

# EXPECTATIONS ON FRACTAL SETS

DAVID H. BAILEY, JONATHAN M. BORWEIN, RICHARD E. CRANDALL,  
AND MICHAEL G. ROSE

ABSTRACT. Using fractal self-similarity and functional-expectation relations, the classical theory of box integrals—being expectations on unit hypercubes—is extended to a class of fractal “string-generated Cantor sets” (SCSs) embedded in unit hypercubes of arbitrary dimension. Motivated by laboratory studies on the distribution of brain synapses, these SCSs were designed for dimensional freedom—a suitable choice of generating string allows for fine-tuning the fractal dimension of the corresponding set. We also establish closed forms for certain statistical moments on SCSs, develop a precision algorithm for high embedding dimensions, and report various numerical results. The underlying numerical quadrature issues are in themselves quite challenging.

## 1. INTRODUCTION

The development of the following mathematics was motivated by recent laboratory data concerning the spatial locations of  $10^6$  mouse-brain synapses naturally embedded in three dimensions. The synapses were found to be distributed in a fractal manner, with fractal box dimension strictly less than the limiting value of 3 that would be expected from a set of points randomly distributed throughout a cuboid (see [11] for full details and [15] for a recent popular account).<sup>1</sup>

Another way that laboratory distributions loom non-random is that statistical moments, such as expected pairwise distances, do not follow the statistics of random distributions. Indeed, the aforementioned brain data possesses certain expectations from non-random distributions. To understand such phenomena further, we develop herein various statistical measures—in particular, separation moments (box integrals)—on a particular class of

---

*Date:* September 4, 2012.

*2010 Mathematics Subject Classification.* Primary 28A80 Secondary 11-04, 65-04.

*Key words and phrases.* Expectations, Fractals, self-similarity, numerical quadrature, Monte Carlo methods.

Supported in part by the Director, Office of Computational and Technology Research, Division of Mathematical, Information, and Computational Sciences of the U.S. Department of Energy, under contract number DE-AC02-05CH11231.

<sup>1</sup>The brain-synapse work to which we refer uses cuboids, being right-parallelepipeds of sides  $(a, b, c)$ , for which box integrals are taken over the volume  $abc$ . In the present treatment, we restrict our study to the  $n$ -cube  $[0, 1]^n$  and fractal sets embedded in such cubes.

abstract fractal sets, in the hope that these measures can assist the identification of empirical synapse distributions.

Box integrals were first introduced in 1976 as a means of analysing the expected norm on points uniformly distributed through unit hypercubes [1]. Following intensive study over the last decade, a number of closed-form results pertaining to box integrals have been developed; see for example [2], [3] and [9]. The canonical definitions are as follows [3].

**Definition 1.1.** Given dimension  $n$ , complex parameter  $s$  and a fixed point  $q$  in the unit  $n$ -cube, the *box integral*  $X_n(s, q)$  is defined as the expectation of a certain norm  $|r - q|^s$ , with  $q$  fixed and  $r$  chosen at random from a uniform distribution over the unit  $n$ -cube. That is,

$$(1.1) \quad \begin{aligned} X_n(s, q) &:= \langle |r - q|^s \rangle_{r \in [0,1]^n} \\ &= \int_{r \in [0,1]^n} |r - q|^s \mathcal{D}r \end{aligned}$$

where  $\mathcal{D}r := dr_1 \dots dr_n$  is the  $n$ -space volume element.

**Example 1.2 (Classical Box Integrals).** Three important instances of the  $X$ -integrals follow. First is  $B_n(s)$ , the order- $s$  moment of separation between a random point and a vertex of the  $n$ -cube (such as the origin):

$$(1.2) \quad B_n(s) := X_n(s, 0) = \langle |r|^s \rangle_{r \in [0,1]^n} = \int_{r \in [0,1]^n} |r|^s \mathcal{D}r;$$

second is  $\Gamma_n(s)$ , the order- $s$  moment of separation between a random point and the centroid  $\mathbf{1}/2 = (1/2, 1/2, \dots, 1/2)$  of the  $n$ -cube:

$$(1.3) \quad \Gamma_n(s) := X_n(s, \mathbf{1}/2) = \langle |r - \mathbf{1}/2|^s \rangle_{r \in [0,1]^n} = \int_{r \in [0,1]^n} |r - \mathbf{1}/2|^s \mathcal{D}r;$$

and finally  $\Delta_n(s)$ , the order- $s$  moment of separation between *two* random points in the  $n$ -cube:

$$(1.4) \quad \Delta_n(s) := \langle X_n(s, q) \rangle_{q \in [0,1]^n} = \langle |r - q|^s \rangle_{r, q \in [0,1]^n} = \int_{r, q \in [0,1]^n} |r - q|^s \mathcal{D}r \mathcal{D}q.$$

It is these three expectations which we aim to generalize, with particular emphasis on  $\Delta_n(s)$  as the measure most relevant to the analysis of empirical synapse distributions.  $\diamond$

The existence of well-defined analytic continuations for  $B_n(s)$  and  $\Delta_n(s)$  over the complex  $s$ -plane was established in [2] and [3]. In particular,  $B_n(s)$  was found to have the following absolutely convergent analytic series:

$$(1.5) \quad B_n(s) = \frac{n^{1+s/2}}{s+n} \sum_{k=0}^{\infty} \gamma_{n-1,k} \left(\frac{2}{n}\right)^k$$

where the  $\gamma_{m,k}$  are fixed real coefficients defined by the following two-variable recursion [12]:

$$(1.6) \quad (1 + 2k/m)\gamma_{m,k} = (k - 1 - s/2)\gamma_{m,k-1} + \gamma_{m-1,k}$$

for  $m, k \geq 1$ ; this recursion being ignited by  $\gamma_{0,k} := \delta_{0,k}$ ,  $\gamma_{m,0} := 1$ . Of particular interest to the present work is the fact that the analytic series (1.5) exhibits a pole at  $s = -n$ , the negated dimension of the sample space (cf Theorem 8.1).

**1.1. The goal of this paper.** Our goal to establish a solid theoretical foundation and closed form results for box integrals over a restricted class of fractal sets; namely, what we call string-generated Cantor sets (SCSs). These sets are formally introduced in Section 2 and a closed form is provided for the relationship between their fractal dimensions and generating strings. We begin to extend the classical box integrals  $B_n(s)$ ,  $\Gamma_n(s)$  and  $\Delta_n(s)$  over the SCSs in Section 3, using a heuristic notion of expectation. A spectral formalism is brought to bear in Section 4. This with the addition of fractal self-similarity relations developed in Section 5 enables closed forms for the expectations to be obtained. Exact results are obtained for arbitrary moments in embedding-dimension one in Section 6, and for second moments in arbitrary dimensions in Section 7. Self-similarity relations are used to analyse the complex poles of  $B_n(s)$  (as defined over SCSs) in Section 8. We finish with some numerical results and open questions in Section 9. Some numerical data and plots regarding the fractal dimensions and second-order separation moments on various SCSs is given in Appendix A—along with a discussion of the methods employed.

Let us emphasise that in this paper we take very much a physicist’s view of the expectations we will uncover. We leave the considerable work of producing a mathematically fully rigorous accounting (in terms of abstract measure theory, and invariant measures on fractals) for a subsequent paper.

## 2. STRING-GENERATED CANTOR SETS

As a first step towards full generalisation of classical box integrals to arbitrary fractal sets, we introduce the concept of a *string-generated Cantor set* (SCS). Though our definition will perhaps appear arbitrary, or not general enough, our motive was simply to construct a large class of Cantor-like fractals embedded in dimension  $n$  with controllable *fractal box dimension* [4] lying in some convenient interval. The brain-synapse research in [11] showed that experimental fractal dimensions tended to be near  $n = 3$  for the 3-dimensional data, yet *varying* along the penetration axis of the tissue. Thus a fractal model is motivated in which one may “tune” the dimension as required. A similar motivation is apparent in a recent paper on scattering [10] which considers only a somewhat coarser fractal model.

**2.1. Formal Definition of a String-generated Cantor Set.** The complete structure of any SCS is encoded within its generating string. For known embedding dimension  $n$ , let  $P = P_1P_2 \dots P_p$  denote a periodic string of digits with period  $p$ , some positive integer. For instance, with  $n = 1$ ,  $P = 01$  will denote the period-2 string  $\overline{01} = 010101 \dots$  (the over-line being henceforth omitted from our notation for convenience). In the embedding space  $[0, 1]^n$  we will consider only those strings for which  $P_i \leq n$  for all  $i$ . This restriction enables us to define a family of SCSs embedded in the unit  $n$ -cube for a given embedding dimension  $n$ .

Consider the ternary expansion for coordinates of  $x = (x_1, \dots, x_n) \in [0, 1]^n$ , namely:

$$\begin{aligned} x_1 &= 0.x_{11} x_{12} x_{13} \dots \\ x_2 &= 0.x_{21} x_{22} x_{23} \dots \\ &\vdots \\ x_n &= 0.x_{n1} x_{n2} x_{n3} \dots \\ &\quad \uparrow \quad \uparrow \quad \uparrow \\ &\quad c_1 \quad c_2 \quad c_3 \quad \dots \end{aligned}$$

with every digit  $x_{jk} \in \{0, 1, 2\}$  (or, when working with “balanced” ternary vectors,  $x_{jk} \in \{-1, 0, 1\}$ ) and the vectors  $c_k = (x_{1k}, \dots, x_{nk})$  denoting respective columns of digits. Each periodic string defines an SCS by singling out points with admissible ternary expansions (the SCS being the collection of such admissible points). More precisely, in a given generating string  $P$  the value of  $P_k$  determines the maximum number of coordinates of  $x$  that are permitted to have the digit 1 in the  $k$ th (and  $(k+p)$ th,  $(k+2p)$ th,  $\dots$ ) place of the ternary expansion.

For the purpose of enumerating the digits that are restrained by the generating strings, we define the following counting functions. First, the “unit” counter, appropriate to vectors  $c$  having all elements  $\in \{0, 1, 2\}$ :

$$U(c) := \#\{1\text{'s in ternary vector } c\},$$

and for later use with balanced-ternary vectors  $b$  having all elements  $\in \{-1, 0, 1\}$ , the “zero” counter:

$$Z(b) := \#\{0\text{'s in balanced-ternary vector } b\}.$$

(Typically we create a balanced-ternary vector  $b$  from a standard ternary vector  $c$  simply by  $b = c - \mathbf{1}_n$ .)

We are now ready to formally define a string-generated Cantor set.

**Definition 2.1 (String-generated Cantor set).** Fix positive integers  $n$  and  $p$ . Given an embedding space  $[0, 1]^n$  and an entirely-periodic string  $P = P_1P_2 \dots P_p$  of non-negative integers with  $P_i \leq n$  for all  $i = 1, 2, \dots, p$ , the *String-Generated Cantor Set (SCS)*, denoted  $C_n(P)$ , is the set of all

admissible  $x \in [0, 1]^n$ , where

$$(2.1) \quad x \text{ admissible} \iff U(c_k) \leq P_k \quad \forall k \in \mathbb{N}$$

with notational periodicity assumed:  $P_{p+k} := P_k$  for all  $k \geq 1$ .

*Remark 2.2.* We make the following preliminary observations:

- (1)  $C_1(0)$  is the classical “middle-thirds-removed” Cantor set on  $[0, 1]$ , as a point  $x \in C_1(0)$  is defined to be admissible iff its ternary expansion is entirely devoid of 1’s (i.e.  $U(c_k) \leq 0$ ).
- (2)  $C_n(n)$  is the full unit  $n$ -cube  $[0, 1]^n$ , as all points  $x \in [0, 1]^n$  are admissible (every ternary vector is allowed for every column  $c_k$ ).
- (3) For a fixed embedding dimension  $n$ , if a single element  $P_k < n$  of a periodic string  $P$  is increased, the restrictions on admissible points are relaxed and so the resulting SCS contains the original SCS. Every SCS  $C_n(P)$  contains the maximally-restricted SCS,  $C_n(0)$ , and is contained by the full  $n$ -cube,  $C_n(n)$  (which places no restrictions on admissible points).
- (4) The Lebesgue measure of (the always uncountable set)  $C_n(P)$  is zero *unless*  $P = n$ , in which case the Lebesgue measure is  $n$  (since the SCS is exactly the full  $n$ -cube)—this follows from the Borel Zero-One law [5].<sup>2</sup>

Some pictorial examples of string-generated Cantor sets in one, two and three dimensions are shown in Figure (2.1, showing both dependence on the defining string and different styles of visualization. The upper image shows the ‘middle-thirds-removal’ procedure for generating the standard *Cantor ternary set*  $C_1(0)$  of dimension  $\log_3 2$ —the thin, wide rectangle showing the ‘dusty’ set remaining after 6 recursive removals. At middle left is  $C_2(0)$ —sometimes called *Cantor dust*, of dimension  $\log_3 4$ . Middle right shows another 2-dimensional variant,  $C_2(1)$ —sometimes called a *gasket* or the *Sierpinski carpet*, of dimension  $\log_3 8$ . Lower-left shows  $C_3(0)$ , also of dimension  $\log_3 8$  (image due to A. Baserinia (2006)) and lower-right shows  $C_3(1)$ , of dimension  $\log_3 20$  (image due to R. Dickau (2008)).  $\diamond$

---

<sup>2</sup>A brief outline of the argument, which generalises to higher dimensions  $n$ , is as follows. A point chosen at random from the set  $C_1(0)$  is admissible only if the digit 1 is avoided in every position of the ternary expansion. The probability that this occurs at each appropriate digit in the expansion is  $2/3$ , and

$$\sum_{n=1}^{\infty} \left(\frac{2}{3}\right)^n = 2 < \infty$$

so the Borel Zero-One law implies a zero probability of admissibility—and so Lebesgue measure 0 by definition (over strings). Uncountability follows immediately upon mapping digits  $0 \rightarrow 0, 2 \rightarrow 1$ , so the cardinality of  $C_1(0)$  is that of  $[0, 1]$ .

