MEASURE OF THE VERY FAT CANTOR SET

LINAS VEPSTAS <LINAS@LINAS.ORG>

ABSTRACT. This breif note defines the idea of a "very fat" Cantor set, and breifly examines the measure associated with such a very fat Cantor set.

The canonical Cantor set is "thin" in that it has a measure of zero. There are a variety of methods by which on can construct "fat" Cantor sets (also known as Smith-Volterra-Cantor sets) which have a measure greater than zero. One of the commonest constructions, based on the dyadic numbers, has a continuously-varying parameter that is associated with the measure. The Smith-Volterra-Cantor set attains a measure of one for a finite value of the parameter; this paper then explores what happens when the parameter is pushed beyond this value. These are the "very fat" Cantor sets refered to in the title. The results consist almost entirely of a set of graphs showing this behaviour.

This paper is part of a set of chapters that explore the relationship between the real numbers, the modular group, and fractals.

1. INTRO

THIS IS A DRAFT. The intro hasn't been written yet, but if it was, it would work like this:

The "very fat" Cantor set and the concept of its measure is defined, and some of these are graphed.

Although much could be said about the topology associated with this construction, nothing is said here.

Although the symmetry properties of the resulting measures could be analyzed in light of the SL(2,Z) period-doubling monoid symmetry, they are not. Maybe later.

2. AN OPEN DYADIC TOPOLOGY

Consider the the collection $\mathcal{A}$ of open intervals given by

$$\begin{aligned} \mathcal{A} &= \left\{ \left(\frac{m}{2^n}, \frac{m+1}{2^n} \right) \middle| \; n \in \mathbb{N}_0 \wedge 0 \leq m < 2^n \right\} \\ &= \left\{ (0,1), \left(0, \frac{1}{2}\right), \left(\frac{1}{2}, 1\right), \left(0, \frac{1}{4}\right), \left(\frac{1}{4}, \frac{1}{2}\right), \cdots \right\} \end{aligned} \tag{2.1}$$

and let $\mathcal{B}$ be the disjoint union of $\mathcal{A}$ and the empty set. That is, $\mathcal{B}$ is the set containing the elements of $\mathcal{A}$ together with the empty set.

The set $\mathcal{B}$ of open intervals forms a base for a topology on the open unit interval $X = (0,1)$. One may verify that it satisfies the axioms for being a topological base. A base for X must it cover X, and clearly $\mathcal{A}$ does this. A base must be closed under the intersection of its elements. Clearly, for any two elements $A_i, A_j \in \mathcal{A}$, one has that $A_i \cap A_j$ is equal to either A_i or $A_j$or is the empty set. Thus $A_i \cap A_j \in \mathcal{B}$, which is why $\mathcal{B}$ was created from $\mathcal{A}$: adjoining the emptyset turns $\mathcal{B}$ into the closure of $\mathcal{A}$ under the operation of intersection.

Date: 31 December 2005 (updates 20 Oct 2006).

Since $\mathcal{B}$ may be obtained by means of a finite intersection of elements of $\mathcal{A}$, one has that $\mathcal{A}$ is a sub-base for $\mathcal{B}$.

The basis $\mathcal{B}$ generates a topology $\mathcal{T}$, so that elements of $\mathcal{T}$ may be expressed as the countable union of elements of $\mathcal{B}$. That is, every element $T \in \mathcal{T}$ may be written as the finite or possibly countably-infinite union of elements of $\mathcal{B}$:

$$T = A_1 \cup A_2 \cup \cdots \tag{2.2}$$

where each $A_k \in \mathcal{B}$.

The topology $\mathcal{T}$ is a topology on the reals, but this topology is clearly coarser than the natural topology. In particular, the interval $(0,1/3)$ is not in $\mathcal{T}$. The closest that one can come is the countable union

$$\left(0,\frac{1}{4}\right) \cup \left(\frac{1}{4},\frac{5}{16}\right) \cup \left(\frac{5}{16},\frac{21}{64}\right) \cup \cdots \tag{2.3}$$

which is properly contained in $(0,1/3)$, and is missing the points at $1/4$, $5/16$ and so on.

The topology $\mathcal{T}$ is *not* a sigma-algebra. This is because $\mathcal{T}$ is not closed under complementation. In fact, $\mathcal{T}$ does not contain the complement of any of its elements, except for the empty set and the whole space X itself. The complement $\neg T$ of $T \in \mathcal{T}$ is defined as the set such that $\neg T \cap T = \varnothing$ while $\neg T \cup T = X$. Clearly $\mathcal{T}$ fails to contain its complements.

XXX Elaborate on some other topological properties XXX.

