, we shall focus on the application to cosmic structure, namely, to the “cosmic web” or “cosmic foam” models, because the dynamics of cosmic structure formation is relatively well understood. This dynamics consists of the growth by gravitational instability of small primordial matter fluctuations. Even though this process is very nonlinear and, therefore, difficult to study with analytic methods, it is arguably scale invariant (within some range of scales), due to the scale invariance of gravity. In addition, void-finding algorithms have been developed regarding cosmic voids, so the connection between the fractal sets and their complementary void sets is more developed in this context.

3.1. Cosmic structure

The fractal geometry of the distribution of galaxies has been studied for years [7], and the study of galaxy clustering actually stimulated the development of fractal geometry [1]. Mandelbrot considered in his book [1] the presence of voids in the distribution of galaxies but, according to the observational situation at that time, favored small voids and, actually, introduced the concept of *lacunarity* to account for that feature. The observation of large voids in the distribution of galaxies is more recent, but the cosmological literature about voids [6] hardly treats their fractal properties. Trying to fill this gap, we began a program to adapt the algorithmic studies of cosmic voids to general fractals and, vice versa, to discern fractal features in cosmic voids, namely, the Zipf law for voids [4,5]. In particular, we devised a void-finding algorithm based on discrete geometry methods that produces a sequence of voids of polyhedral but otherwise arbitrary shape [5]. This void-finder has a free parameter that controls the shape of voids. We showed that it is best to tune it to get compact (non-degenerate) shapes, which are not necessarily convex but tend to be so.

There is a successful model of cosmic structure formation that produces walls as first structures, namely, the adhesion model [15] (the walls are called “pancakes” in this context). The full structure produced by this model is a self-similar pattern of interlocking walls that has been dubbed the “cosmic web” (or the “cosmic foam”). As matter concentrates in the walls, there appear increasingly depleted regions, that is, voids. Due to the self-similarity, these voids form a *hierarchy*, akin to the void distribution given by the Zipf law. However, the model makes no prediction about void shapes.

The definition of void as a totally empty region is surely not adequate for the full matter distribution, since galaxy voids are likely to contain underdense dark matter. It is normally

assumed in cosmology that the galaxy distribution follows the dark matter distribution but it has some *bias*. In the cosmic foam model, it is natural that galaxies form in the regions with higher mass concentration, while voids become more and more depleted but not totally empty. It is also natural that the dark matter and the galaxy distributions share two general characteristics: they are self-similar on small scales and become homogeneous on large scales. Of course, the fractal properties belong to the small-scale self-similar regime, where the dynamics is fully nonlinear. In particular, the size of the largest voids should not be much larger than the homogeneity scale. However, the transition to homogeneity in the rank-ordering of voids appears as a flattening on small ranks of the constant-slope part of its log–log plot [5]; but this flattening is progressive and may span a non-negligible range of void sizes. This remark may help to explain a possible disagreement between the size of the largest galaxy voids, of tens of h^{-1} Mpc, and the homogeneity scale, which is believed to be smaller.

3.1.1. The Voronoi foam model

In connection with the adhesion model, Rien van de Weygaert and collaborators introduced a model based on Voronoi tessellations, in which Voronoi cells represent void regions while the matter is concentrated in their walls (a comprehensive reference is [16]). The cell centres represent void germs (contrary to their role in our void-finding algorithm, in which they correspond to matter particles). This *Voronoi foam model* relies on a model of ellipsoidal void expansion: a set of points that are initially the peaks of the gravitational potential, where the matter is underdense, become the “expansion centres” of matter flowing outwards with *uniform* velocity. Furthermore, when the flow from one void encounters the flow from an adjacent one, a wall forms halfway between their centres. The resulting distribution is a set of Voronoi cells, that is, a “Voronoi foam”. Voronoi cells are convex polyhedra and if the expansion centres form a self-similar pattern, so do the Voronoi cells. Moreover, their shapes are non-degenerate, according to the “bubble theorem” [16]: the evolution of an underdense region is such that its initial slight asphericity decreases (contrary to the evolution of an overdense region); so the evolution of underdense regions is essentially described by the expansion of ellipsoidal voids that become more spherical, until they collapse with other voids, forming walls. Therefore, these Voronoi foams are particular cases of cut-out sets with non-degenerate convex polyhedral voids.