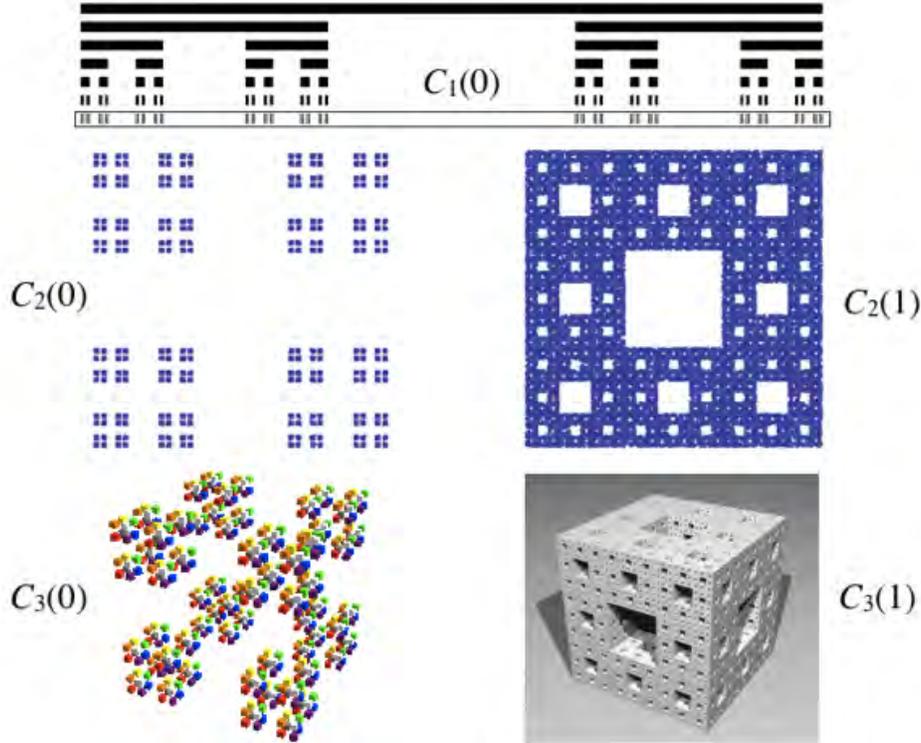


FIGURE 1. Images of various string-generated Cantor sets (SCSs).

**2.2. Fractal dimension of an SCS.** The periodic string formulation in Definition 2.1 immediately gives rise to a closed form for the fractal dimension of an SCS.<sup>3</sup> The proof of this closed form result utilises the machinery of iterated function systems, which we briefly introduce below [14].

**Definition 2.3 (Iterated function system).** A mapping  $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a *contraction* if there exists a *contraction factor*  $0 < c < 1$  such that  $|S(x) - S(y)| \leq c|x - y|$  for all  $x, y \in \mathbb{R}^n$ . If equality holds for all  $x$  and  $y$ , the mapping is said to be a *similarity*. An *iterated function system (IFS)* is a finite family of contractions  $\{S_1, S_2, \dots, S_m\}$  with  $m \geq 2$ . Every IFS has a unique *attractor* - a non-empty compact subset  $A \subset \mathbb{R}^n$  such that

$$A = \bigcup_{i=1}^m S_i(A).$$

<sup>3</sup>In this paper, the phrase “fractal dimension” will always refer to both the Hausdorff and box-counting dimensions [14, Chapters 2 & 3], which are shown in the proof of Proposition 7 to be identical for any given SCS.

An IFS is said to satisfy the *open set condition* if there exists a non-empty bounded open set  $V \subset \mathbb{R}^n$  such that

$$V \supset \bigcup_{i=1}^m S_i(V).$$

**Theorem 2.4 (Fractal dimension of a self-similar set).** *Suppose that the open set condition holds for the IFS  $\{S_1, S_2, \dots, S_m\}$  on  $\mathbb{R}^n$  (with associated contraction factors  $\{c_1, c_2, \dots, c_m\}$ ). Then the Hausdorff dimension and box-counting dimension of the IFS attractor are equal and take the value  $\delta$ , where:*

$$(2.2) \quad \sum_{i=1}^m (c_i)^\delta = 1.$$

As a final preliminary observation, note that for a given periodic string  $P$  the set of admissible columns  $c_k$  is enumerated by the formula:

$$(2.3) \quad N_k(P, n) := N_k = \#\{\text{admissible columns } c_k\} = \sum_{j=0}^{P_k} \binom{n}{j} 2^{n-j},$$

which follows directly from the observation that  $j \leq P_k$  coordinates may attain the value 1, leaving  $n - j$  positions each able to attain the value 0 or 2. Note also that the minimum fractal dimension (corresponding to the sparsest SCS) has  $N_k = 2^n$ , while the maximum dimension (corresponding to the unit  $n$ -cube) has  $N_k = 3^n$ .

**Proposition 2.5 (Fractal dimension—closed form).** *The fractal dimension (in both the Hausdorff and box-counting sense)  $\delta(C_n(P))$  of the SCS  $C_n(P)$  is given by the closed form*

$$(2.4) \quad \delta(C_n(P)) = \frac{\log \prod_{k=1}^p N_k(P, n)}{p \log 3}.$$

*Proof.* The set  $C_n(P)$  is the union of finitely many copies of itself scaled by a factor of  $3^{-p}$  (cf Figure 2.1). Overlay the unit  $n$ -cube with  $3^p$  hypercubes of side length  $3^{-p}$  and consider the similarities  $S_i$  that map the unit  $n$ -cube into these hypercube subsets. Then all such similarities have contraction factor  $3^{-p}$ . The IFS with attractor  $C_n(P)$  consists of those ‘admissible’ similarities  $S_i$  which map the unit  $n$ -cube into a hypercube subset that intersects  $C_n(P)$ . To enumerate these we use our column-counting formula (2.3).

Consider the fractal approximation to  $C_n(P)$  obtained by truncation of the periodic string  $P$  to its first  $p$  digits (ie. its first complete period). This set is identical to  $C_n(P)$  on scales greater than  $3^{-p}$ , but contains no fine structure below this limit. Equivalently, the equivalence classes contain those points in the unit  $n$ -cube whose coordinate ternary expansions are equal up to and including the  $p$ -th digit. Each equivalence class (represented by a coordinate ternary expansion consisting of all 0’s after the  $p$ -th place) can be mapped one-to-one onto its own  $3^{-p}$ -scaled hypercube.

The number of such hypercubes that intersect  $C_n(P)$  is therefore equal to the number of admissible equivalence classes. Each admissible equivalence class corresponds to a ordered concatenation of admissible  $c_k$  columns for  $k = 1, \dots, p$ .

For the  $k$ th ternary place there are  $N_k$  admissible columns  $c_k$ , so the number of admissible similarities is enumerated by:

$$m = \prod_{k=1}^p N_k.$$

Assign each admissible similarity an unique index  $i$  ranging from 1 to  $m = \prod_{k=1}^p N_k$ . Then, the self-similar structure of  $C_n(P)$  is encapsulated by

$$C_n(P) = \bigcup_{i=1}^{\prod N_k} S_i(C_n(P)).$$

The similarities  $S_i$  satisfy the open set condition with  $V$  the interior of the unit  $n$ -cube. It remains to verify the summation in (2.2) for  $\delta$ :

$$\sum_{i=1}^m (c_i)^\delta = \sum_{i=1}^{\prod N_k} (3^{-p})^{\frac{\log \prod N_k}{p \log 3}} = \prod N_k \left( 3^{-\log_3 \prod N_k} \right) = 1.$$

Having satisfied all conditions of Theorem 2.4, the result immediately follows.  $\square$

**Example 2.6 (Fractal dimensions).** Note that the fractal dimension  $\delta(C_n(P))$  of the SCS  $C_n(P)$  depends only on the embedding dimension  $n$  and the generating string  $P = (P_1 \dots P_p)$ , as exemplified by the following:

- (1) The dimension of the full unit  $n$ -cube is  $\delta(C_n(n)) = n$ , as expected.
- (2) The dimension of the classical Cantor set is the celebrated result

$$\delta(C_1(0)) = \log_3(2).$$

- (3) In  $n = 6$  embedding dimensions, the Cantor set  $C_6(1012)$  has dimension

$$\delta(C_6(1012)) = \frac{26 \log 2 + \log 31}{4 \log 3} \approx 4.8848.$$

- (4) We can derive arbitrary-dimension formulae for a given string  $P$ . For example, when  $P = 2$  (and  $n \geq 2$ ):

$$\delta(C_n(2)) = \log_3(2^{n-3}(n^2 + 3n + 8)).$$

$\diamond$

*Remark 2.7.* Note also that the fractal dimension  $\delta(C_n(P))$  of the SCS  $C_n(P)$  depends only on the elements present in the generating string, and *not* the order in which they occur. This has a nice consequence for modelling applications—permuting the generating string of an SCS allows for

fine-tuning of the statistical measures without altering the fractal dimension of the model.  $\diamond$

**2.3. Density theorem for SCS fractal dimensions.** The range of available fractal dimensions in any given class of SCSs is quantified by the following proposition, which is a straightforward deduction from the explicit formula (2.4). Recall that the minimum fractal dimension for a given embedding space has  $N_k = 2^n$ , while the maximum dimension has  $N_k = 3^n$ .

**Proposition 2.8 (Density).** *For given embedding dimension  $n$ , the set of possible fractal dimensions  $\delta(C_n(P))$  is dense on the interval  $[n \log_3 2, n]$ .*

Proposition 2.8 ensures that for 3-dimensional embedding we can always select some SCS with fractal dimension dense the interval  $[1.90, 3.00]$ . The motivating research in [11] experimentally requires this sort of span—certainly in  $[2.4, 3.0]$ —as brain-synapse dimensions vary importantly across neural layers. Indeed, the simple form  $\delta(C_n(P)) = (\log_3 \prod_{k=1}^p N_k)/p$  has various computational applications, such as “matching” an SCS to an actual laboratory distribution.

**2.4. Expectations for the full cube.** In Tables 1 and 2 we show various results for expectations over the *full* cube in two and three dimensions as established in [2]. The evaluations are hypergeometric in stark contrast with our later rationality results such as Corollary 8.7 of Section 8.4.

### 3. STATISTICAL DEFINITIONS

To extend the classical box integrals of Section 1 to string-generated Cantor sets as in Section 2, we proceed from their interpretation as expectations over the unit  $n$ -cube. We will later connect the generalised integrals  $B(s, C_n(P))$ ,  $\Gamma(s, C_n(P))$  and  $\Delta(s, C_n(P))$  to explicit integrals over the unit  $n$ -cube; for now, we denote certain expectations over an arbitrary SCS  $C_n(P)$  by the following formal assignments (after all, we have not yet defined *expectations* over an SCS): for  $\Re(s) > 0$

$$(3.1) \quad B(s, C_n(P)) := \langle |r|^s \rangle_{r \in C_n(P)},$$

$$(3.2) \quad \Gamma(s, C_n(P)) := \langle |r - \mathbf{1}/2|^s \rangle_{r \in C_n(P)},$$

$$(3.3) \quad \Delta(s, C_n(P)) := \langle |r - q|^s \rangle_{r, q \in C_n(P)}.$$

A pictorial of the basic notion of expectation is Figure 3.

*Remark 3.1.* Note that we have suppressed the subscript  $n$  on  $B$ ,  $\Gamma$  and  $\Delta$ , since the embedding dimension  $n$  is now implicit in the choice of  $C_n(P)$ . Although expectations on a fractal SCS have yet to be defined, we will require from our definition that

$$B(s, C_n(n)) = B_n(s), \quad \Gamma(s, C_n(n)) = \Gamma_n(s) \quad \text{and} \quad \Delta(s, C_n(n)) = \Delta_n(s)$$

since  $C_n(n)$  is always the full  $n$ -cube  $[0, 1]^n$ .  $\diamond$

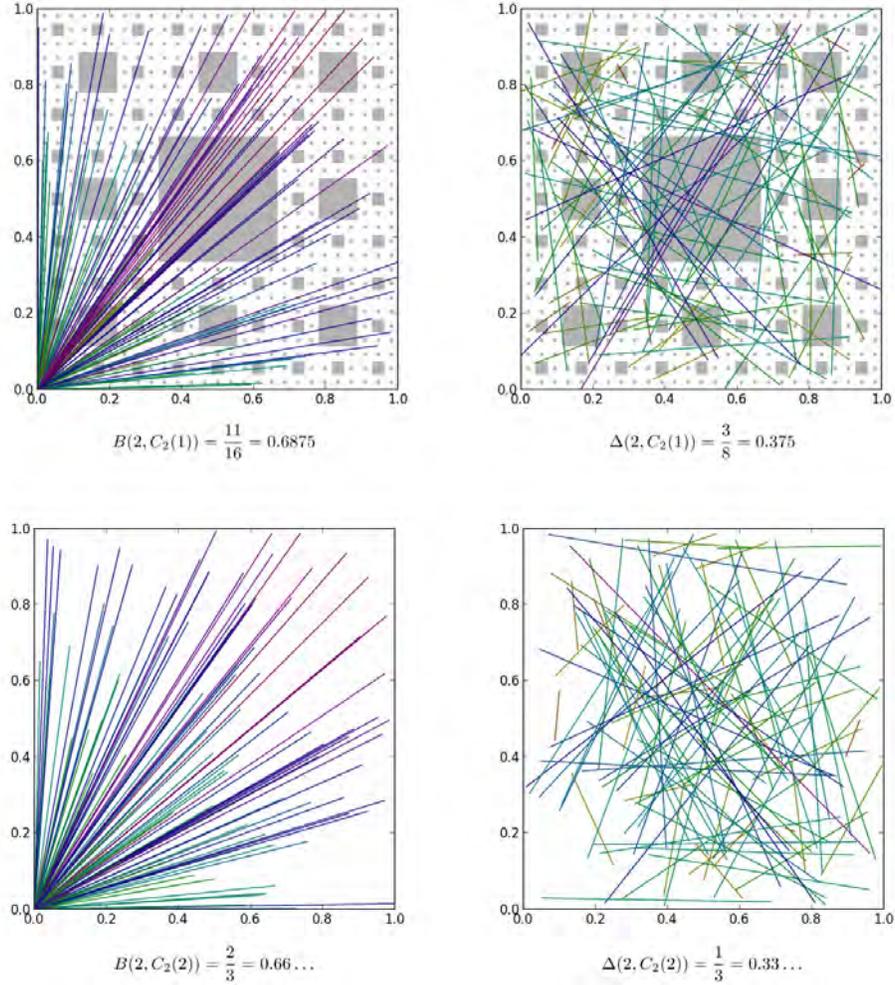


FIGURE 2. The fractal-box integral  $B(2, C_2(1))$  is (top-left) the average squared distance of a carpet point from the origin. The number  $\Delta(2, C_1(1))$  is the expected squared separation (top-right) between two carpet points. Over the classical unit square, the corresponding quantities are  $B(2, C_2(2))$  (bottom-left) and  $\Delta(2, C_1(2))$  (bottom-right). As distance increases, the colour is shifted further towards the violet end of the visible spectrum. The figure gives exact theoretical values for these  $B, \Delta$  cases. Numerical verification of such results is nontrivial (see Section 9 and Appendix A).