2.1. **A measure on this topology.** The set $\mathcal{B}$ has a countable number of elements, which may be conveniently enumerated by the mid-points of the intervals. In particular, the elements of $\mathcal{B}$ are in one-to-one correpondance with the dyadic rationals. Insofar as one may want to define a measure on $\mathcal{T}$, this measure may be specified by a function taking values on the dyadic rationals. A measure is a function

$$\mu : \mathcal{T} \to \mathbb{R}^+ \tag{2.4}$$

that to each element of $\mathcal{T}$ assigns a non-negative real number, the measure of the element. Measures must have the additional property of being additive, in that one must have

$$\mu(T_j \cup T_k) = \mu(T_j) + \mu(T_k) \tag{2.5}$$

whenever $T_j \cap T_k = \varnothing$. Now, the topology $\mathcal{T}$ has the curious property that every element of $\mathcal{T}$ may be written as the (countable) union of pairwise-disjoint sets, and furthermore, this collection of pair-wise disjoint sets is a subset of the basis $\mathcal{B}$. Thus, to define an additive measure on $\mathcal{T}$, it is sufficient to define an additive measure on $\mathcal{B}$, which can then be extended uniquely to an additive measure on $\mathcal{T}$.

Because the the base $\mathcal{B}$ is countable, the measure may be uniquely defined by specifying a countable set of real numbers. Obe begins with the usual normalization, namely that $\mu(X) = 1$. To the interval $(0,1/2)$ one may assign a measure a while to the interval $(1/2,1)$ one may assign the measure b. In order to have this measure correspond to the usual sigma-additive measures on the reals, one might demand that $a+b=1$. However, there is nothing in particular about the topology $\mathcal{T}$ that demands this, and one might suggest that $a+b<1$. Such a measure would still obey the containment axiom, namely, that $\mu(A) \leq \mu(B)$ whenever $A \subseteq B$. Since $\mathcal{T}$ is not a sigma-algebra, as demonstrated above, one cannot invert this axiom by taking the complement of A and B, thus forcing the equality $a+b=1$.

This deficiency of the measure can be codified with a definition. Enumerate elements of the set $\mathcal{B}$ by a pair of integers (m,n) so that $(2m+1)/2^n$ is the midpoint of the interval $I_{(m,n)}$.

Definition 2.1. A measure μ is said to be **full** if

$$\mu\left(I_{(m,n)}\right)=\mu\left(I_{(2m,n+1)}\right)+\mu\left(I_{(2m+1,n+1)}\right) \tag{2.6}$$

for all intervals; else the measure is **deficient** if

$$\mu\left(I_{(m,n)}\right)\geq\mu\left(I_{(2m,n+1)}\right)+\mu\left(I_{(2m+1,n+1)}\right) \tag{2.7}$$

If strict inequality holds, then the measure is **strictly deficient**.

Any non-negative real function on the dyadic rationals which obeys the above conditions can be taken to be a measure on this topology.

3. A Closed Dyadic Topology

Consider the collection $\mathcal{C}$ of closed intervals given by

$$\mathcal{C}=\left\{\left[\frac{m}{2^n},\frac{m+1}{2^n}\right]\,\middle|\, n\in\mathbb{N}_0\wedge 0\leq m<2^n\right\} \tag{3.1}$$

This collection is a sub-base for a topology on the reals. That is, $\mathcal{C}$ is not closed under intersection, but its closure $\mathcal{D}$ is straight-forward enough: one adjoins the dyadic rationals. That is,

$$\mathcal{D}=\left\{c\,\middle|\, c\in\mathcal{C}\vee c=\frac{m}{2^n}\text{ for }n\in\mathbb{N}_0\wedge 0\leq m<2^n\right\} \tag{3.2}$$

Each dyadic rational may be obtained as the intersection of two elements of $\mathcal{C}$. Thus, $\mathcal{D}$ is the basis for a topology $\mathcal{S}$ on the closed unit interval. The topology $\mathcal{S}$ is finer than the natural topology on the reals.

XXX a listing of properties of $\mathcal{S}$. XXX

4. A Dyadic Lattice Topology

Consider the set Ω consisting of all half-infinite strings in the two symbols -1 and $+1$ (or A and B, or L and R, etc.). That is,

$$\Omega=\{\sigma\,|\,\sigma=(\sigma_1,\sigma_2,\sigma_3,\cdots)\text{ with }\sigma_k\in\{-1,+1\}\text{ for }k\in\mathbb{N}\} \tag{4.1}$$

The set Ω can be considered to be the set of all lattice configurations one the half-infinite lattice, where each lattice position may take one of two values. Such lattices have a natural topology, the product topology, whose basis elements are the cylinder sets. A cylinder set $C_{(k,s)}$ is specified by a pair (k,s) where k is a non-negative ineteger, and $s=(s_1,s_2,s_3,\cdots)$ is a *finite*-length string in two letters. The cylinder sets are then the subsets of Ω where the lattice values starting at position k are matched by the string s. That is,

$$C_{(k,s)}=\{\sigma\in\Omega\,|\,\sigma_k=s_1\wedge\sigma_{k+1}=s_2\wedge\cdots\} \tag{4.2}$$

The cylinder sets then provide a basis for a topology on the lattice, in that the intersection of any two cylinder sets is again a cylinder set, or the empty set.

