Trees, the Cantor set and the irrational numbers

David Gauld*

Abstract

The Cantor set, first described by Smith in about 1874, sits naturally as the limit points of the complete binary tree, showing that the Cantor set is homeomorphic to the topological power $\mathbf{2}^{\mathbb{N}}$, where $\mathbf{2}$ denotes the discrete two-point space. It also illustrates why there is no topological change if we replace $\mathbf{2}$ here by a discrete n-point space for any $n > 1$. What if we replace $\mathbf{2}$ by $\mathbb{N}$? Perhaps surprisingly the answer is that $\mathbb{N}^{\mathbb{N}}$ is homeomorphic to the irrational numbers.

The *Cantor set*, which apparently was first described by H.J.S. Smith in 1874 or 1875 [3, p. 147] eight years before Cantor's independent discovery (and between those two discoveries also described by du Bois-Reymond and Volterra; see [1] for a lot of the background), is usually described as follows. Define $S_0 = [0, 1]$, the closed unit interval. Given S_n, which is a union of 2^n closed intervals each of length 3^{-n}, obtain S_{n+1} from S_n by removing the open middle third of each of the closed intervals making up S_n. Thus $S_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, $S_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$, and so on. The Cantor set, which is here denoted by $\mathbb{S}$ in honour of its discoverer, is the intersection $\cap_{n=0}^{\infty} S_n$. Of course all of the end points of the intervals making up any of the sets S_n will be in $\mathbb{S}$ but $\mathbb{S}$ contains many other points, for example $\frac{1}{4}$; indeed, $\mathbb{S}$ has the cardinality of the set $\mathbb{R}$ of real numbers.

The Cantor set leads to a quick way to construct a space-filling curve. Firstly note that the Cantor set consists precisely of those numbers in $[0, 1]$ which have a ternary representation consisting solely of 0s and 2s. Of course some numbers like $\frac{1}{3}$ have two ternary representations, $0.10000\ldots$ and $0.02222\ldots$ and the latter contains only 0s and 2s as expected. Since we need to use both ternary and binary representations, we will indicate the base by a preceding superscript such as ${}^{3}0.02222\ldots$ etc. Define $f : \mathbb{S} \to [0, 1]$ by $f({}^{3}0.c_1c_2c_3\ldots) = {}^{2}0.\frac{c_1}{2}\frac{c_2}{2}\frac{c_3}{2}\ldots$ whenever each c_i is 0 or 2, so that $\frac{c_i}{2}$ is 0 or 1. Thus, for example, $f(\frac{1}{3}) = f({}^{3}0.0222\ldots) = {}^{2}0.0111\cdots = \frac{1}{2} = f(\frac{2}{3})$. It is easy to see that f is continuous and that $f(\mathbb{S}) = [0, 1]$, so the 0-dimensional Cantor set can be mapped onto the 1-dimensional interval $[0, 1]$. There is little more we need to do to construct a continuous surjection $g : \mathbb{S} \to [0, 1]^2$; for example, we could let

$$g({}^{3}0.c_1c_2c_3c_4\ldots) = \left({}^{2}0.\frac{c_1}{2}\frac{c_3}{2}\frac{c_5}{2}\ldots, {}^{2}0.\frac{c_2}{2}\frac{c_4}{2}\frac{c_6}{2}\ldots\right).$$

Received 22 June 2009; accepted for publication 2 June 2010.
*Department of Mathematics, The University of Auckland, Private Bag 92019, Auckland, New Zealand. E-mail: d.gauld@auckland.ac.nz

The function g easily extends to continuous $\hat{g}\colon [0,1] \to [0,1]^2$ by extending linearly over the intervals which have been cut from $[0,1]$ to get $\mathbb{S}$.

The description above of the Cantor set readily leads to the *Cantor tree*, $T_{\mathbb{S}}$. Recall that a *tree* is a partially ordered set $\langle T, < \rangle$ such that for each $t \in T$ the set of predecessors $\hat{t} = \{s \in T\colon s < t\}$ is well ordered by $<$. Let $T_{\mathbb{S}}$ consist of the union over n of all of the intervals making up S_n and declare $I < J$, for $I, J \in \mathbb{S}$, whenever $I \supset J$. Often trees are represented by joining any element to each of its immediate successors. Figure 1 shows part of the Cantor tree, where the horizontal segments are the elements of the tree and the upward sloping lines show how elements are related. The tree continues upwards with a *level* for each natural number and each vertex has two branches emanating from it.

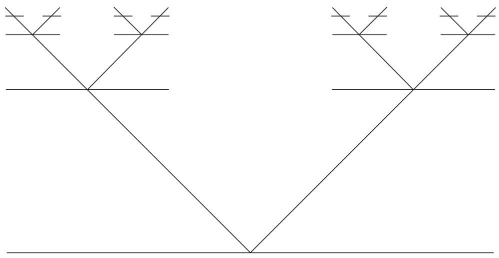

Figure 1. The Cantor tree.