$n$	$s$	$B(s, C_2(2))$
2	-4	$-\frac{1}{4} - \frac{\pi}{8}$
2	-3	$-\sqrt{2}$
2	-2	$\infty$
2	-1	$2 \log(1 + \sqrt{2})$
2	1	$\frac{1}{3}\sqrt{2} + \frac{1}{3} \log(1 + \sqrt{2})$
2	3	$\frac{7}{20}\sqrt{2} + \frac{3}{20} \log(1 + \sqrt{2})$
$n$	$s$	$B(s, C_3(3))$
3	-5	$-\frac{1}{6}\sqrt{3} - \frac{1}{12}\pi$
3	-4	$-\frac{3}{2}\sqrt{2} \arctan \frac{1}{\sqrt{2}}$
3	-3	$\infty$
3	-2	$-3G + \frac{3}{2}\pi \log(1 + \sqrt{2}) + 3 \operatorname{Ti}_2(3 - 2\sqrt{2})$
3	-1	$-\frac{1}{4}\pi + \frac{3}{2} \log(2 + \sqrt{3})$
3	1	$\frac{1}{4}\sqrt{3} - \frac{1}{24}\pi + \frac{1}{2} \log(2 + \sqrt{3})$
3	3	$\frac{2}{5}\sqrt{3} - \frac{1}{60}\pi + \frac{7}{20} \log(2 + \sqrt{3})$

TABLE 1. Evaluations for  $B_{2,3}(s)$ . For  $C_n(n)$  all integer values for  $1 \leq n \leq 5$  have closed forms.  $\operatorname{Ti}_2$  is a generalized tangent (polylog) value and  $G$  is Catalan's constant.

**3.1. Precise definition of expectation.** For a complex-valued function  $F : \mathcal{R}^n \rightarrow \mathcal{C}$  we consider the evaluations of  $F$  at every admissible point in an SCS and adopt the definition:

**Definition 3.2 (Expectation of an SCS).** The expectation of  $F : \mathcal{R}^n \rightarrow \mathcal{C}$  on an SCS  $C_n(P)$  is defined by

$$\langle F(r) \rangle_{r \in C_n(P)} := \lim_{j \rightarrow \infty} \frac{1}{N_1 \cdots N_j} \sum_{U(c_i) \leq P_i} F(c_1/3 + c_2/3^2 + \cdots + c_j/3^j),$$

$$\langle F(r - q) \rangle_{r, q \in C_n(P)} := \lim_{j \rightarrow \infty} \frac{1}{N_1^2 \cdots N_j^2} \sum_{U(c_i), U(d_i) \leq P_i} F((c_1 - d_1)/3 + \cdots + (c_j - d_j)/3^j),$$

when the respective limits exist.

For functions such as  $F(r) := |r|^s$ , such limits exist, uniformly on compact sets with  $\Re(s) \geq 0$  and so are analytic.<sup>4</sup> In Section 4 we determine a

<sup>4</sup>Analytic continuation beyond this half  $s$ -plane is possible. We allow “primed sums, excluding  $F$  arguments of the 0 vector on the right. Or—for the first part of the

$n$	$s$	$\Delta(s, C_2(2))$
2	-5	$\frac{4}{3} + \frac{8}{9}\sqrt{2}$
2	-2,-3,-4	$\infty$
2	-1	$\frac{4}{3} - \frac{4}{3}\sqrt{2} + 4 \log(1 + \sqrt{2})$
2	1	$\frac{2}{15} + \frac{1}{15}\sqrt{2} + \frac{1}{3} \log(1 + \sqrt{2})$
$n$	$s$	$\Delta(s, C_3(3))$
3	-7	$\frac{4}{5} - \frac{16\sqrt{2}}{15} + \frac{2\sqrt{3}}{5} + \frac{\pi}{15}$
3	-3,-4,-5,-6	$\infty$
3	-2	$2\pi - 12 G + 12 \text{Ti}_2(3 - 2\sqrt{2}) + 6\pi \log(1 + \sqrt{2})$ $+ 2 \log 2 - \frac{5}{2} \log 3 - 8\sqrt{2} \arctan\left(\frac{1}{\sqrt{2}}\right)$
3	-1	$\frac{2}{5} - \frac{2}{3}\pi + \frac{2}{5}\sqrt{2} - \frac{4}{5}\sqrt{3} + 2 \log(1 + \sqrt{2})$ $+ 12 \log\left(\frac{1+\sqrt{3}}{\sqrt{2}}\right) - 4 \log(2 + \sqrt{3})$
3	1	$-\frac{118}{21} - \frac{2}{3}\pi + \frac{34}{21}\sqrt{2} - \frac{4}{7}\sqrt{3}$ $+ 2 \log(1 + \sqrt{2}) + 8 \log\left(\frac{1+\sqrt{3}}{\sqrt{2}}\right)$
3	3	$-\frac{1}{105} - \frac{2}{105}\pi + \frac{73}{840}\sqrt{2} + \frac{1}{35}\sqrt{3}$ $+ \frac{3}{56} \log(1 + \sqrt{2}) + \frac{13}{35} \log\left(\frac{1+\sqrt{3}}{\sqrt{2}}\right)$

TABLE 2. Evaluations for  $\Delta_{2,3}(s)$ . For  $C_n(n)$  all integer values for  $1 \leq n \leq 5$  have closed forms.

probability measure such that, at least formally,

$$(3.4) \quad \langle F(r) \rangle_{r \in C_n(P)} = \int_{r \in [0,1]^n} F(r) \phi(r) \mathcal{D}r,$$

where  $\phi$  is a density that vanishes on inadmissible  $r \in [0, 1]^n \setminus C_n(P)$ .<sup>5</sup>

#### 4. SPECTRAL FORMALISM OF THE SCS RANDOM WALK

**4.1. Random-walk approach.** Let us view the vector  $r$  (of Definition 2.1) as a random walk whose  $N$ -th step has position

$$r^{(N)} = c_1/3 + c_2/3^2 + \cdots + c_N/3^N,$$

definition—to replace  $c_1/3 + c_2/3^2 + \cdots + c_j/3^j \rightarrow c_1/3 + c_2/3^2 + \cdots + (c_j + \mathbf{1}/2)/3^j$ , thereby offsetting the argument vector to coincide with the centroid of a given box.

<sup>5</sup>Such a measure-based definition requires  $\phi$  to be pathological, as in mathematical physics. Even for the classical Cantor set  $C_1(0)$ , such a  $\phi$ —a Cantor function derivative—must take infinite values on a fractal of measure 0, and so corresponds to a singular measure.

with  $U(c_k) \leq P_k$ .<sup>6</sup> Then the *transition probability density* for displacement  $u := r^{(k)} - r^{(k-1)}$  can be formally cast as

$$\tau_k(u) = \frac{1}{N_k} \sum_{U(c_k) \leq P_k} \delta^n(u - c_k/3^k).$$

The physical analogy is simple: the random walk jumps to its  $k$ -th position according to a comb of  $n$ -dimensional delta functions. This is compatible with the expectation definition of (3.2). Now the density  $\tau_k$  enjoys a Fourier transform

$$T_k(\omega) = \int \tau_k(u) e^{i\omega \cdot u} \mathcal{D}u,$$

where the integral is taken over  $R^n$ .

**4.2. The underlying density.** The convolution principle then establishes—at least formally—that the random walk’s overall density is

$$(4.1) \quad \phi(r) = \frac{1}{(2\pi)^n} \int \prod_{k=1}^{\infty} T_k(\omega) e^{-i\omega \cdot r} \mathcal{D}\omega.$$

This formal result is quite powerful if we carefully denote a *spectral kernel*

$$(4.2) \quad {}_n G_P(\omega) := \prod_{k=1}^{\infty} \frac{1}{N_k} \sum_{Z(b_k) \leq P_k} e^{i\omega \cdot b_k/3^k},$$

where we have transformed to *balanced ternary* vectors  $b_k$ —being vectors of elements  $x_{jk} \in \{-1, 0, 1\}$ , each vector having zero-count  $Z(b_k)$  not exceeding  $P_k$  (see Definition 2.1). Note that  $Z(b_k) = Z(|b_k|)$  where  $|b_k|$  represents the coordinate-wise absolute value. It is useful to observe that

$$(4.3) \quad {}_n G_P(\omega) := \prod_{k=1}^{\infty} \frac{1}{N_k} \sum_{Z(b_k) \leq P_k} \cos(\omega \cdot b_k/3^k),$$

since  $-b_k$  is admissible exactly when  $b_k$  is.

When it is clear from context we will denote  $G_P := {}_n G_P$ . With this notation in mind we have the fundamental density relation

$$(4.4) \quad \phi(r) = \frac{1}{(2\pi)^n} \int e^{-i\omega \cdot r} e^{i\omega \cdot \mathbf{1}/2} G_P(\omega) \mathcal{D}\omega,$$

and a similarly derived representation for the density  $\Phi$  of a two-vector difference  $r - q$ :

$$(4.5) \quad \Phi(d := r - q) = \frac{1}{(2\pi)^n} \int e^{-i\omega \cdot d} |G_P(\omega)|^2 \mathcal{D}\omega,$$

---

<sup>6</sup>Links to classical lattice theory of random walks arise on restricting to finitely many steps.

It is instructive (and as we shall see, lucrative) to observe that our spectral kernel has a simple interpretation as an expectation; that is, application of Definition 3.2 to (4.2) immediately yields

$$G_P(\omega) = \left\langle e^{i\omega \cdot (r-1/2)} \right\rangle_{r \in C_n(P)}.$$

In addition,

$$|G_P(\omega)|^2 = \left\langle e^{i\omega \cdot (r-q)} \right\rangle_{r, q \in C_n(P)}.$$

It is likewise instructive simply to insert such functions as  $f(r) := e^{i\omega \cdot r}$  into the expectation Definition 3.2.

In Section 5, we record cases where the spectral kernel associated with an SCS reduces to a product of simpler kernels. In particular, both  ${}_nG_n$  and  ${}_nG_0$  factor into a  $n$ -fold product of 1-dimensional kernels (see Example 5.3).

**Example 4.1 (Cantor dust).** The 2-dimensional Cantor Dust  $C_2(0)$  has an associated spectral kernel which can be factored as

$${}_2G_0(\omega_1, \omega_2) = {}_1G_0(\omega_1) \cdot {}_1G_0(\omega_2)$$

and so

$$\left\langle e^{i\omega \cdot (r-1/2)} \right\rangle_{r \in C_2(0)} = \left\langle e^{i\omega_1 \cdot (x-1/2)} \right\rangle_{x \in C_1(0)} \left\langle e^{i\omega_2 \cdot (x-1/2)} \right\rangle_{x \in C_1(0)}$$

Similarly, the 3-dimensional Cantor Dust  $C_3(0)$  has an associated spectral kernel

$${}_3G_0(\omega_1, \omega_2, \omega_3) = {}_1G_0(\omega_1) \cdot {}_1G_0(\omega_2) \cdot {}_1G_0(\omega_3)$$

and so the pattern continues.  $\diamond$

## 5. FUNDAMENTAL FRACTAL RELATIONS

Remarkably, many of the exact results we shall derive depend exclusively on self-similarity relations. At the very core of such analysis is the following.

**Proposition 5.1 (Spectral self-similarity).** *The spectral kernel for a given SCS in  $n$ -dimensions with string  $P$  satisfies*

$$G_P(\omega) = S_P(\omega) G_P(\omega/3^p),$$

where  $p$  is the period of the generating string  $P$ , with the similarity factor  $S_P = {}_nS_P$  given by the finite product

$$(5.1) \quad S_P(\omega) := \prod_{k=1}^p \frac{1}{N_k} \sum_{Z(b_k) \leq P_k} e^{i\omega \cdot b_k/3^k}$$

$$(5.2) \quad = \prod_{k=1}^p \frac{1}{N_k} \sum_{Z(b_k) \leq P_k} \cos(\omega \cdot b_k/3^k).$$

It is fairly easy to compute  ${}_nS_P$  symbolically. We look at the one dimensional case in more detail in the following section.

**Example 5.2 (Similarity relations in two and three dimensions).**  
 We have

$$\begin{aligned}
 {}_1S_0(\omega_1) &= \cos\left(\frac{1}{3}\omega_1\right) \\
 {}_2S_0(\omega_1, \omega_2) &= {}_1S_0(\omega_1){}_1S_0(\omega_2) \\
 {}_2S_1(\omega_1, \omega_2) &= \frac{1}{4}({}_1S_0(\omega_1) + {}_1S_0(\omega_2) + {}_2S_0(\omega_1, \omega_2)) \\
 {}_2S_2(\omega_1, \omega_2) &= {}_1S_1(\omega_1){}_1S_1(\omega_2) \\
 {}_3S_0(\omega_1, \omega_2, \omega_3) &= {}_1S_0(\omega_1){}_1S_0(\omega_2){}_1S_0(\omega_3) \\
 {}_3S_1(\omega_1, \omega_2, \omega_3) &= \frac{1}{5}({}_2S_0(\omega_1, \omega_2) + {}_2S_0(\omega_1, \omega_3) + {}_2S_0(\omega_2, \omega_3) + {}_3S_0(\omega_1, \omega_2, \omega_3)) \\
 {}_3S_2(\omega_1, \omega_2, \omega_3) &= \frac{1}{13}({}_1S_0(\omega_1) + {}_1S_0(\omega_2) + {}_1S_0(\omega_3) \\
 &\quad + {}_2S_0(\omega_1, \omega_2) + {}_2S_0(\omega_1, \omega_3) + {}_2S_0(\omega_2, \omega_3) + {}_4S_0(\omega_1, \omega_2, \omega_3)) \\
 {}_3S_3(\omega_1, \omega_2, \omega_3) &= {}_1S_1(\omega_1){}_1S_1(\omega_2){}_1S_1(\omega_3)
 \end{aligned}$$

as illustrative of the structure of all similarity relations for SCSs.  $\diamond$

**Example 5.3 (Similarity relations for period-1 strings in  $n$  dimensions).**

$${}_nS_0(\omega_1, \omega_2, \dots, \omega_n) = \prod_{k=1}^n {}_1S_0(\omega_k) \quad \text{and} \quad {}_nS_n(\omega_1, \omega_2, \dots, \omega_n) = \prod_{k=1}^n {}_1S_1(\omega_k)$$

and there is more to say as we shall see in Section 8.4.  $\diamond$

Such self-similarity triggers a cascade of results attendant on fractal theory. Using the  $S$ -factor to rescale integral representations such as (4.4, 4.5) yields:

**Proposition 5.4 (Scaling relations).** *For  $r, q$  in  $R^n$ , the probability densities, respectively pertaining to the box integrals  $B, \Gamma$  and  $\Delta$ , satisfy the scaling relations:*

$$(5.3) \quad \phi(r) = \frac{3^{pn}}{\prod_{k=1}^p N_k} \sum_{U(c_k) \leq P_k} \phi(3^p(r - c_1/3 - c_2/3^2 - \dots - c_p/3^p))$$

$$(5.4) \quad \phi(d := r - \mathbf{1}/2) = \frac{3^{pn}}{\prod_{k=1}^p N_k} \sum_{Z(b_k) \leq P_k} \phi(3^p(d - b_1/3 - b_2/3^2 - \dots - b_p/3^p))$$

$$(5.5) \quad \Phi(d := r - q) = \frac{3^{pn}}{\prod_{k=1}^p N_k^2} \sum_{Z(b_k), Z(a_k) \leq P_k} \Phi\left(3^p\left(d - \sum_{j=1}^p (b_j - a_j)/3^j\right)\right).$$

Finally, we arrive at a functional relation for general expectations; by substitution of the Proposition (5.4) formulae into expectation integrals such as (3.4) we achieve:

**Proposition 5.5 (Functional equations for expectations).** *For  $r, q$  in  $R^n$ , and appropriate  $F$  we have:*

(5.6)

$$\langle F(r) \rangle_{r \in C_n(P)} = \frac{1}{\prod_{j=1}^p N_j} \sum_{U(c_k) \leq P_k} \langle F(r/3^p + c_1/3 + \cdots + c_p/3^p) \rangle$$

(5.7)

$$\langle F(d := r - \mathbf{1}/2) \rangle_{r \in C_n(P)} = \frac{1}{\prod_{j=1}^p N_j} \sum_{Z(b_k) \leq P_k} \langle F(d/3^p + b_1/3 + \cdots + b_p/3^p) \rangle$$

(5.8)

$$\langle F(d := r - q) \rangle_{r, q \in C_n(P)} = \frac{1}{\prod_{j=1}^p N_j^2} \sum_{Z(b_k), Z(a_k) \leq P_k} \langle F(d/3^p + \sum_{j=1}^p (b_j - a_j)/3^j) \rangle$$

It is typical of such functional equations—at least in one dimension—that any solution is either absolutely continuous or is singular with respect to Lebesgue measure, as described in [8] or [7, Chapter 2].