The cylinder sets impose a topology on the unit interval by means of the dyadic or Cantor mapping

$$x(\sigma)=\sum_{k=1}^{\infty}\frac{b_k}{2^k} \tag{4.3}$$

where for convenience we define $b_k=(\sigma_k+1)/2$ so that $b_k\in\{0,1\}$ can be interpreted as a binary digit. This mapping is a homomorphism of the set Ω into the closed unit interval $[0,1]$.

XXX Vet this. The point-set topology on Ω is a finer topology than that on the reals.

This expansion is not unique, because for every dyadic number, there are two inequivalent expansions. For example, $1/2 = 0.1000.. = 0.0111....$

5. Dyadic numbers, Cantor Sets and Measures

The canonical Cantor function may be cconstructed from this expansion, as

$$c_z(x) = (1-z)\sum_{k=1}^{\infty} b_k z^k \tag{5.1}$$

Clearly, for $z = 1/2$, we have $c_{1/2}(x) = x$. For $z = 1/3$, we have essentially the inverse of the classic Cantor function, as show in the figure 5.1. For $1/2 < z < 1$, one obtains a fractal sawtooth function which more than covers the unit interval, as shown in figure 5.2.

We now want to consider how this function maps a measure defined on the domain to the range of the function. For the most part, we'll focus on the case $1/2 < z < 1$.

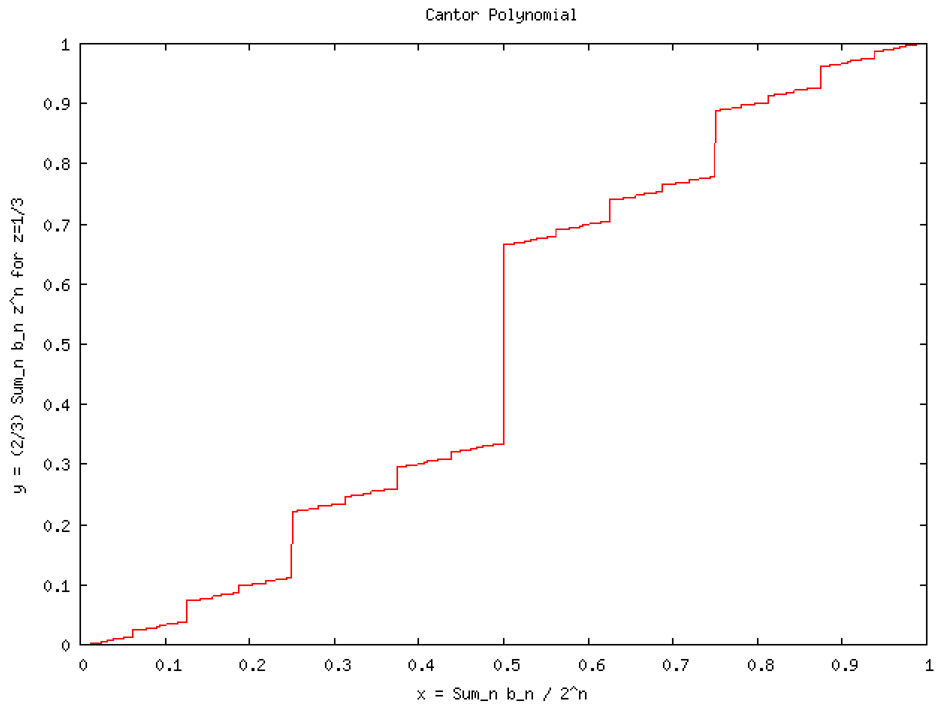
Each of the following figures shows the distribution of the uniform measure on the domain as it is mapped to the codomain. This is done by box-counting over a finite set of intervals. For example, one starts by dividing up the unit interval into 2^nequal-length intervals, whose endpoints lie on integral multiples of 2^{-n}. The function $c_z(x)$, for fixed z, is used to map each of these intervals. To avoid the ambiguity of which binary expansion to use, the map that results in the shortest possible interval is used. Each of these intervals is then histogrammed over a set of N boxes, and then normalized, to generate an image. The normalization is such that the sum of the histogram (the sum total of the boxes) is normalized to one. In principle, each such figure generated will depend not only on the number of dyadic intervals 2^n but also on the number of bars N in the histogram. In paractice, most (but not all) such figures will exhibit a limiting behaviour for large n and N, and care is taken to try to always present a figure in this limiting regime. Exceptions to this are noted.

The most remarkable aspect of this series of figures is how drastically different each of these look, for even small changes in the parameter z. Also remarkable is that as z gets larger and larger, the figures become smooth.