4. Discussion

Zipf’s power-law rank-ordering is naturally associated with scale invariance, but its precise relation to fractal geometry needs to be qualified: fractal void sizes fulfill the Zipf law. We have proved the relation between the box dimension of a cut-out set E and the exponent of the corresponding Zipf law, for convex voids, in particular, for convex polygonal voids in $d = 2$ and convex polyhedral voids in $d = 3$. The particular forms of Zipf’s law for voids in $d = 2$ and $d = 3$ are Eqs. (10) and

(17), respectively. So we have extended Falconer's results for $d = 1$ and $d = 2$ (further extension to convex voids in any dimension d seems straightforward). If the Zipf law exponent is $e (= d\alpha)$, the relation is $\dim_B E = d/e$. Sufficient conditions for this relation to hold are that $1 < e < d/(d-1)$ and, in addition, the exclusion of *degenerate* void shapes.

It is useful to make a few remarks about the box dimension formula. We have emphasized that cut-out sets must have topological codimension 1, that is, topological dimension $d-1$. This is why $e < d/(d-1)$, which implies $\dim_B E = d/e > d-1$, according to the known order of box and topological dimensions: if $e > d/(d-1)$, the $d-1$ -measure of the boundaries of voids converges and $\dim_B E = d-1$. On the other hand, the order of box, Hausdorff–Besicovitch and topological dimensions is $\dim_T E \leq \dim_H E \leq \dim_B E$, so $e > 1$ implies that $\dim_H E < d$. However, we cannot determine whether the inequality $\dim_T E \leq \dim_H E$ is strict, so we cannot tell whether E is in fact a fractal (according to the usual definition). As a one-dimensional example, consider the “convergent sequence sets” $E^{(p)} = \{0, 1, 2^{-p}, 3^{-p}, 4^{-p}, \dots\}$, $p > 0$, with $\dim_B E^{(p)} = 1/(p+1)$ [3]: they have all but one of their points isolated and $\dim_H E^{(p)} = 0$. Let us recall that the box dimension only depends on the size of the voids and not on their arrangement, but the Hausdorff–Besicovitch dimension can be altered by a rearrangement of the voids.

Nevertheless, for the case of cut-out sets with physical interest (Section 3), which are statistically self-similar, we can say that $\dim_H E = \dim_B E$ *almost surely*. The equality holds for strictly self-similar fractals [3], and holds almost surely in the statistically self-similar cases (although the proofs are more involved). Intuitively, it is necessary to arrange the voids in a cut-out set in a very precise fashion to achieve the inequality $\dim_H E < \dim_B E$, while most arrangements produce the equality, that is, produce a cut-out set with the highest Hausdorff–Besicovitch dimension.

The mathematical requirement of having topological codimension 1 can be difficult to test for point sets obtained from physical observations, which are necessarily finite. In fact, the definition of void itself becomes uncertain for a finite set of points. This is why one must resort to void-finding algorithms. These algorithms may have free parameters, giving rise to different sets of voids according to their value. We noted in Ref. [5] the possibility of “percolation of voids”, which must be avoided, by selecting parameter values that produce small, convex-like voids. In this way, when the number of sampling points of a cut-out fractal set grows, we expect the set of voids found to approach the real set of voids. However, questions of convergence are difficult to treat in a rigorous way. If a fractal set E has $\dim_B E > d-1$, but has topological codimension larger than 1, the cut-out set defined by the set of voids found by some algorithm is a fractal set $E' \supset E$ with the same box dimension $\dim_B E' = \dim_B E$, and with topological codimension 1. Naturally, the boundary of the voids includes $E' - E$, with $\dim_H(E' - E) \leq \dim_B(E' - E) \leq d-1$, so it is negligible with respect to E . We have mentioned in Section 2.1 an example belonging to Cantor-like fractals.