*Remark 5.6.* We shall see at the end of the next section that Proposition 5.1 allows us to show that  $G_P(\omega)$  is everywhere convergent as a product. That is, the product is finite and vanishes only if one of the  ${}_n S_P(\omega)$  terms does.  $\diamond$

## 6. EXACT ANALYSES IN $n = 1$ DIMENSION

There are but two SCS with period  $p = 1$  in dimension  $n = 1$ ; namely  $C_1(0)$  (the standard middle-thirds Cantor set) and  $C_1(1)$  (the full interval  $[0, 1]$ ).

**6.1. The case of  $C_1(1)$ .** From our results on self-similarity we have, for the SCS  $C_1(1) = [0, 1]$ , that

$${}_1 G_1(\omega) = \prod_{k \geq 1} \frac{1}{3} \left( 1 + e^{i\omega/3^k} + e^{-i\omega/3^k} \right)$$

This expression is reducible by the following lemma in which  $\text{sinc}(x) := \sin(x)/x$ :

**Lemma 6.1.** *For all  $\omega$  in  $R$  we have*

$$(6.1) \quad {}_1 G_1(\omega) = \prod_{k \geq 1} \frac{1}{3} \left( 1 + e^{i\omega/3^k} + e^{-i\omega/3^k} \right) = \text{sinc} \left( \frac{\omega}{2} \right).$$

*Proof.* Note as above, that  $e^{i\omega/3^k} + e^{-i\omega/3^k} = 2 \cos(\omega/3^k)$ . Define the function:

$$Q(x) := x \prod_{k \geq 1} \frac{1}{3} \left( 1 + 2 \cos(2x/3^k) \right).$$

Lemma 6.1 is equivalent, on setting  $x = \omega/2$ , to:  $Q(x) = \sin(x)$ . Now,

$$\begin{aligned} \frac{Q(3x)}{Q(x)} &= \frac{3x \prod_{k \geq 1} \frac{1}{3} \left( 1 + 2 \cos(2x/3^{k-1}) \right)}{x \prod_{k \geq 1} \frac{1}{3} \left( 1 + 2 \cos(2x/3^k) \right)} \\ &= 3 \prod_{k \geq 1} \left( \frac{1 + 2 \cos(2x/3^{k-1})}{1 + 2 \cos(2x/3^k)} \right) \\ &= 3 \left( \frac{1 + 2 \cos(2x)}{1 + 2 \cos(2x/3)} \right) \left( \frac{1 + 2 \cos(2x/3)}{1 + 2 \cos(2x/3^2)} \right) \dots \\ &= 3 \left( \frac{1 + 2 \cos(2x)}{1 + 2 \cos(0)} \right) = 1 + 2 \cos(2x) = \frac{\sin(3x)}{\sin(x)} \end{aligned}$$

and by recursion,

$$\frac{Q(x)}{\sin(x)} = \frac{Q(x/3)}{\sin(x/3)} = \frac{Q(x/9)}{\sin(x/9)} = \dots$$

Therefore,

$$\begin{aligned} \frac{Q(x)}{\sin(x)} &= \lim_{x \rightarrow 0} \frac{Q(x)}{\sin(x)} \\ &= \lim_{x \rightarrow 0} \frac{1}{x} \cdot x \prod_{k \geq 1} \frac{1}{3} \left( 1 + 2 \cos(2x/3^k) \right) \\ &= \lim_{x \rightarrow 0} \prod_{k \geq 1} \left( 1 - \frac{4}{3} \sin^2(x/3^k) \right) = 1, \end{aligned}$$

since the final product is absolutely convergent and is continuous at zero.  $\square$

Returning to the SCS  $C_1(1) = [0, 1]$  with  ${}_1G_1(\omega) = \text{sinc}(\omega/2)$ , we see that this special case (of fractal dimension 1) has probability density

$$\phi(r) = \frac{1}{2\pi} \int e^{-i\omega(r-1/2)} {}_1G_1(\omega) d\omega$$

which is a “square pedestal” situated on  $[0, 1]$ —the transform of *sinc* (which is only conditionally convergent) is of course the characteristic function of the interval.

**Example 6.2 (The easiest classical box integral).** Likewise, continuing to imaginary  $\omega = it$ , we have an expectation for parameter  $t$ :

$$\langle e^{-tr} \rangle_{r \in [0,1]} = \frac{1 - e^{-t}}{t},$$

giving in turn the (trivially known) box integral

$$(6.2) \quad B(s, C_1(1)) = B_1(s) = \frac{1}{s+1}.$$

When  $s \geq 0$  is integer, this follows from the series for  $(1 - e^{-t})/t$ . When  $s$  is non-integer, consider the region  $\Re(s) \in (-1, 0)$ , where

$$\begin{aligned} \langle r^s \rangle &= \frac{1}{\Gamma(-s)} \int_0^\infty t^{-s-1} \langle e^{-rt} \rangle dt \\ &= \frac{1}{\Gamma(-s)} \int_0^\infty t^{-s-2} (1 - e^{-t}) dt = \frac{1}{s+1}. \end{aligned}$$

We may then infer (6.2) by analytic continuation on appealing to analyticity of the box representation as discussed at the end of §3.1.  $\diamond$

The scaling relation for this SCS is also instructive. We have  $N_1 = 3$  admissible ternary digits, so from Proposition 5.4 we have

$$\phi(r) = \phi(3r) + \phi(3r - 1) + \phi(3r - 2),$$

which (almost everywhere) admits the unit pedestal solution. Perhaps most interesting in this simple case is the two-point density  $\Phi(d := r - q)$ , with scaling relation for  $r, q \in [0, 1]$ :

$$\Phi(d) = \frac{1}{3} (3\Phi(3d) + 2\Phi(3d - 1) + 2\Phi(3d + 1) + \Phi(3d + 2) + \Phi(3d - 2)).$$

This is satisfied a.e. by the “unit tent,” with graph a triangle with base  $[-1, 1]$  and apex at  $(0, 1)$  and likewise for the appropriate density for the difference  $d = r - q$ .

**6.2. The case of  $C_1(0)$ .** As one would expect, the above discussion for the degenerate case  $C_1(1) = [0, 1]$  only becomes more complicated as the defining string is modified. For the standard Cantor set  $C_1(0)$  we have the scaling relation

$$\phi(r) = \frac{3}{2}(\phi(3r) + \phi(3r - 2)),$$

which is satisfied a.e. in the sense of distribution theory.<sup>7</sup> However, one may often *integrate* over such pathological densities with impunity. The spectral kernel as defined in (4.2) now becomes:

$$(6.3) \quad {}_1G_0(\omega) = \prod_{k \geq 1} \cos(\omega/3^k),$$

from which we develop generating-function relations, starting with:

$$(6.4) \quad \sum_{m \geq 0} B(m, C_1(0)) \frac{t^m}{m!} = \langle e^{tr} \rangle_{r \in C_1(0)} = e^{t/2} \prod_{k \geq 1} \cosh(t/3^k).$$

<sup>7</sup>Recall this  $\phi$  is nonzero only on the Cantor set—and yet should have measure one, so it belongs to the world of delta-functions and their generalizations (or of singular measures).

Now a simple scaling  $t \rightarrow 3t$  results in another series

$$(6.5) \quad \sum_{k \geq 0} B(k, C_1(0)) \frac{3^k t^k}{k!} = \langle e^{3tr} \rangle_{r \in C_1(0)} = e^{3t/2} \cosh t \prod_{k \geq 1} \cosh(t/3^k) \\ = e^t \cosh t \sum_{m \geq 0} B(m, C_1(0)) \frac{t^m}{m!}.$$

Matching the coefficients of powers of  $t$  for the far left and far right series in (6.5) immediately resolves *all* box integrals  $B(m, C_1(0))$  for nonnegative integers  $m$ . Note importantly that both Theorem 6.3 below and the  $\Delta$ -counterpart theorem following can alternatively be established via self-similarity relations starting with (8.1) below, without direct recourse to series algebra. We thus have:

**Theorem 6.3 (Closed form for certain  $B(s, C_1(0))$ ).** *For any non-negative integer  $m$ , the expectation  $\langle |r|^m \rangle$  for  $r$  in the standard Cantor set  $C_1(0)$  is given exactly by the recursion, ignited by  $B(0, C_1(0)) = 1$ , for  $m \geq 1$ ,*

$$(6.6) \quad B(m, C_1(0)) = \frac{1}{3^m - 1} \sum_{j=0}^{m-1} \binom{m}{j} 2^{m-j-1} B(j, C_1(0)).$$

Corresponding considerations to those used to derive Theorem 6.3 yield an attractive recursion for the even moments:

**Theorem 6.4 (Closed form for even moments  $B(s, C_1(0))$ ).** *For any non-negative even integer  $2m$ , the expectation  $\langle |r|^{2m} \rangle$  for  $r$  in the standard Cantor set  $C_1(0)$  is given exactly by the recursion, ignited by  $B(0, C_1(0)) = 1$ , for  $m \geq 1$ ,*

$$(6.7) \quad B(2m, C_1(0)) = \frac{1}{3^{2m} - 1} \sum_{j=1}^m \binom{2m}{2j} (2^{2j} - 1) B(2(m - j), C_1(0)).$$

Let us illustrate as follows:

**Example 6.5.** The first eight values of the box integral  $B$  over  $C_1(0)$  are thus:

$$\{B(0, C_1(0)), \dots, B(7, C_1(0)), \dots\} = \left\{ 1, \frac{1}{2}, \frac{3}{8}, \frac{5}{16}, \frac{87}{320}, \frac{31}{128}, \frac{10215}{46592}, \frac{2675}{13312}, \dots \right\},$$

and so on. ◇

A very similar argument—this time involving the square of  ${}_1G_0(\omega)$ , yields a companion result for any separation moment  $\Delta(m, C_1(0))$  for  $m = 0, 1, 2, 3, \dots$ :

**Theorem 6.6 (Closed form for certain  $\Delta(s, C_1(0))$ ).** *For any non-negative integer  $m$ , the expectation  $\langle |r - q|^m \rangle$  for  $r, q$  in the standard Cantor set  $C_1(0)$*

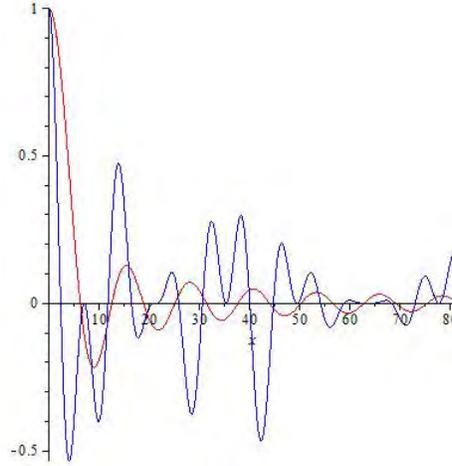


FIGURE 3. Sinc function  ${}_1G_1(\omega)$  (6.1) and more oscillatory  ${}_1G_0(\omega)$  (6.3).

is given exactly by the ignition  $\Delta(0, C_1(0)) := 1$  and the recursion (6.8)

$$\Delta(m, C_1(0)) = \frac{2^m}{2 \cdot 3^m - 2 + (m \bmod 2)} \sum_{k=0}^{\lfloor (m-1)/2 \rfloor} \binom{m}{2k} 4^{-k} \Delta(2k, C_1(0)).$$

Again we illustrate :

**Example 6.7.** The first eight values of the box integral  $\Delta$  over  $C_1(0)$  are thus:

$$\{\Delta(0, C_1(0)), \dots, \Delta(7, C_1(0)), \dots\} = \left\{ 1, \frac{2}{5}, \frac{1}{4}, \frac{19}{106}, \frac{11}{80}, \frac{427}{3880}, \frac{529}{5824}, \frac{139681}{1819168}, \dots \right\},$$

and so on.  $\diamond$

Theorems 6.3 and 6.6 immediately imply the following striking result:

**Corollary 6.8 (Rationality of moments for the Cantor set).** *For any non-negative integer  $s$ , all moments  $B(s, C_1(0))$  and  $\Delta(s, C_1(0))$  are rational.*

The complexity of  ${}_nG_P$  is illustrated in Figure 6.2 just for  $n = 1$ . It is also instructive to compute and plot  ${}_1G_{01}$  and  ${}_1G_{10}$  which share their fractal dimension but not their moments (see Appendix A). Note that

$${}_1S_{01}(\omega_1) = \frac{1}{3} (\cos(2/9 \omega_1) + \cos(3/9 \omega_1) + \cos(4/9 \omega_1)),$$

and

$${}_1S_{10}(\omega_1) = \frac{1}{3} (\cos(1/9 \omega_1) + \cos(2/9 \omega_1) + \cos(4/9 \omega_1)),$$

while

$${}_2S_{20}(\omega_1, \omega_2) = {}_1S_{10}(\omega_1) \cdot {}_1S_{10}(\omega_2).$$