–

References

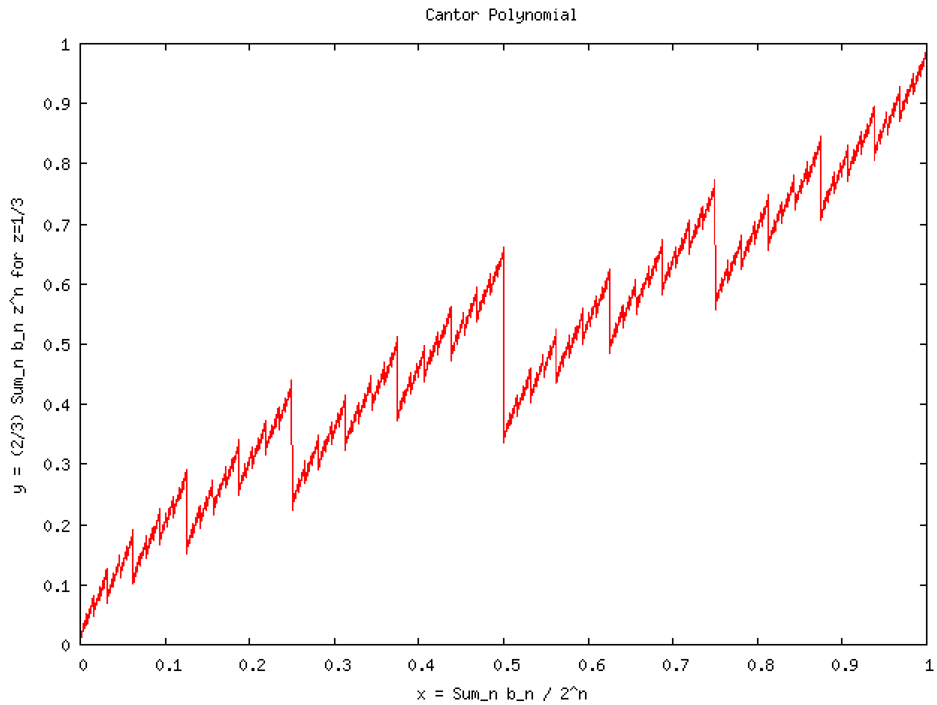
FIGURE 5.1. Cantor Polynomial



This figure shows a graph of the function $c_z(x)$ for a value of $z = 1/3$. This graph provides possibly the easiest visual proof that the cardinality of the Cantor Set is equal to the cardinality of the unit interval[**?**], a result which can sometimes be difficult to visualize. The function $c_{1/3}(x)$ maps the unit interval into the unit interval. It is strictly monotonically increasing, and if $x \neq y$ then $c_{1/3}(x) \neq c_{1/3}(y)$. Thus, this function is one-to-one, and thus the cardinality of the image is equal to the cardinality of the unit interval. But the image of this function is just the canonical Cantor set, constructed by recursively removing the open middle-third interval. The largest "middle-third" corresponds to $x = 1/2$; the next two largest correspond to $x = 1/4$ and $x = 3/4$, and so on: each removed interval corresponds to a rational number, expressed as its binary expansion. It is not hard to see that that the sum of the lengths of the "middle thirds" adds up to one, and thus, the measure of the image of this function is zero. Thus, we have quickly sketched that the cardinality of the Cantor set is that of the real numbers, but the measure of the Cantor Set is zero. It is not hard to see that one gets a similar result using $c_z(x)$ for any value of $0 < z < 1/2$.

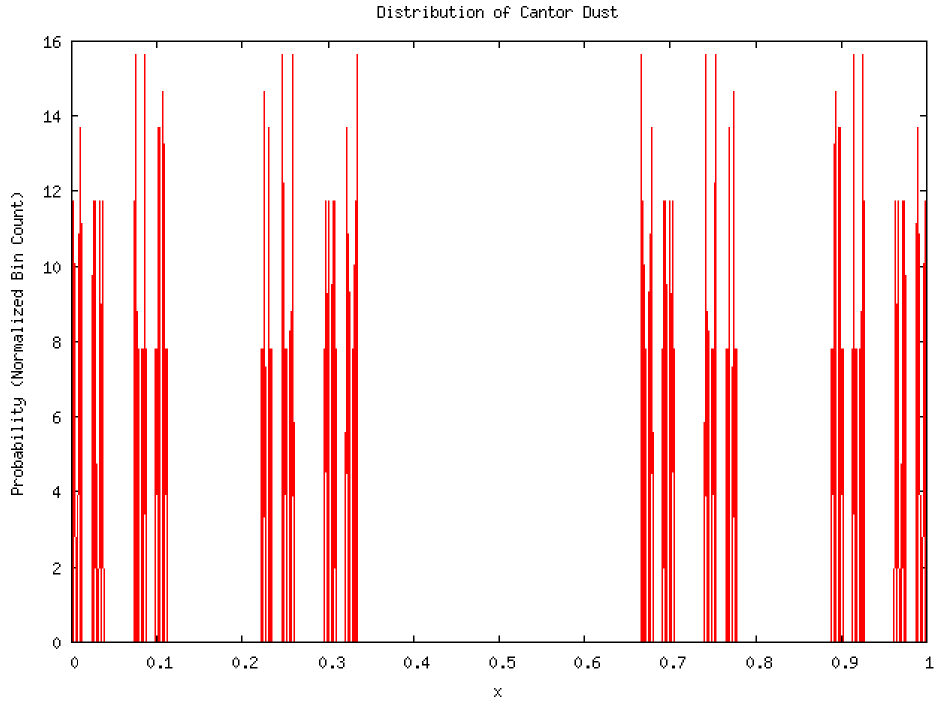
There were only two "tricky" parts to this demonstration. One is the assertion that $c_z(x) \neq c_z(y)$ whenever $x \neq y$, but this can be deduced easily enough by contemplating the definition of $c_z(x)$. The other trick we slid by here was that $c_z(x)$ is somewhat vaguely defined for the rationals: every dyadic rational number $x = p/2^n$ has two inequivalent binary expansions. For example, x=1/2=0.1000...=0.0111... and the first binary expansion yeilds $c_z(1/2) = 1 - z$ while the second yeilds $c_z(1/2) = z$. But this doesn't affect the proof: however we choose to tighten up the definition of $c_z(x)$, it is still one-to-one and monotonic, and its image still has the cardinality of the unit interval.