We have explored possible application of our results in physics, in too many areas to be considered in detail. Regarding the application to the cosmic structure, both observations and theory support the existence of voids and, to some extent, scale invariance. In particular, the adhesion model leads to the formation of a self-similar “cosmic foam”, as commented in Section 3. For the moment, it is a moot point whether or not the galaxies are distributed along walls (even though the expression “wall galaxy” is in use, especially in the cosmological literature about voids). As to the scaling of cosmic voids, our analysis of catalogues in Ref. [4] did not show any evidence. However, a recent analysis of observations of the local group of galaxies has found a set of non-degenerate convex-like “minivoids” that satisfy the Zipf law [17].

Acknowledgments

I thank M. Santander and D. Alarcos for a conversation on integral geometry. My work is supported by the “Ramón y Cajal” program and by grant BFM2002-01014 of the Ministerio de Educación y Ciencia.

References

- [1] B.B. Mandelbrot, *The Fractal Geometry of Nature*, W.H. Freeman and Company, NY, 1977.
- [2] J. Feder, *Fractals*, Plenum Press, New York, 1988.
- [3] K. Falconer, *Techniques in Fractal Geometry*, John Wiley & Sons, Chichester, UK, 1997.
- [4] J. Gaite, S.C. Manrubia, *Mon. Not. R. Astron. Soc.* 335 (2002) 977.
- [5] J. Gaite, *Eur. Phys. Jour. B* 47 (2005) 93;
J. Gaite, Scaling of voids in the large scale distribution of matter, in: J.C. del Toro, E.J. Alfaro, J.G. Gorgas, E. Salvador-Solé (Eds.), *The Many Scales of the Universe — JENAM 2004*, *Astrophysics Reviews*, [astro-ph/0510328](https://doi.org/10.1007/978-3-540-25103-2_8).
- [6] J. Einasto, M. Einasto, M. Gramann, *Mon. Not. R. Astron. Soc.* 238 (1989) 155;
G. Kauffmann, A.P. Fairall, *Mon. Not. R. Astron. Soc.* 248 (1991) 313;
H. El-Ad, T. Piran, *Astrophys. J.* 491 (1997) 421;
J. Aikio, P. Mähönen, *Astrophys. J.* 497 (1998) 534.
- [7] F. Sylos Labini, M. Montuori, L. Pietronero, *Phys. Rep.* 293 (1998) 61;
A. Gabrielli, F. Sylos Labini, M. Joyce, L. Pietronero, *Statistical Physics for Cosmic Structures*, Springer, 2005.
- [8] G.K. Zipf, *Human Behavior and the Principle of Least Effort*, Addison-Wesley Press Inc., Massachusetts, 1949.
- [9] D. Sornette, *Critical Phenomena in Natural Sciences*, Springer-Verlag, Berlin, 2000.
- [10] H.G. Eggleston, *Convexity*, Cambridge UP, Cambridge, 1958;
- [11] L.A. Santaló, *Integral Geometry and Geometric Probability*, Addison-Wesley, 1976.
- [12] K.R. Mecke, *Additivity, Convexity, and beyond: Applications of Minkowski functionals in statistical physics*, in: K.R. Mecke, D. Stoyan (Eds.), *Statistical Physics and Spatial Statistics*, in: *Lecture Notes in Physics*, vol. 554, Springer, July 2000, pp. 111–184.
- [13] K. Falconer, *Fractal Geometry*, John Wiley & Sons, Chichester, UK, 2003.
- [14] S.V. Anishchik, N.N. Medvedev, *Phys. Rev. Lett.* 75 (1995) 4314.
- [15] S.W. Tyler, S.W. Wheatcraft, *Soil Sci. Soc. Am. J.* 56 (1992) 362.
- [16] S.N. Gurbatov, A.I. Saichev, S. Shandarin, *Mon. Not. R. Astron. Soc.* 236 (1989) 385.
- [17] R. Van de Weygaert, in: M. Plionis, S. Cotsakis (Eds.), *Modern Theoretical and Observational Cosmology*, vol. 276, Kluwer, Dordrecht, 2002, p. 119.
- [18] A.V. Tikhonov, I.D. Karachentsev, *Minivoids in the local volume*, *Special Astrophysics Observatory, Astrophys. J.* (2006) (in press) (preprint March 2006).