**6.3. Estimation of  ${}_nG_P$ .** We conclude this section with a simple estimate of the size of  $G_P(\omega)$  for arbitrary  $P$  of period length  $p$  and any  $n$ . We first observe that

$$(6.9) \quad {}_nG_P(\omega) = \prod_{m=1}^{\infty} {}_nS_p\left(\frac{\omega}{3^{mp}}\right)$$

where

$$(6.10) \quad {}_nS_P(\omega) = \prod_{k=1}^p \frac{1}{N_k} \sum_{Z(b_k) \leq P_k} \cos\left(\frac{\omega \cdot b_k}{3^k}\right)$$

is as given by (5.2) of Proposition 5.1. Fix  $\omega$  and select  $m_0$  so that  $|\omega/3^{mp}| \leq \pi/2$  for  $m \geq m_0$ . Since (6.10) represents  ${}_nS_P(\omega/3^{mp})$  as a product of weighted arithmetic means of cosine values with domain in  $[-\pi/2, \pi/2]$  each term is larger than  $\cos(\|\omega\|_1/3^{mp+k})$  and no greater than 1 we deduce that

$$(6.11) \quad \begin{aligned} \prod_{m=1}^{m_0-1} \left| {}_nS_p\left(\frac{\omega}{3^{mp}}\right) \right| &\geq |{}_nG_P(\omega)| \geq \prod_{m=1}^{m_0-1} \left| {}_nS_p\left(\frac{\omega}{3^{mp}}\right) \right| \prod_{m=m_0}^{\infty} \prod_{k=1}^p \cos\left(\frac{\|\omega\|_1}{3^{mp+k}}\right) \\ &= \prod_{m=1}^{m_0-1} \left| {}_nS_p\left(\frac{\omega}{3^{mp}}\right) \right| \prod_{i=m_0}^{\infty} \cos\left(\frac{\|\omega\|_1}{3^i}\right). \end{aligned}$$

From (6.11) and Lemma 6.1 we deduce:

**Lemma 6.9.** *For all  $\omega$  in  $R^n$  we have*

$$(6.12) \quad \prod_{m < m_0} \left| {}_nS_p\left(\frac{\omega}{3^{mp}}\right) \right| \geq |{}_nG_P(\omega)| \geq \prod_{m < m_0} \frac{|{}_nS_p\left(\frac{\omega}{3^{mp}}\right)|}{\left| \cos\left(\frac{\|\omega\|_1}{3^m}\right) \right|} \cdot {}_1G_0(\|\omega\|_1).$$

*In particular, the kernel  ${}_nG_P(\omega) = \prod_{m=1}^{\infty} {}_nS_p(\omega/3^{pm})$  is an everywhere convergent product which can vanish only at zeros of  ${}_nS_p$  or of cosine.*

In (6.12) we note that the denominator on the right can only vanish when  ${}_1G_1$  does and so the apparent singularities disappear.

## 7. SECOND MOMENTS FOR A GENERAL SCS

The functional expectation relations of Proposition 5.5 can be used directly to yield all expectations  $B(2, C_n(P))$  as rational numbers depending only on the defining string  $P$  and embedding dimension  $n$ , as follows:

**Theorem 7.1 (Closed forms for  $B(2, C_n(P))$  and  $\Delta(2, C_n(P))$ ).** *For any embedding dimension  $n$  and SCS  $C_n(P)$  the box integral  $B(2, C_n(P))$  is rational, given by the closed form:*

$$(7.1) \quad B(2, C_n(P)) = \frac{n}{4} + \frac{1}{1-9^{-p}} \sum_{k=1}^p \frac{1}{9^k} \frac{\sum_{j=0}^{P_k} \binom{n}{j} 2^{n-j} (n-j)}{\sum_{j=0}^{P_k} \binom{n}{j} 2^{n-j}},$$

and the corresponding box integral  $\Delta(2, C_n(P))$  is also rational, given by:

$$(7.2) \quad \Delta(2, C_n(P)) = 2B(2, C_n(P)) - \frac{n}{2}.$$

*Proof.* For (7.1), take  $F(r) := |r|^2$  in Proposition 5.5, so that for  $d := r - \mathbf{1}/2$  the expectation  $\langle d \cdot d \rangle$  is proportional to a sum of expectations

$$\langle (d/3^p + b_1/3 + \cdots + b_p/3^p)^2 \rangle.$$

Observe that expectations of dot-products  $b_j \cdot b_k$  all vanish (this is how the second-moment problem especially simplifies), and one is left with a simple if tedious combinatorial end-argument for this sum of  $b_k$ -dependent expectations.

For (7.2), we argue as follows. We have, simply, for vectors  $r, q$  each ranging over the given SCS,

$$\begin{aligned} \Delta(2, C_n(P)) &= \langle |r - q|^2 \rangle \\ &= \langle r^2 \rangle + \langle q^2 \rangle - 2 \langle r \cdot q \rangle \\ &= 2B(2, C_n(P)) - 2 \langle r \cdot q \rangle. \end{aligned}$$

Now there is a trick to evaluate the dot-product expectation, namely write

$$(r - \mathbf{1}/2) \cdot (q - \mathbf{1}/2) = 0,$$

by symmetry of any SCS around the centroid vector  $\mathbf{1}/2$ . Expanding out this vanishing expectation, we find  $\langle r \cdot q \rangle = n/4$ , which proves the  $\Delta$  relation of the theorem.  $\square$

**Example 7.2 (Values of  $B$  and  $\Delta$ ).** The first few cases of Theorem 7.1 for period-1 strings  $P$  are:

$$\begin{aligned} B(2, C_n(0)) &= \frac{3}{8}n, & B(2, C_n(1)) &= \frac{n(3n+5)}{8n+16}, \\ B(2, C_n(2)) &= \frac{n(3n^2+7n+22)}{8n^2+24n+64}, & B(2, C_n(n-1)) &= \frac{n}{4} \left( 1 + \frac{3^{n-1}}{3^n-1} \right), \end{aligned}$$

together with:

$$\begin{aligned} \Delta(2, C_n(0)) &= \frac{1}{4}n, & \Delta(2, C_n(1)) &= \frac{n(n+1)}{4n+8}, \\ \Delta(2, C_n(2)) &= \frac{n(n^2+n+6)}{4n^2+12n+32}, \\ \Delta(2, C_n(n-1)) &= \frac{n}{6} \left( \frac{3^n}{3^n-1} \right), \end{aligned}$$

Note for comparison that the classical box integrals over the unit  $n$ -cube are:

$$(7.3) \quad B_n(2) = \frac{n}{3} \quad \text{and} \quad \Delta_n(2) = \frac{n}{6}$$

which matches the output of Theorem 7.1 for  $P = n$ .  $\diamond$

Exact rational values for various strings in dimensions 1, 2 and 3 are shown in Appendix A. Computational data derived from Proposition and Theorem 7.1 concerning the relationship between fractal dimension and the box integrals  $B(2, C_n(P))$  and  $\Delta(2, C_n(P))$  is shown in tabulated values and scatterplots in the appendix. These scatterplots show a great deal of fractal structure in and of themselves.

*Remark 7.3.* An example of ordering of box-integral values is the amusing pair:

$$B_3(2) = 1, \text{ but } B(2, C_3(2)) = \frac{105}{104}.$$

This kind of observation provokes thoughts of a monotonicity-ordering principle. One might hypothesise that any two SCSs  $C_n(P)$  and  $C_n(P')$  will satisfy

$$\delta(C_n(P)) \geq \delta(C_n(P')) \stackrel{?}{\implies} B(s, C_n(P)) \leq B(s, C_n(P')),$$

and similarly for  $\Delta(s, C_n(P))$ . However, our computational experiments show that no such ordering principle exists. For instance, in embedding dimension  $n = 1$ ,

$$\begin{aligned} \delta(C_1(100)) &= 0.75\dots \\ \delta(C_1(01)) &= 0.81\dots, \end{aligned}$$

while

$$\begin{aligned} B(C_1(100)) &= \frac{123}{364} = 0.33\dots \\ B(C_1(01)) &= \frac{89}{240} = 0.37\dots \end{aligned}$$

Appendix A holds more computational data for second separation moments of SCSs.  $\diamond$

## 8. SELF-SIMILARITY AND ANALYTICITY

What can be said about integrals  $B(s, C_n(P))$  in general? Certainly the relative ease of analysis for a simple string  $P$  or moment  $s = 2$  will not carry over generally. For one thing, the classical box integrals  $B_n(s)$ —which are as our fractal cases for  $C_n(n) = [0, 1]^n$ —are currently unknown in closed form past  $n = 5$ , see [9]. Nonetheless, one may still exploit self-similarity for general  $s$ .

For example, it follows from Proposition 5.5 that for the standard middle-thirds Cantor set:

(8.1)

$$B(s, C_1(0)) := \langle |r^s| \rangle_{r \in C_1(0)} = \frac{1}{2} \left\langle \left( \frac{r}{3} \right)^s \right\rangle + \frac{1}{2} \left\langle \left( \frac{r+2}{3} \right)^s \right\rangle,$$

(8.2)

$$\Delta(s, C_1(0)) := \langle |d := r - q|^s \rangle = \frac{1}{2} \frac{1}{3^s} \langle |d|^s \rangle + \frac{1}{4} \frac{1}{3^s} \langle (2+d)^s + (2-d)^s \rangle.$$

These are powerful self-similarity relations, yielding both theoretical knowledge and exact expressions for certain values of  $s$ . An immediate result emerges from an equivalent form of (8.1), using the fact that the first expectation, that of  $(r/3)^s$ , is a scaled expectation itself. This leads to

$$(8.3) \quad B(s, C_1(0)) = \frac{1}{2 \cdot 3^s - 1} \langle (r+2)^s \rangle.$$

**8.1. Pole theorem.** This last expression (8.3) reveals a pole in the  $s$ -plane, namely at  $s = -\log_3 2$ . It is not a coincidence that the pole location is the negated fractal dimension. Indeed, self-similarity implies a general result:

**Theorem 8.1** (Poles of  $B(s, C_n(P))$ ). *For any embedding dimension  $n$  and any SCS  $C_n(P)$ , the (analytically continued) box integral  $B(s, C_n(P))$  has a pole at*

$$(8.4) \quad s = -\delta(C_n(P)).$$

*Proof.* The functional relations of Proposition 5.5 let us write

$$\begin{aligned} B(s, C_n(P)) &:= \langle |r|^s \rangle_{r \in C_n(P)} = \frac{1}{\prod_{j=1}^p N_j} \sum_{U(c_k) \leq P_k} \langle |r/3^p + c_1/3 + \dots + c_p/3^p|^s \rangle \\ &= \frac{1}{\prod_{j=1}^p N_j} 3^{-ps} B(s, C_n(P)) + \frac{1}{\prod_{j=1}^p N_j} \sum'_{U(c_k) \leq P_k} \langle |r/3^p + c_1/3 + \dots + c_p/3^p|^s \rangle, \end{aligned}$$

where the last, primed sum indicates—importantly—that not all  $c_k$  can be 0 in the sum. But said primed sum is always finite for any complex  $s$ , as it is a finite sum of expectations of  $3^{ps} |r + b|^s$  for nonzero vectors  $b$ .

Now on regrouping we have

(8.5)

$$\left( 1 - \frac{3^{-ps}}{\prod_j N_j} \right) B(s, C_n(P)) = \frac{1}{\prod_{j=1}^p N_j} \sum'_{U(c_k) \leq P_k} \langle |r/3^p + c_1/3 + \dots + c_p/3^p|^s \rangle,$$

Now the companion factor to  $B$  in (8.5), namely  $(1 - 3^{-ps} / \prod_j N_j)$  vanishes at the fractal dimension  $s = -\delta(C_n)$  (given in Proposition 7) while the right side remains bounded away from zero. It follows that  $B$  must have a pole at such  $s$ .  $\square$

Note the attractive corollary that for the full  $n$ -cube  $[0, 1]^n$ , the pole is at  $s = -n$ , which is entirely consistent with the classical theory (as in Equation (1.5)).

*Remark 8.2.* We have been unable to determine pole structure for integrals  $\Delta(s, C_n)$ , with arbitrary  $n$ . For classical box integrals  $\Delta_n(s)$  always has  $(n + 1)$  complex poles, unlike  $B_n(s)$  which only has one. And yet, we have seen no evidence of multiple poles for any SCS of fractal dimension less than the embedding dimension.  $\diamond$

**8.2. Fractal-box series.** Relation (8.3) also allows us to expand the expectation bracket to obtain an analytic series in  $s$  for the Cantor case  $C := C_1(0)$ :

$$(8.6) \quad B(s, C) = \frac{2^s}{2 \cdot 3^s - 1} \sum_{j \geq 0} \binom{s}{j} \frac{1}{2^j} B(j, C),$$

where—remarkably enough—we know *every*  $B(j, C)$  via Theorem 6.3. As for a  $\Delta$ -series again for  $C := C_1(0)$ , we may use (8.2) similarly to develop

$$(8.7) \quad \Delta(s, C) = \frac{2^s}{2 \cdot 3^s - 1} \sum_{\text{even } j \geq 0} \binom{s}{j} \frac{1}{2^j} \Delta(j, C),$$

and knowledge of  $\Delta(j, C)$  for  $j = 0, 1, 2, 3, \dots$  gives rise to a fine numerical series.