FIGURE 5.2. Cantor Polynomial



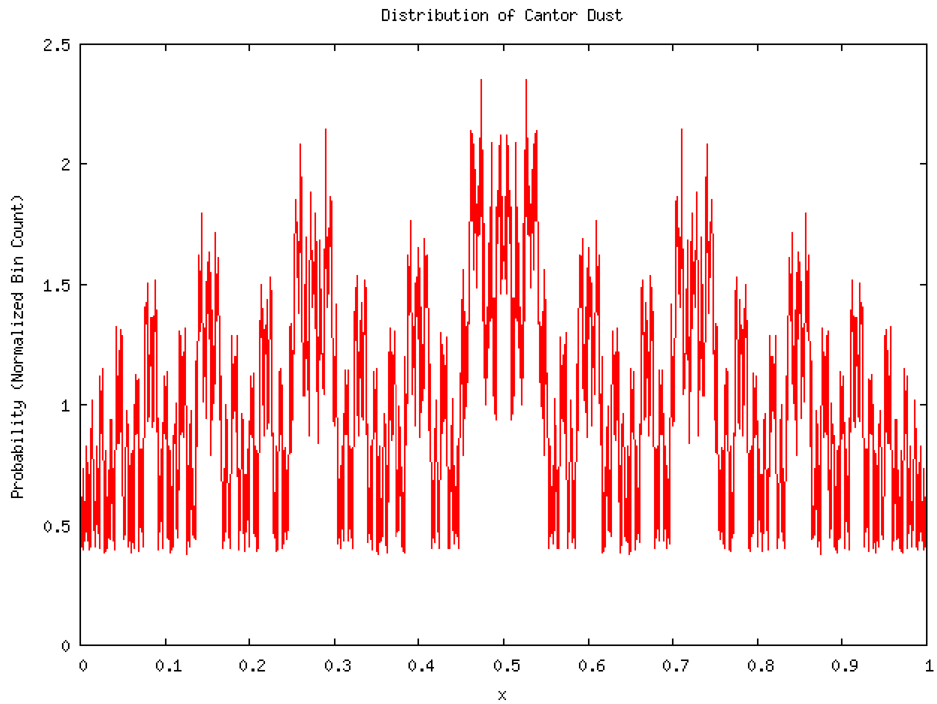
This figure shows a graph of the function $c_z(x)$ for a value of $z = 2/3$.

FIGURE 5.3. Measure of the Cantor Dust



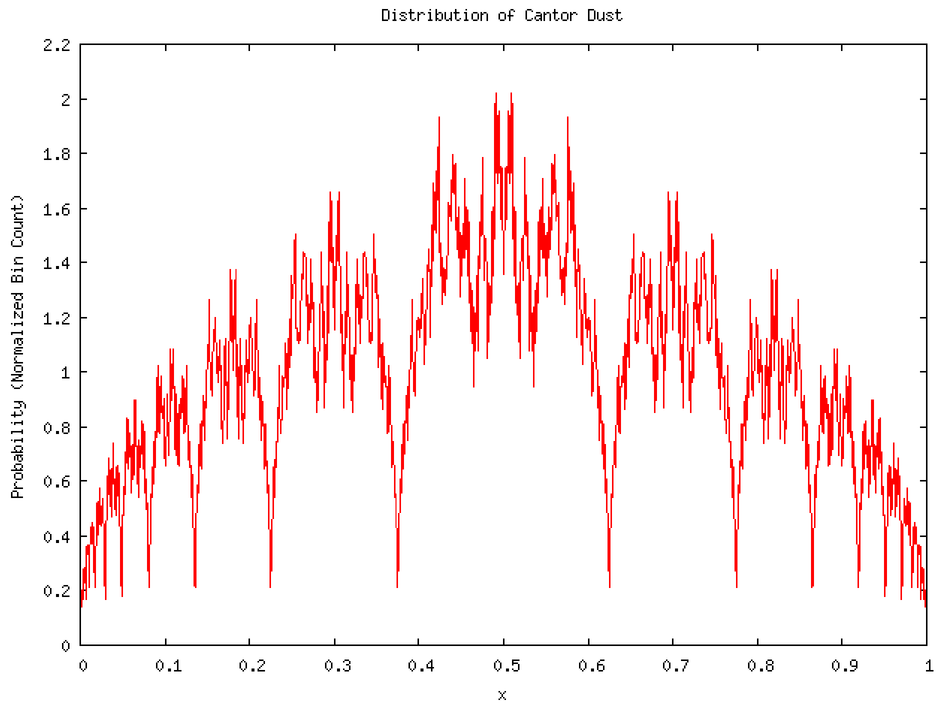
This figure shows the histogram for the image of 2^{22} intervals of the form $[p/2^{22},(p+1)/2^{22}]$, mapped by the function $c_z(x)$, for fixed $z=1/3$, histogrammed into $N=2001$ boxes. This is essentially a diagram of the classic, canonical Cantor set. The height of the bars is a simple function of the number of bars N, given by XXXX give the formula. The un-evenness of the bar heights is a symptom of the fact that $N=2001$ is not divisible by 3, and so one has a jittering effect.