Notice that each *branch* of the Cantor tree corresponds to an element of the Cantor set and vice versa. Each branch may be coded by a sequence consisting of the numbers 0 and 1. In this way we obtain a bijection between the points of the Cantor set and the set of all (infinite) binary sequences. If we denote by **2** the two-point set $\{0,1\}$ then the binary sequences are precisely $\mathbf{2}^{\mathbb{N}}$, the set of all functions from $\mathbb{N}$ to **2**, so the bijection above may be given explicitly as $h : \mathbb{S} \to \mathbf{2}^{\mathbb{N}}$ where $h({}^3 0.c_1c_2c_3\ldots) = \langle \frac{c_1}{2}, \frac{c_2}{2}, \frac{c_3}{2}, \ldots \rangle$. Now **2** has a natural metric, where the distance between 0 and 1 is 1. This gives rise to a metric topology on the product space $\mathbf{2}^{\mathbb{N}}$, the distance from $\langle c_n \rangle$ to $\langle d_n \rangle$ in this metric being $\sum_{n=1}^{\infty} |c_n - d_n|/2^n$. With this topology the bijection $h\colon \mathbb{S} \to \mathbf{2}^{\mathbb{N}}$ becomes a homeomorphism. What makes it work are the gaps imposed between the branches of $T_{\mathbb{S}}$ caused by the removal of the middle thirds of the intervals.

What happens if instead of removing the middle third we remove more intervals, for example two fifths? Actually this situation was also considered by Smith in [3]. Our tree representation still works but now more branches emanate from each vertex. For example, if at each stage we removed the second and fourth fifths then the resulting space would be $\mathbf{3}^{\mathbb{N}}$, which is homeomorphic to $\mathbf{2}^{\mathbb{N}}$, confirming the fact that removing a different number of subintervals in this way does not change the topological type of the space.

Can we obtain a similar interpretation for $\mathbb{N}^{\mathbb{N}}$? Yes, and some might find the answer surprising. Actually we will look at $\mathbb{Z}^{\mathbb{N}}$ but $\mathbb{Z}$ and $\mathbb{N}$ are homeomorphic so $\mathbb{Z}^{\mathbb{N}}$ and $\mathbb{N}^{\mathbb{N}}$ are also homeomorphic. The metric we use on $\mathbb{N}^{\mathbb{N}}$ and $\mathbb{Z}^{\mathbb{N}}$ is similar to that on $\mathbf{2}^{\mathbb{N}}$ except that to ensure convergence of the series we truncate the individual coordinate distances at 1; more precisely the distance from $\langle c_n \rangle$ to $\langle d_n \rangle$ in this metric is $\sum_{n=1}^{\infty} \min\{1, |c_n - d_n|\}/2^n$.

Denote by $\mathbb{P}$ and $\mathbb{Q}$ the sets of irrational and rational numbers respectively. It is well known that $\mathbb{Q}$ is countably infinite; let $\alpha\colon \mathbb{Q} \to \mathbb{N}$ be a bijection. When s is a finite sequence of integers we denote by $s(i)$ the ith term when that exists. Each nonempty *finite* sequence s has a natural immediate successor of the same length. We denote this immediate successor by s^+. If s has length l then $s^+(i) = s(i)$ for each $i < l$ while $s^+(l) = s(l) + 1$. If s has length $l > 1$ then by $s|$ we mean the sequence of length $l - 1$ obtained by removing the last term in s, i.e. removing $s(l)$. For any sequence s of length l by $\langle s, n \rangle$ we mean the sequence of length $l+1$ whose value at $i \leq l$ is that of s and whose value at $l+1$ is n.

By induction on the length of a sequence we construct for each nonempty finite sequence s of integers a rational number q_s satisfying:

(a) $\{q_s \colon s \text{ has length } 1\} = \mathbb{Z}$;
(b) $q_s < q_{s^+}$ for any sequence s;
(c) $q_{s|} < q_s < q_{s|^+}$ for any sequence s of length at least 2;
(d) $\lim_{n\to\infty} q_{\langle s,n\rangle} = q_{s^+}$, and $\lim_{n\to-\infty} q_{\langle s,n\rangle} = q_s$ for any sequence s.

If s has length 1 let $q_s = s(1)$. Then conditions (a) and (b) are satisfied.

Suppose that q_s has been defined for sequences of length less than n so that (a), (b), (c) and (d) are satisfied and let s have length $n > 1$. Both $q_{s|}$ and $q_{s|^+}$ are already defined and $q_{s|} < q_{s|^+}$. Let $\bar{q}$ be the rational $q \in (q_{s|}, q_{s|^+})$ for which $\alpha(q)$ is minimal and set

$$q_s = \begin{cases} \bar{q} + \left(1 - \dfrac{1}{2^{s(n)}}\right)(q_{s|^+} - \bar{q}) & \text{if } s(n) \geq 0, \\ \bar{q} - \left(1 - \dfrac{1}{2^{s(n)}}\right)(\bar{q} - q_{s|}) & \text{if } s(n) \leq 0. \end{cases}$$

We have two observations about the numbers q_s.