It seemed to us natural to look at such expansions and infer asymptotic behavior for large  $s$ . Numerical experiments are embodied in accurate  $B$ -plots of Figure 4, and made good use of the fractal-box series (8.6); all of this leading us to

**Conjecture 8.3 (Asymptotics).** *Let  $\delta := \log_3 2$ , the fractal dimension of  $C_1(0)$ .*

(1) *For large  $s$  we have*

$$B(s, C_1(0)) \sim \frac{A}{s^\delta},$$

*where  $A \approx 0.7$  is an absolute constant.*

(2) *Similarly,*

$$\Delta(s, C_1(0)) \sim \frac{A'}{s^{2\delta}},$$

*where the absolute constant  $A' \approx 1.07$ .*

*Remark 8.4.* Conjecture 1 seems reasonable, on the very loose heuristic that if, in the sum (8.6), we have  $B_j \sim A/j^\delta$ , then a complicated integral-approximation argument gives the left-hand side  $\sim A/s^\delta$ . We have left out the heuristic argument, but we do emphasize the conjecture. The corresponding conjecture for  $\Delta(s, C_1(0))$ , follows from the numerical observation that  $2B(s, C_1(0))^2/\Delta(s, C_1(0))$  appears to tend to a limit of approximately 0.73 for large real  $|s|$ .  $\diamond$

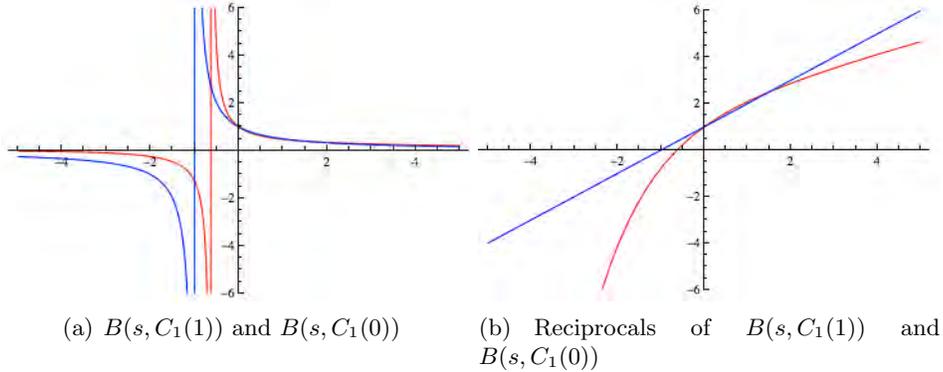


FIGURE 4. Analytic plots of box integrals for real  $s$ . Figure 4(a) shows  $B_1(s) = B(s, C_1(1)) = 1/(s+1)$  (with a pole at  $s = -1$ ) and  $B(s, C_1(0))$  (with a proven pole at  $s = -\log_3 2$ ). Figure 4(b) plots reciprocals: a perfectly straight line  $\sim s$  itself and  $1/B(s, C_1(0)) \sim s^{\log_3 2}$  for large  $s$  (unproven).

**8.3. High-precision algorithm for general  $B$ .** We next exhibit a novel algorithm that allows not only high-precision expectations,<sup>8</sup> but also *rigorous* bounds on such expectations. Note first the principle—immediate from the SCS definition—that we have the equality

$$C_n(P) = C_n(P \cdot P \cdots P),$$

where  $P \cdot P \cdots P$  is any finite string of copies of  $P$ . This trivial observation gives rise to a powerful computational expedient. We first recast (8.5) as

$$(8.8) \quad \left( 3^{pqs} \prod_{j=1}^p N_j^q - 1 \right) B(s, C_n(P)) = \sum_{U(c_k) \leq P_k} \langle |r + c_{p+1}3^{p'-p-1} + \cdots + c_{p'}|^s \rangle + \sum_{\substack{U(c_k) \leq P_k \\ (c_1, \dots, c_p) \neq (0_n, \dots, 0_n)}} \langle |r + c_13^{p'-1} + c_23^{p'-2} + \cdots + c_{p'}|^s \rangle,$$

where now the sequence  $(P_k)$  is interpreted as having period  $p' := pq$ . Freedom of choice on  $q$  allows arbitrarily large powers of 3 within the expectation terms on the right of (8.8) and allows recursion on the multiplicity factor

<sup>8</sup>Meaning, say, 20 digits. In previous works we have used the phrase “extreme precision” to mean at least 100 digits, or certainly enough to discover identities via integer-relation detection.

$q$ .<sup>9</sup> We arrive at

$$(8.9) \quad B(s, C_n(P)) = \frac{1}{Q(s, q, P)} \sum_{\substack{U(c_k) \leq P_k \\ (c_1, \dots, c_p) \neq (\mathbf{0}_n, \dots, \mathbf{0}_n)}} \langle |r + c_1 3^{p'-1} + c_2 3^{p'-2} + \dots + c_p|^s \rangle,$$

where

$$Q(s, q, P) := 3^{ps(q-1)} \prod_{j=1}^p N_j^{q-1} \left( 3^{ps} \prod_{j=1}^p N_j - 1 \right).$$

Since relation (8.9) is valid for every factor  $q = 1, 2, 3, \dots$  we have:

**Algorithm 8.5 (High-precision  $B$  computation with rigorous bounds).**

Given an SCS  $C_n(P)$  and a complex power  $s$ , the algorithm returns a precise value for  $B(s, C_n(P))$ , and—if desired, and for suitable  $s$ —rigorous bounds on  $B$ .

(1) For increasing  $q$ , calculate  $B$  from  $B(s, C_n(P))$

(8.10)

$$= \lim_{q \rightarrow \infty} \frac{1}{Q(s, q, P)} \sum_{\substack{U(c_k) \leq P_k \\ (c_1, \dots, c_p) \neq (\mathbf{0}_n, \dots, \mathbf{0}_n)}} |(\mathbf{1}/2)_n + c_1 3^{p'-1} + c_2 3^{p'-2} + \dots + c_p|^s,$$

where  $p' := pq$ , and we have fixed  $r$  to be the centroid of the unit  $n$ -cube  $[0, 1]^n$ , so expectation brackets have been removed.<sup>10</sup>

- (2) Approximations for increasing  $q$  approach the true  $B$  value, often smoothly enough that Aitken-like extrapolation will significantly increase accuracy.
- (3) For rigorous bounds, observe that all components of vectors within the expectation brackets in (8.8) are nonnegative. For real  $s \geq 0$  we may take  $r$  to be either the origin  $\mathbf{0}_n$  or the far apex  $\mathbf{1}_n$  of the unit  $n$ -cube, to deduce

$$\begin{aligned} \frac{1}{Q(s, q, P)} \sum_{\substack{U(c_k) \leq P_k \\ (c_1, \dots, c_p) \neq (\mathbf{0}_n, \dots, \mathbf{0}_n)}} |\mathbf{0}_n + c_1 3^{p'-1} + c_2 3^{p'-2} + \dots + c_p|^s \\ \leq B(s, C_n(P)) \leq \\ \frac{1}{Q(s, q, P)} \sum_{\substack{U(c_k) \leq P_k \\ (c_1, \dots, c_p) \neq (\mathbf{0}_n, \dots, \mathbf{0}_n)}} |\mathbf{1}_n + c_1 3^{p'-1} + c_2 3^{p'-2} + \dots + c_p|^s. \end{aligned}$$

Numerical examples of this algorithm in action are given in Section 9.  $\square$

<sup>9</sup>Indeed, the primed-sum in (8.8) is—up to a constant factor—a representation for the same  $B$  but involving  $(q-1)$  copies.

<sup>10</sup>Certainly this approximation sequence works for  $\Re(s) \geq 0$ ; we conjecture that it works for all complex  $s$ , based on numerical trials.

**8.4. General monomial moments.** We define a monomial moment in terms of an  $n$ -vector  $m = (m_1, \dots, m_n)$  of nonnegative integers

$$\mathcal{M}(m, C_n(P)) := \langle x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n} \rangle_{x \in C_n(P)}.$$

When the SCS  $C_n(P)$  is understood, we shall simplify to  ${}_n\mathcal{M}(m) = \mathcal{M}(m)$ .

First we list some elementary properties of monomial moments  $\mathcal{M}$ :

- (1)  $\mathcal{M}(m) = \mathcal{M}(m')$  where  $m'$  is any coordinate permutation of  $m$ . This follows from the symmetry in the definition of an SCS.
- (2) For any SCS,  $\mathcal{M}((1, 0, 0, 0, \dots)) = \frac{1}{2}$ . Again this follows from symmetry.
- (3) For any Cantor dust  $C_n(0)$ , the monomial moments are separable, that is

$$\mathcal{M}(m) := \left\langle \prod x_h^{m_h} \right\rangle = \prod \langle x_h^{m_h} \rangle.$$

This follows in any of several ways. One is to observe that the spectral  $G$  kernel separates for  $C_n(0)$ , so that expectation integrals completely factor. Another is to work through the combinatorics of Theorem 8.6 below.<sup>11</sup>

- (4) An immediate generalization is to *difference-monomial moments*. For separation problems involving  $\Delta$  expectations, we define

$$\mathcal{D}(m) := \langle (x_1 - q_1)^{m_1} \cdots (x_n - q_n)^{m_n} \rangle,$$

with elementary properties similar to those above. Now a coordinate separation variable  $x_i - q_i$  is bipolar, running over  $[-1, 1]$ .

We now provide a theorem generalizing the explicit results of Theorem 7.1:

**Theorem 8.6 (Monomial rationality).** *For any SCS  $C_n(P)$ , every monomial moment  $\mathcal{M}(m)$  is rational, and can be given an explicit closed form. The same properties hold for difference-monomial moments  $\mathcal{D}(m)$ .*

*Proof.* Now from the first functional relation in Proposition 5.5 we may use  $F(x) := \prod x_h^{m_h}$  to obtain

$$(8.11) \quad \mathcal{M}(m) = \frac{1}{\prod_{j=1}^p N_j} \sum_{U(c_k) \leq P_k} \left\langle \prod_{h=1}^n (x_h/3^p + c_{1h}/3 + \cdots + c_{ph}/3^p)^{m_h} \right\rangle,$$

where  $c_{jh}$  is the  $h$ -th element of column  $c_j$  from display (2.1). Now define the *weight* of a monomial as  $W(m) := \sum m_h$ , and observe that in this functional

<sup>11</sup>Caveat: we do not know precisely which monomial moments do *not* separate. For example, in 2 dimensions,  $\langle xy \rangle = \langle x \rangle \langle y \rangle$  for any SCS, meaning for any of  $C_2(0), C_2(1), C_2(2) = [0, 1]^2$ .

relation there are exactly  $\prod_{j=1}^p N_j$  summands, each having a leading term  $\langle \prod (x_h/3^p)^{m_h} \rangle$ , so that

$$(8.12) \quad \left(1 - \frac{1}{3^{pW}}\right) \mathcal{M}(m) = \sum R(m') \mathcal{M}(m'),$$

where the  $R$  coefficients are all rational and  $m'$  runs over a set of  $n$ -vectors each of whose weights  $W(m')$  being strictly less than  $W(m)$ . Therefore one has a multilevel recursion that has to finitely terminate with  $\mathcal{M}(m)$  being a rational number.  $\square$

**Corollary 8.7 (Even moment rationality).** *For any SCS  $C_n(P)$  and any nonnegative even integer  $u$ , both  $B(u, C_n(P))$  and  $\Delta(u, C_n(P))$  are rational and each can be given an explicit closed form.*

*Proof.* Everything follows with nonnegative even integer  $u$ , from

$$|x|^u = (x_1^2 + \cdots + x_n^2)^{u/2},$$

whence the right-hand side can be expanded in monomials.  $\square$

**Example 8.8 (Explicit coefficients).** An instance of Corollary 8.7 is

$$B(4, C_2(1)) = \langle (x^2 + y^2)^2 \rangle = 2\mathcal{M}((4, 0)) + 2\mathcal{M}((2, 2)) = \frac{429}{640},$$

while for the dust SCS, we obtain  $B(4, C_2(0)) = \frac{33}{40}$ . Moreover, in the case of  $n$ -dimensional Cantor dust we have complete separability of the underlying density (as discussed above for the monomials). Thence, for integers  $m > 0$  we obtain from the binomial theorem that

$$(8.13) \quad B(2m, C_n(0)) = \sum_{\substack{k_1, k_2, \dots, k_n \geq 0 \\ \sum_j k_j = m}} \binom{m}{k_1, \dots, k_n} \prod_{j=1}^n b_{2k_j}$$

where  $b_n := B(n, C_1(0))$  is as given by Theorem 6.3.  $\diamond$

**Example 8.9 (Monomial recursion).** An example of using the recursion in the proof of Theorem 8.6 is as follows. A collection of moments  $\mathcal{M}((i, j)) = \langle x^i y^j \rangle$  for vectors  $(x, y)$  on the SCS  $C_2(1)$  is

$$\left\langle \begin{array}{cccc} 1 & y & y^2 & y^3 \\ x & xy & xy^2 & xy^3 \\ x^2 & x^2y & x^2y^2 & x^2y^3 \\ x^3 & x^3y & x^3y^2 & x^3y^3 \end{array} \right\rangle = \begin{pmatrix} 1 & \frac{1}{2} & \frac{11}{32} & \frac{17}{64} \\ \frac{1}{2} & \frac{1}{4} & \frac{11}{64} & \frac{17}{128} \\ \frac{11}{32} & \frac{11}{64} & \frac{601}{5120} & \frac{923}{10240} \\ \frac{17}{64} & \frac{17}{128} & \frac{923}{10240} & \frac{1409}{20480} \end{pmatrix},$$

whereas for the SCS  $C_2(0)$ , the corresponding matrix of moments is

$$= \begin{pmatrix} 1 & \frac{1}{2} & \frac{3}{8} & \frac{5}{16} \\ \frac{1}{2} & \frac{1}{4} & \frac{3}{16} & \frac{5}{32} \\ \frac{3}{8} & \frac{3}{16} & \frac{9}{64} & \frac{15}{128} \\ \frac{5}{16} & \frac{5}{32} & \frac{15}{128} & \frac{25}{256} \end{pmatrix}.$$

Remarkably,  $\langle xy \rangle = 1/4$  separates as  $(1/2) \cdot (1.2)$  for each SCS here, and for  $C_2(2) = [0, 1]^2$ , but some other elements of the matrix for  $C_2(1)$  are *not* separable.  $\diamond$

## 9. NUMERICAL ALGORITHMS, RESULTS, AND CHALLENGES

We have given rational closed  $B(s, \cdot), \Delta(s, \cdot)$  forms for all non-negative integer  $s$  in  $n = 1$  dimension, and for all dimensions  $n$  we have proven the existence of rational closed forms when  $s$  is even. Yet, we run into trouble with regards to exact evaluation when  $n \geq 2$  and  $s$  is not an even integer, say  $s = 1, n = 2$  as the canonical ‘open case.’

In a word: We do not yet know the expectation of distance from the origin  $B(1, \cdot)$  or the expected separation  $\Delta(1, \cdot)$  on any non-trivial SCS embedded in  $n \geq 2$  dimensions. Even more stultifying: the scientifically important case  $\Delta(-1, C_3(P))$  seems to be inaccessible in closed form.<sup>12</sup>

**Example 9.1.** [Selected numerical excursions] We list some of our successes and some of the challenges remaining.

1. An initial numerical foray was that of V. Klungre and D. Bailey (2010), who used ‘offset-box’ integrals (see Appendix A) to achieve numerical values:

$$\Delta(1, C_2(0)) = 0.63644048(5) \dots, \quad \Delta(1, C_2(1)) = 0.553861543(7) \dots,$$

where the symbol (digit) means it is questionable. By contrast, the full-square classical case is known exactly [3]

$$\Delta(1, C_2(2)) = \Delta_2(1) = \frac{1}{15} \left( 2 + \sqrt{2} + 5 \log \left( 1 + \sqrt{2} \right) \right) = 0.521405433164721 \dots,$$

where fractal dimension and expected separation are both monotonically ordered over the three SCS cases. (We have previously, though, by way of Remark 7.3 demolished any general monotonicity-ordering conjecture.)