FIGURE 5.4. Measure of the Cantor Dust



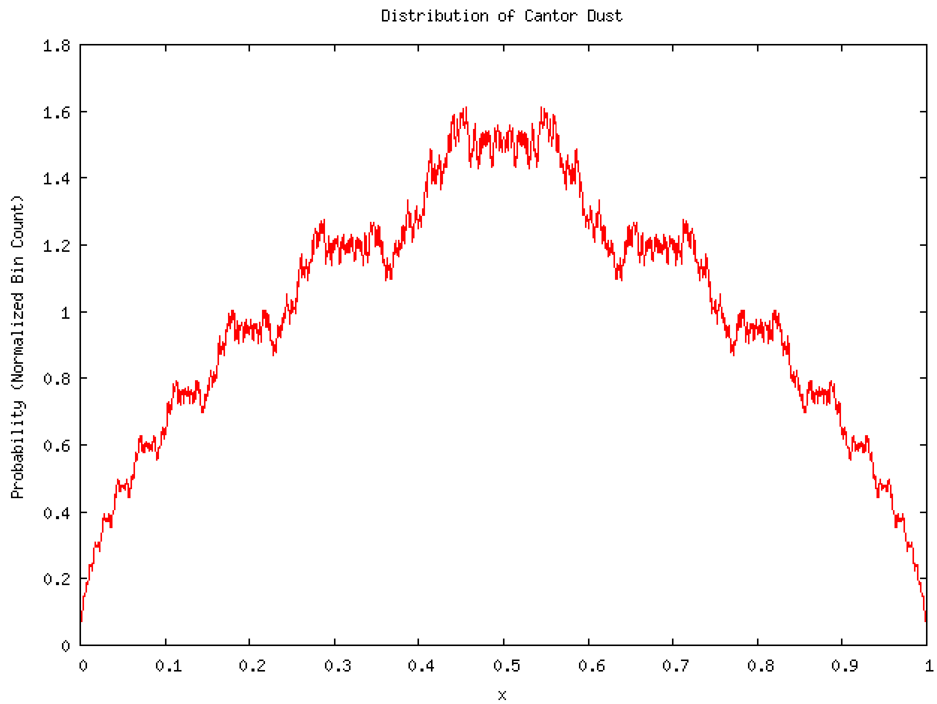
This figure shows the histogram for the image of 2^{22} dyadic intervals of $c_z(x)$ for $z = 0.55$, histogrammed into $N = 2001$ boxes. The shape of this figure, namely, the height of the bars, is essentially independent of N. Going to larger N or graphing a finer number of intervals does not change the general visual appearance of this graph.