1. Every rational number is uniquely represented in the form q_s. Firstly every rational number is represented in this form because by (a)–(d), for any $n \in \mathbb{N}$ the union of all intervals of the form (q_s, q_{s^+}), as s ranges through sequences of length n, will consist of all real numbers except those rationals of the form q_s for s of length at most n. Then we choose a rational $\bar{q}$ in each of these intervals so that $\alpha(\bar{q})$ is minimal. So given a rational $q \in \mathbb{Q}$, we will have $q = q_s$ for some sequence s of length at most $\alpha(q)$. Uniqueness follows because those intervals are mutually disjoint.

2. For any given nonempty finite sequence s the numbers $\{q_t \colon t| = s\}$ subdivide the interval (q_s, q_{s^+}) in much the same way as $\mathbb{Z}$ subdivides $\mathbb{R}$, with the ordering within $\mathbb{R}$ determined by the usual ordering of the last terms of the sequences t for which $t| = s$. See Figure 2.

Now we are in a position to identify $\mathbb{Z}^{\mathbb{N}}$ as a familiar set. Just as the branches of the Cantor tree are coded by infinite binary sequences so are the branches of

Figure 2. Subdividing $\mathbb{R}$.

the tree $\mathbb{Z}^{\mathbb{N}}$ coded by infinite sequences of integers. Then just as the branches of the Cantor tree correspond to the gaps used to make the Cantor set, the branches of the tree $\mathbb{Z}^{\mathbb{N}}$ correspond to gaps which in turn create a subset of $\mathbb{R}$ which is homeomorphic to the metric space $\mathbb{Z}^{\mathbb{N}}$. What is this set? The gaps are precisely the rationals $\mathbb{Q}$. Thus $\mathbb{Z}^{\mathbb{N}}$ is homeomorphic to the irrationals $\mathbb{P}$.

The homeomorphism $\theta\colon \mathbb{Z}^{\mathbb{N}} \to \mathbb{P}$ is not as easy to describe as that between $\mathbf{2}^{\mathbb{N}}$ and $\mathbb{S}$. Here is one way to see how it works. Suppose $s \in \mathbb{Z}^{\mathbb{N}}$ is given. Let $s|n$ denote the finite sequence of length n (a natural number) whose ith term is $s(i)$. Then the sequence of intervals $\langle [q_{s|n}, q_{(s|n)^+}] \rangle$ is a nested sequence with the interior of each interval containing its successor. Because the length of these intervals goes to zero as $n \to \infty$, by the nested sets theorem there is a single point in $\cap_{n=1}^{\infty}[q_{s|n}, q_{(s|n)^+}]$; this is $\theta(s)$. The point $\theta(s)$ cannot be rational because if it were rational then at some stage $\theta(s)$ would have the least image under α amongst those rationals lying in a particular one of those intervals so should have been chosen as $\bar{q}$, thereby excluding $\theta(s)$ from all subsequent intervals. On the other hand, given an irrational r, we can inductively build up a sequence $s \in \mathbb{Z}^{\mathbb{N}}$ by choosing $s(1)$ so that $s(1) < r < s(1) + 1$ and, given $s(i)$ for $i < n$, choose $s(n)$ so that $q_{s|n} < r < q_{(s|n)^+}$.

In [2, Theorem 2], Sierpiński used continued fractions to prove a more general result: every separable 0-dimensional metric space is homeomorphic to a subset of $\mathbb{P}$.

Comment. We can easily modify the description above for certain other subsets of the real numbers. All that is really needed is that the set $\mathbb{Q}$ should be countable and dense. We did use the field structure of $\mathbb{Q}$ to specify the points q_s but we could instead have made do with the density, making sure that the sequences $\langle q_{\langle s,n\rangle} \rangle$ and $\langle q_{\langle s,-n\rangle} \rangle$ converge respectively to q_{s^+} and q_s while the length of any interval created at the nth stage converges to 0 as n goes to ∞. Thus we could take the complement in $\mathbb{R}$ of any countable dense subset. One familiar example is the set of transcendental numbers (though we might observe that the set of algebraic numbers does form a countable field so the method used above may be copied in this case). Another example involves the set of dyadic rationals, that is, those rationals whose denominators are powers of 2; these are countable and dense but do not form a field. The complement of the dyadic rationals is also homeomorphic to $\mathbb{Z}^{\mathbb{N}}$.

References

[1] Hawkins, T. (1970). *Lebesgue's Theory of Integration*. University of Wisconsin Press.
[2] Sierpiński, W. (1921). Sur les ensembles connexes et non connexes. *Fund. Math.* **2,** 81–95.
[3] Smith, H.J.S. (1874). On the Integration of discontinuous functions. *Proc. London Math. Soc.* **s1-6,** 140–153.