---

<sup>12</sup>This lies in embedding dimension  $n = 3$  (relevant to laboratory work with brain tissue) and is the expectation of  $1/r$ , and so works to reject somewhat the effects of finite cuboid boundaries. Pragmatically even better might be such expectations as the *Yukawa* form  $\langle e^{-\lambda r}/r \rangle$  [11].

2. Application of the high-precision Algorithm 8.5 in Section 8.3 for the  $(n = 2)$ -dimensional dust  $C_2(0)$  involves the numerically powerful approximation (note  $p = 1$  for this SCS, with the approximation exact as  $q \rightarrow \infty$ ):

(9.1)

$$B(s, C_2(0)) \approx \frac{3^{(1-q)s} 4^{1-q}}{4 \cdot 3^s - 1} \sum_{\substack{U(c_k)=0 \\ c_1 \neq (0,0)}} |(\mathbf{1}/2)_n + c_1 3^{q-1} + c_2 3^{q-2} + \dots + c_q|^s,$$

where it will be observed that the summation is over  $q$  ternary 2-vectors with the constraint that  $c_1$  is not all 0's. This approach leads via the algorithm to *rigorous* bounds such as

$$\frac{791059}{1000000} < B(1, C_2(0)) < \frac{791062}{1000000}.$$

The approximation scheme of the algorithm, on the other hand, yields results (using multiplicity  $q \geq 10$ ) such as

$$\begin{aligned} B(4, C_2(0)) &= 0.8250000000000000(0), \\ B(1, C_2(0)) &= 0.7910607171001881694140(5) \dots, \\ B(1/2, C_2(0)) &= 0.85759154619636804162(5) \dots, \\ B(-1, C_2(0)) &= 3.22851042553756462173(5) \dots, \\ B(-12618595/1000000, C_2(0)) &= 105529978.23182819(5) \dots, \\ B(-2, C_2(0)) &= -0.7222518765084439(9) \dots, \\ B(-3, C_2(0)) &= -0.1820952173493284(1) \dots \end{aligned}$$

where only the parenthetic digit is in doubt. This is evidently very difficult to match in precision using Monte Carlo (MC) methods. (Of course, MC methods are still useful, see the Appendix, as strong checks on any other algorithm.) Note that the exact value of  $B(4, C_2(0))$  is  $33/40$ , so the reported value lends credibility to the algorithm. Also the  $B$  values for sufficiently negative  $s$  are *negative*—for these are analytic continuation values to the left of the pole at  $s = \log_3 4$ ; note a  $B$  value in excess of  $10^8$  for  $s$  near this pole is on the list above.

3. More recently, we applied self-similarity expansions to effect more precise box-integral values. By self-similarity expansions we refer to the use of formula such as the following one, appropriate for the function  $f(x, y) = (x^2 + y^2)^{s/2}$ , with the fractal  $C_2(0)$  assumed. Note that  $B(s, C_2(0)) = \langle f(x, y) \rangle$ :

(9.2)

$$(4 - 3^{-s}) B(s, C_2(0)) = 2 \left\langle f \left( \frac{x+2}{3}, \frac{y}{3} \right) \right\rangle + \left\langle f \left( \frac{x+2}{3}, \frac{y+2}{3} \right) \right\rangle.$$

One uses the binomial expansion on the right-hand-side and inserts known, exact even-power moments. Applying self-similarity a few times recursively gives intricate expansions with perhaps superior convergence. In this way

we achieved (absent symbolic/numeric details attendant on 2-dimensional Taylor expansions):

$$\Delta(1, C_2(0)) = 0.636440485697895310368137114931019(3) \dots$$

Though this is probably too imprecise to allow a modern experimental closed-form search, it is nevertheless hard to imagine obtaining such 30+ decimal resolution via direct point-counting or offset-box summations.

This notion of moment expansions suggests there may be a sharp improvement to Algorithm 8.5; namely, instead of forcing  $r \rightarrow (\mathbf{1}/\mathbf{2})_n$ , the centroid value, we can attempt expansion around the centroid and use known even moments. Our best precision to date for ( $n = 2$ ) dimensions being the 32 digits for  $\Delta(1, C_2(0))$  above indicates that such future research is called for.

4. This self-similar expansion method is instructive in ( $n = 1$ )-dimensional cases. We know from Section 8 that for  $d := r - 1/2$ ,

$$B(1/2, C_1(0)) = \frac{2}{2 - \frac{1}{\sqrt{3}}} \sqrt{\frac{5}{6}} \left\langle \sqrt{1 + \frac{2d}{5}} \right\rangle.$$

Binomial expansion of the  $d$ -dependent term, with insertion into the expansion of the  $d$ -moments, results in extreme precision in reasonable time, e.g.,  $B(1/2, C_1(0)) =$

$$0.640051038674413046291777407650533688744217331985844782542398 \dots$$

with many more places possible. Alternatively, series (8.6) leads to the same results.

By contrast, estimation with fractional  $s$  in higher dimensions is trickier. Even for  $n = 2$ ,  $s = -1$  we have been able to do any better than

$$\Delta(-1, C_2(0)) = 3.927(1) \dots$$

Incidentally, this expectation of (separation)<sup>-1</sup>—speaking formally—should agree with

$$\Delta(-1, C_2(0)) = \frac{2}{\pi} \int_0^\infty \int_0^\infty \frac{G(u)^2 G(v)^2}{\sqrt{u^2 + v^2}} du dv,$$

where we refer to the kernel  $G(w) := {}_1G_0(w) = \prod_{m \geq 1} \cos(w/3^m)$ , yet we have been unable to get even the rough value 3.927... via quadrature.

5. Indeed such numerical quadrature is for whatever reason generally problematic. As derived  $\Delta(1, C_1(0)) = 2/5$ , leading to the peculiar integral identity

$$\int_0^\infty \frac{\cos k + k \sin k - 1}{k^2} G(k)^2 dk = \frac{\pi}{5},$$

which, again, does not seem to yield readily to numerical quadrature techniques beyond just a few good decimals.  $\diamond$

10. SOME OPEN QUESTIONS

We finish by emphasizing some outstanding questions:

- How can we get expectations such as  $\Delta(1, C_2(0))$  to enough accuracy (say 100+ decimals) for experimental-mathematical analysis as explored in [6]? (Note the final observations of Example 9.1 on possible refinements on Algorithm 8.5.)
- It is shown in [12] that if the standard Cantor set is charged (say with a total charge of +1), then the electrostatic potential at the position  $(0, y)$ —so directly above the left-hand edge of  $C_1(0)$ —is given by

$$V(0, y) = \frac{1}{2\pi} \int_0^\infty J_0(yk) G_1(k) dk,$$

with  ${}_1G_1$  as in Section 6 and  $J_0$  the Bessel function. To date, this integral has not been evaluated, even asymptotically. However, it is known that  $V(0, y)$  behaves something like  $1/y^{1-\delta}$  where  $\delta = \log_3 2$ . Might any of the new techniques in the present paper apply?

- For brain-synapse analysis, we would like to know at least a good numerical values for  $\Delta(s, C_3(P))$ —meaning separation moments in  $(n = 3)$  embedding dimensions—and  $s = -2, -1, 1$ . It would be useful, therefore, to work out a  $\Delta$ -analogue of Algorithm 8.5.
- How much can we build on these results to evaluate expectations on a larger class of fractal sets embedded in the unit hypercube? Perhaps on all fractals covered by Theorem 2.4.

ACKNOWLEDGEMENTS

The authors are grateful to Roland Girgensohn for bringing Theorem 2.4 to our attention and thus suggesting a more elegant approach to the proof of Proposition 1. They are also indebted to Dirk Nuyens and Josef Dick for discussions on Monte Carlo integration. Thomas Wieting aided us generously in matters of fractal theory.

REFERENCES

1. R.S. Anderssen et al, *Concerning  $\int_0^1 \cdots \int_0^1 (x_1^2 + \cdots + x_k^2)^{1/2} x_1 \cdots x_k$  and a Taylor Series Method*, SIAM Journal on Applied Mathematics **30** (1976), no. 1, 22-30.
2. D.H. Bailey, J.M. Borwein and R.E. Crandall, *Advances in the theory of box integrals*, Mathematics of Computation **79** (2009), 1839-1866.
3. D.H. Bailey, J.M. Borwein and R.E. Crandall, *Box integrals*, Journal of Computational and Applied Mathematics **206** (2007), 196-208.
4. M. Barnsley, *Fractals everywhere*. Academic Press, Inc., Boston, MA, 1988.
5. A.K. Basu, *Measure Theory and Probability*, PHI Learning Pvt. Ltd., 2004, p.113.
6. J.M. Borwein and D.H. Bailey, *Mathematics by Experiment: Plausible Reasoning in the 21st Century*, A.K. Peters Ltd. Second expanded edition, September 2008.
7. J.M. Borwein, D.H. Bailey and R. Girgensohn, *Experimentation in Mathematics: Computational Paths to Discovery*, A.K. Peters Ltd, 2004.
8. J.M. Borwein and R. Girgensohn, *Functional equations and distribution functions (for Janos Aczel's 70th Birthday)*, Results in Mathematics **26** (1994), 229-237.

9. J.M. Borwein, O. Chan and R.E. Crandall, *Higher-dimensional box integrals*, *Experimental Mathematics* **19** (2010), no. 4, 431-446.
10. A. Y. Cherny, E. M. Anitas, A. I. Kuklin, M. Balasoiu and V. A. Osipov, *Scattering from generalized Cantor fractals*, *J. Appl. Cryst.* **43**, (2010), 790–797. doi:10.1107/S0021889810014184.
11. R.E. Crandall, *On the fractal distribution of brain synapses*, <http://www.perfscipress.com/papers/Synapses.pdf> and to appear, *Computational and Analytical Mathematics, Proceedings*, in Springer Series in Mathematics and Statistics, 2013.
12. R.E. Crandall, *Theory of box series*, in *Scientific reflections: Selected multidisciplinary works*, PSIPress, 2011.
13. R.E. Crandall, *Topics in Advanced Scientific Computation*, Springer, New York, (1997).
14. K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications (2nd edition)*, John Wiley & Sons, West Sussex (2003).
15. *How Mathematics Can Reveal Brain Structure and Diagnose Brain Health*. 5th June 2012. Available at <http://www.sciguru.com/newsitem/13858/how-mathematics-can-reveal-brain-structure-and-diagnose-brain-health>, accessed 11th July 2012.

## APPENDIX A. EXPECTATION DATA FOR SCS CASES

**A.1. Scatterplots of second-order box integrals vs fractal dimension.** The separation moments  $B(2, C_n(P))$  and  $\Delta(2, C_n(P))$  (computed using Theorem 7.1) and fractal dimension (computed using Proposition 7) of various SCSs were plotted for embedding dimensions  $n = 2$  and 3. The resulting scatterplots are shown below.

**A.2. Exact rational values and numerical verification for second-order expectations.** Below, for a selection of string-generated Cantor sets  $C_n(P)$ , we present data comprising fractal dimensions  $\delta(C_n(P))$  and second-order moments  $B(2, C_n(P))$  and  $\Delta(2, C_n(P))$  for embedding dimension  $n = 1, 2$  and 3, with entries arranged in order of increasing fractal dimension. All such quantities are rational and were computed exactly—fractal dimensions via Proposition 7, and separation moments via Theorem 7.1. Decimal equivalents (to six-digit accuracy) are included, together with results of Monte-Carlo computations, as a check.

- (1) One possible Monte-Carlo approach is to utilize a uniform  $(0, 1)$  pseudorandom number generator to generate pairs of  $n$ -tuples at random, then check each pair of  $n$ -tuples so generated to see if it is “admissible” for the given set  $C_n(P)$ . This is certainly a relatively simple and straightforward scheme to code on a computer, and can be used effectively in cases where the average “density” of the fractal set is not too small.
  - The disadvantage of this scheme is that for many fractal Cantor sets that one might wish to examine, the density is so small that only a microscopic fraction of the  $n$ -tuple pairs so generated in each trial are admissible. In such cases, thousands or even millions of pairs of  $n$ -tuples must be generated to find just one

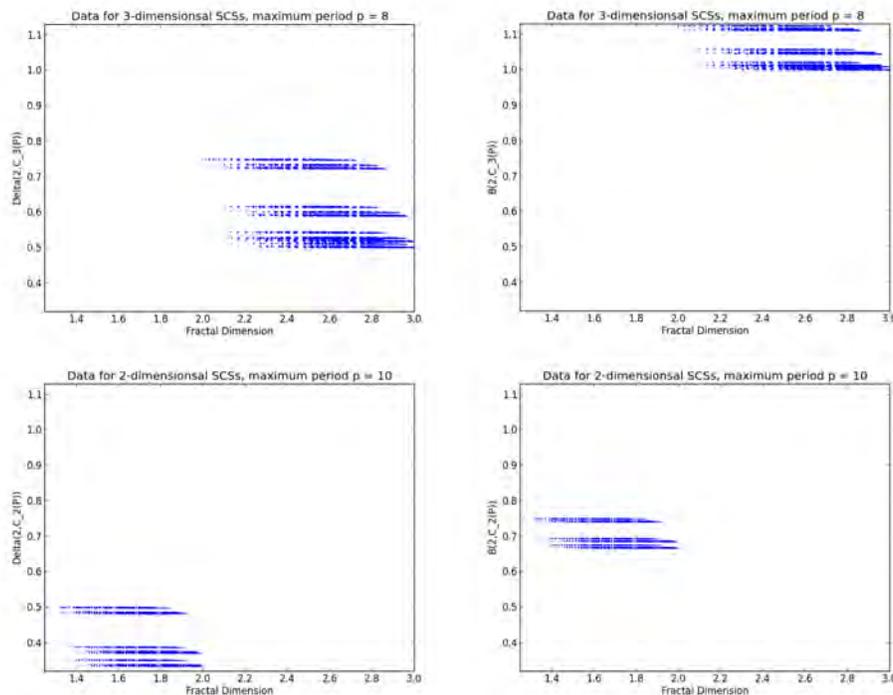


FIGURE 5. Separation moments plotted against fractal dimensions for SCSs. Top: separation moments  $\Delta(2, C_3(P))$  (left) and  $B(2, C_3(P))$  (right) for SCSs of period 8 or less in embedding dimension  $n = 3$ . Bottom: separation moments  $\Delta(2, C_2(P))$  (left) and  $B(2, C_2(P))$  (right) for SCSs of period 10 or less in embedding dimension  $n = 2$ . All plots have the same scale.

admissible pair, and yet millions of such valid pairs must be generated to obtain reliable mean statistics.