FIGURE 5.5. Measure of the Cantor Dust



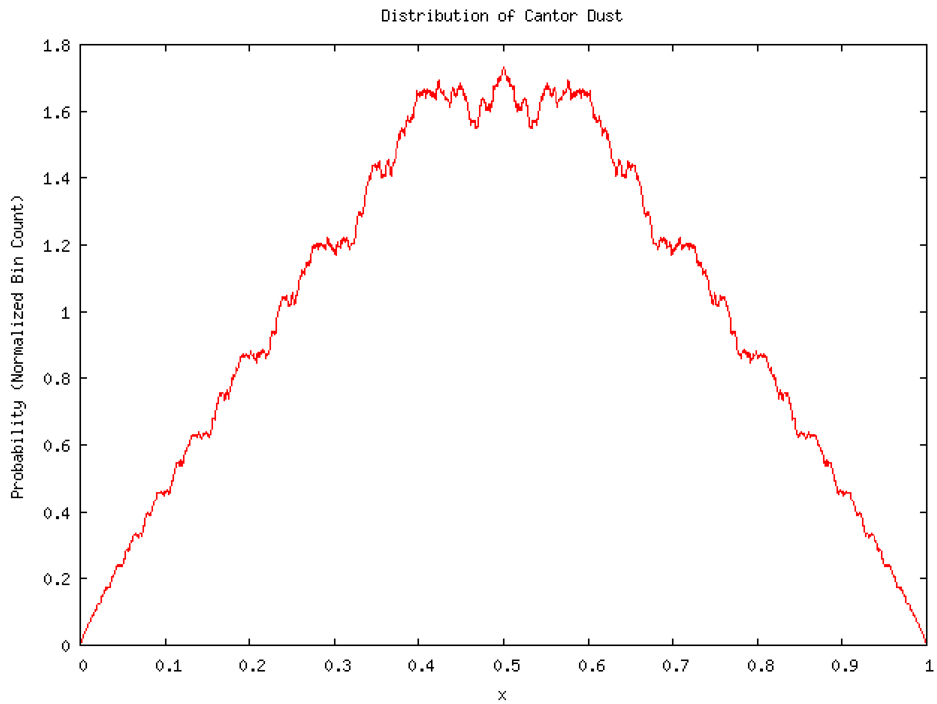
This figure shows the histogram for the image of 2^{22} dyadic intervals of $c_z(x)$ for $z = 0.60$, histogrammed into $N = 2001$ boxes. The shape of this figure, namely, the height of the bars, is essentially independent of N. Going to larger N or graphing a finer number of intervals does not change the general visual appearance of this graph.

FIGURE 5.6. Measure of the Cantor Dust, z=0.63



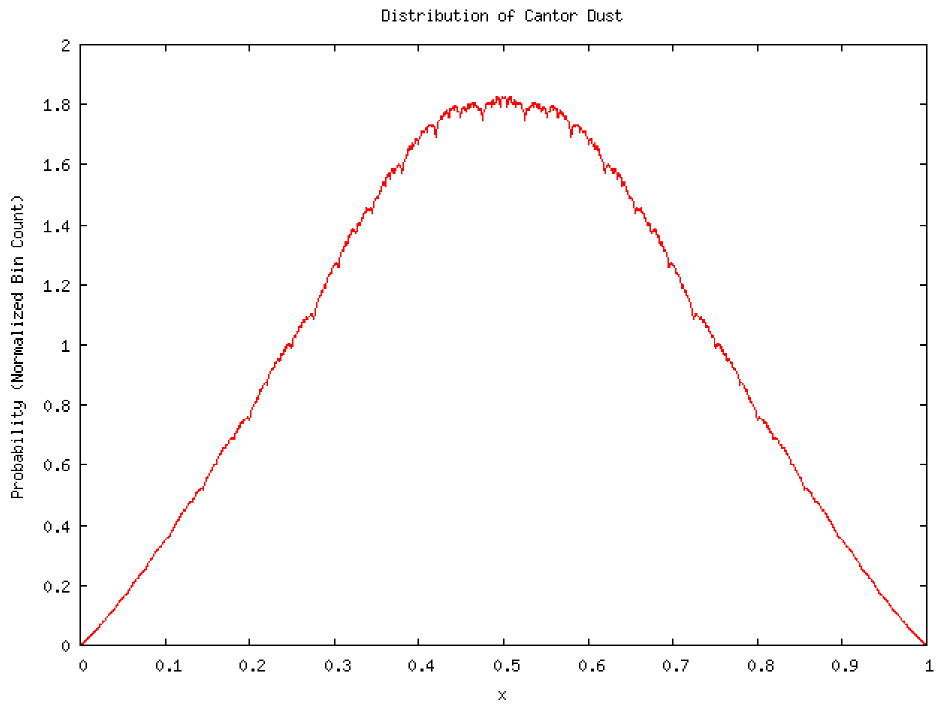
This figure shows the histogram for the image of 2^{22} dyadic intervals of $c_z(x)$ for $z = 0.63$, histogrammed into $N = 2001$ boxes.

FIGURE 5.7. Measure of the Cantor Dust, z=0.69



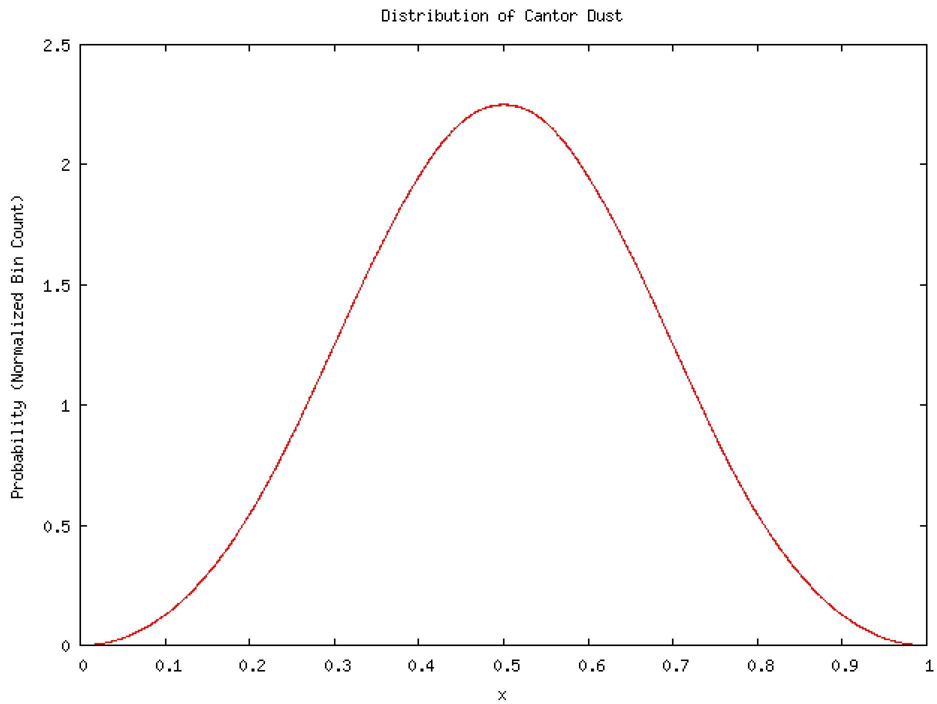
This figure shows the histogram for the image of 2^{22} dyadic intervals of $c_z(x)$ for $z = 0.69$, histogrammed into $N = 2001$ boxes.

FIGURE 5.8. Measure of the Cantor Dust, z=0.725



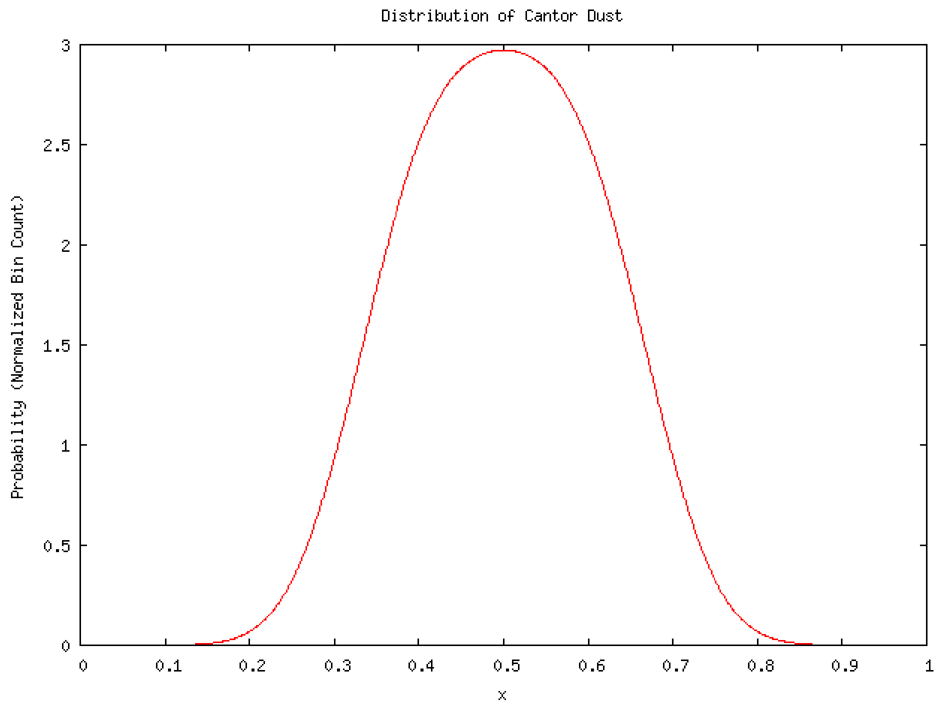
This figure shows the histogram for the image of 2^{22} dyadic intervals of $c_z(x)$ for $z = 0.725$, histogrammed into $N = 2001$ boxes.

FIGURE 5.9. Measure of the Cantor Dust, z=0.8



This figure shows the histogram for the image of 2^{22} dyadic intervals of $c_z(x)$ for $z = 0.8$, histogrammed into $N = 2001$ boxes. Although this figure appears to be smooth, it is not: there are still small-scale discontinuities, of the same general form as in the other figures. Here, the size of these are small enough to give the appearance of smoothness. Note this figure seems sinusouidal; it is not, as the next figure reveals. It also at first seems Gaussian in shape: again, it cannot be, as support vanishes outside of the interval $[0,1]$.

FIGURE 5.10. Measure of the Cantor Dust, z=0.96



This figure shows the histogram for the image of 2^{28} dyadic intervals of $c_z(x)$ for $z = 0.96$, histogrammed into $N = 2001$ boxes. Note that this figure requires consideration of a much larger number of intervals in order to get an accurate limiting shape. Although this figure appears to be smooth, it is not: there are still small-scale discontinuities, of the same general form as in the other figures. It is not Gaussian in shape; it cannot be, as support vanishes outside of the interval $[0,1]$.