- (2) A second approach, which we adopted below, is, for each case to be studied, to first construct a table of all admissible columns  $c_k$  for each of the components of  $P$ , and then, for each of a large number of trials, and for each ternary column  $1 \leq k \leq 20$  of a single trial, pseudorandomly select an admissible column from the appropriate table. While coding this scheme is much more complicated than the more straightforward scheme, it has the distinct advantage that each pair of  $n$ -tuples is guaranteed to be admissible.

- We have tabulated our results, with the columns headed by “Numeric” giving Monte-Carlo results for  $B(2, C_n(P))$  and  $\Delta(2, C_n(P))$ . In each line, figures are based on a computation of  $10^9$  pseudorandom pairs of admissible  $n$ -tuples, each to 20 ternary digit precision. Note that 20 ternary digit precision corresponds to

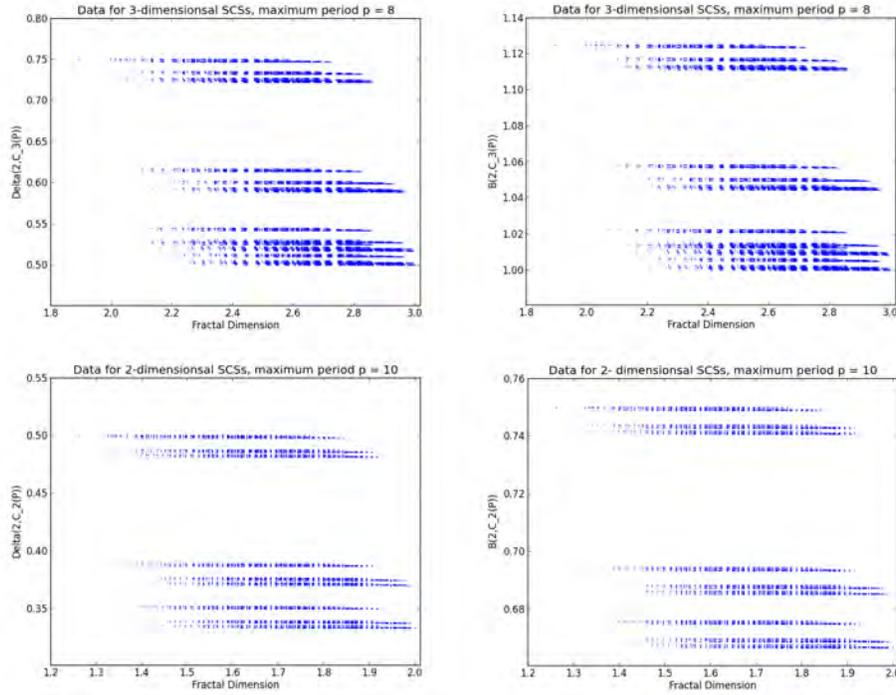


FIGURE 6. Separation moments plotted against fractal dimensions for SCSs. Top: separation moments  $\Delta(2, C_3(P))$  (left) and  $B(2, C_3(P))$  (right) for SCSs of period 8 or less in embedding dimension  $n = 3$ . Bottom: separation moments  $\Delta(2, C_2(P))$  (left) and  $B(2, C_2(P))$  (right) for SCSs of period 10 or less in embedding dimension  $n = 2$ . The scale of each plot has been adjusted to fill the plotting area.

roughly 9.5 decimal digit precision, which is adequate given that fact that even with  $10^9$  pseudorandom trials per case, only about 5 accurate digits can be expected in the output means. At the bottom of each of the three tables are statistics giving the *maximum* and *root-mean-square* error for the Monte-Carlo results in the table. We observe that theory and computation mesh very well.

- As a final set of results, we also exhibit, in Table 6, a few validating results for the box integrals  $B_n(s, C_n(P))$ , kindly provided to the authors by Dirk Nuyens, who employed a multi-level Monte-Carlo scheme to compute moments for  $s = 1$ ,  $s = 2$  and  $s = 1/2$ , for various  $C_n(P)$ .
- (3) We mention one other approach, which was taken by V. Klungre in Example 9.1 of the previous section. Klungre utilized a *Mathematica* program to compute the mean distance, in any given dimension  $n$ ,

between points in two square (or cube, etc.) Cantor patches. He then recursively extended this scheme up to dimension 9. While the results so produced are very accurate, so much computation is required for a single case that it is not practical for a large number of cases, as we required below.

We are investigating more sophisticated quasi-Monte-Carlo schemes, but do not yet have any results.

A.2.1. *Tabulated data for  $n = 1$  (generating string period  $\leq 4$ ), with numeric Monte-Carlo results based on  $10^9$  trials.*

P	$\delta(C_1(P))$	$B(2, C_1(P))$			$\Delta(2, C_1(P))$		
	Decimal	Rational	Decimal	Numeric	Rational	Decimal	Numeric
0	0.630930	$\frac{3}{8}$	0.375000	0.375013	$\frac{1}{4}$	0.250000	0.250009
0001	0.723197	$\frac{7379}{19680}$	0.374949	0.374941	$\frac{2459}{9840}$	0.249898	0.249904
0010	0.723197	$\frac{2457}{6560}$	0.374543	0.374543	$\frac{817}{3280}$	0.249085	0.249080
0100	0.723197	$\frac{2433}{6560}$	0.370884	0.370887	$\frac{793}{3280}$	0.241768	0.241765
1000	0.723197	$\frac{2217}{6560}$	0.337957	0.337937	$\frac{577}{3280}$	0.175915	0.175928
001	0.753953	$\frac{409}{1092}$	0.374542	0.374552	$\frac{68}{273}$	0.249084	0.249092
010	0.753953	$\frac{135}{364}$	0.370879	0.370877	$\frac{22}{91}$	0.241758	0.241751
100	0.753953	$\frac{123}{364}$	0.337912	0.337895	$\frac{16}{91}$	0.175824	0.175824
0011	0.815465	$\frac{737}{1968}$	0.374492	0.374486	$\frac{245}{984}$	0.248984	0.248989
01	0.815465	$\frac{89}{240}$	0.370833	0.370836	$\frac{29}{120}$	0.241667	0.241661
0110	0.815465	$\frac{243}{656}$	0.370427	0.370429	$\frac{79}{328}$	0.240854	0.240850
10	0.815465	$\frac{27}{80}$	0.337500	0.337481	$\frac{7}{40}$	0.175000	0.175014
1001	0.815465	$\frac{665}{1968}$	0.337907	0.337894	$\frac{173}{984}$	0.175813	0.175812
1100	0.815465	$\frac{219}{656}$	0.333841	0.333825	$\frac{55}{328}$	0.167683	0.167686
011	0.876977	$\frac{809}{2184}$	0.370421	0.370397	$\frac{263}{1092}$	0.240842	0.240847
101	0.876977	$\frac{737}{2184}$	0.337454	0.337442	$\frac{191}{1092}$	0.174908	0.174904
110	0.876977	$\frac{243}{728}$	0.333791	0.333781	$\frac{1}{6}$	0.167582	0.167583
0111	0.907732	$\frac{7289}{19680}$	0.370376	0.370350	$\frac{2369}{9840}$	0.240752	0.240757
1011	0.907732	$\frac{6641}{19680}$	0.337449	0.337440	$\frac{1721}{9840}$	0.174898	0.174903
1101	0.907732	$\frac{6569}{19680}$	0.333791	0.333765	$\frac{1649}{9840}$	0.167581	0.167574
1110	0.907732	$\frac{2187}{6560}$	0.333384	0.333386	$\frac{547}{3280}$	0.166768	0.166774
1	1.000000	$\frac{1}{3}$	0.333333	0.333333	$\frac{1}{6}$	0.166667	0.166671
Max error				0.000033			0.000026

RMS error		0.000014	0.000008
-----------	--	----------	----------

A.2.2. *Tabulated data for  $n = 2$  (generating string period  $\leq 2$ ), with numeric Monte-Carlo results based on  $10^9$  trials.*

P	$\delta(C_2(P))$	$B(2, C_2(P))$			$\Delta(2, C_2(P))$		
	Decimal	Rational	Decimal	Numeric	Rational	Decimal	Numeric
0	1.26186	$\frac{3}{4}$	0.750000	0.750012	$\frac{1}{2}$	0.500000	0.500036
01	1.577324	$\frac{119}{160}$	0.743750	0.743744	$\frac{39}{80}$	0.487500	0.487478
10	1.577324	$\frac{111}{160}$	0.693750	0.693716	$\frac{31}{80}$	0.387500	0.387514
02	1.630930	$\frac{89}{120}$	0.741667	0.741664	$\frac{29}{60}$	0.483333	0.483325
20	1.630930	$\frac{27}{40}$	0.675000	0.674983	$\frac{7}{20}$	0.350000	0.349996
1	1.892789	$\frac{11}{16}$	0.687500	0.687486	$\frac{3}{8}$	0.375000	0.374998
12	1.946395	$\frac{329}{480}$	0.685417	0.685414	$\frac{89}{240}$	0.370833	0.370830
21	1.946395	$\frac{107}{160}$	0.668750	0.668743	$\frac{27}{80}$	0.337500	0.337494
2	2.000000	$\frac{2}{3}$	0.666667	0.666646	$\frac{1}{3}$	0.333333	0.333335
Max error				0.000057			0.000036
RMS error				0.000024			0.000010

A.2.3. *Tabulated data for  $n = 3$  (generating string period  $\leq 2$ ), with numeric Monte-Carlo results based on  $10^9$  trials.*

P	$\delta(C_3(P))$	$B(2, C_3(P))$			$\Delta(2, C_3(P))$		
	Decimal	Rational	Decimal	Numeric	Rational	Decimal	Numeric
0	1.892789	$\frac{9}{8}$	1.125000	1.124983	$\frac{3}{4}$	0.750000	0.749994
01	2.309811	$\frac{447}{400}$	1.117500	1.117460	$\frac{147}{200}$	0.735000	0.734995
10	2.309811	$\frac{423}{400}$	1.057500	1.057481	$\frac{123}{200}$	0.615000	0.614989
02	2.429218	$\frac{579}{520}$	1.113462	1.113436	$\frac{189}{260}$	0.726923	0.726925
20	2.429218	$\frac{531}{520}$	1.021154	1.021104	$\frac{141}{260}$	0.542308	0.542313
03	2.446395	$\frac{89}{80}$	1.112500	1.112453	$\frac{29}{40}$	0.725000	0.725009
30	2.446395	$\frac{81}{80}$	1.012500	1.012502	$\frac{21}{40}$	0.525000	0.525011
1	2.726833	$\frac{21}{20}$	1.050000	1.049967	$\frac{3}{5}$	0.600000	0.599985
12	2.846240	$\frac{5439}{5200}$	1.045962	1.045915	$\frac{1539}{2600}$	0.591923	0.591917

$C$	$B(2, C)$	$B(1, C)$	$B(1/2, C)$
$C_3(0)$	1.12500	1.00501	0.98359
$C_3(1)$	1.05000	0.97909	0.97522
$C_3(2)$	1.00961	0.96424	0.96957
$C_3(3)$	1.00000	0.96059	0.96811
$C_2(0)$	0.75000	0.79106	0.85759
$C_2(1)$	0.68750	0.77261	0.85778
$C_2(2)$	0.66667	0.76520	0.85581
$C_1(0)$	0.37500	0.50000	0.64005
$C_1(1)$	0.33333	0.50000	0.66667
$C_3(01)$	1.11750	1.00196	0.98229
$C_3(10)$	1.05750	0.98205	0.97648
$C_3(02)$	1.11346	1.00028	0.98151
$C_2(01)$	0.74375	0.78846	0.85735
$C_2(10)$	0.69375	0.77507	0.85830
$C_2(02)$	0.74167	0.78749	0.85700
$C_2(20)$	0.67500	0.76846	0.85662
$C_2(12)$	0.68542	0.77174	0.85748
$C_2(21)$	0.66875	0.76605	0.85610
$C$	$\Delta(2, C)$	$\Delta(1, C)$	$\Delta(1/2, C)$
$C_3(01)$	0.3675	0.5969	0.7694
$C_3(10)$	0.3075	0.5404	0.7299
$C_3(02)$	0.3634	0.5946	0.7683

TABLE 6. The expectation integrals for various  $s$ , computed by Dirk Nuyens (for  $B$ ) and Josef Dick (for  $\Delta$ ).

21	2.846240	$\frac{5271}{5200}$	1.013654	1.013606	$\frac{1371}{2600}$	0.527308	0.527303
13	2.863417	$\frac{209}{200}$	1.045000	1.045028	$\frac{59}{100}$	0.590000	0.589982
31	2.863417	$\frac{201}{200}$	1.005000	1.004979	$\frac{51}{100}$	0.510000	0.509998
2	2.965647	$\frac{105}{104}$	1.009615	1.009601	$\frac{27}{52}$	0.519231	0.519224
23	2.982824	$\frac{1049}{1040}$	1.008654	1.008633	$\frac{269}{520}$	0.517308	0.517303
32	2.982824	$\frac{1041}{1040}$	1.000962	1.000925	$\frac{261}{520}$	0.501923	0.501918
3	3.000000	1	1.000000	0.999960	$\frac{1}{2}$	0.500000	0.499983
Max error				0.000065			0.000032
RMS error				0.000031			0.000012

LAWRENCE BERKELEY NATIONAL LABORATORY, BERKELEY, CA 94720.

*E-mail address:* `dhbailey@lbl.gov`

CENTRE FOR COMPUTER ASSISTED RESEARCH MATHEMATICS AND ITS APPLICATIONS (CARMA), UNIVERSITY OF NEWCASTLE, CALLAGHAN, NSW 2308, AUSTRALIA. DISTINGUISHED PROFESSOR, KING ABDUL-AZIZ UNIVERSITY, JEDDAH.

*E-mail address:* `jonathan.borwein@newcastle.edu.au`

CENTER FOR ADVANCED COMPUTATION, REED COLLEGE, PORTLAND OR.

*E-mail address:* `crandall@reed.edu`

CENTRE FOR COMPUTER ASSISTED RESEARCH MATHEMATICS AND ITS APPLICATIONS (CARMA), UNIVERSITY OF NEWCASTLE, CALLAGHAN, NSW 2308, AUSTRALIA.

*E-mail address:* `michael.rose@newcastle.edu.au`