

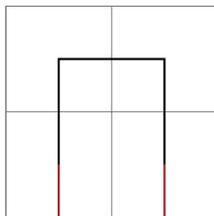
## set.1 Appendix: Hilbert's Space-filling Curves

his:set:hilbertcurve:  
sec

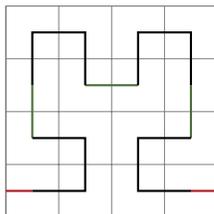
In chapter ??, we mentioned that Cantor's proof that a line and a square have exactly the same number of points (??) prompted Peano to ask whether there might be a space-filling *curve*. He obtained a positive answer in 1890. In this section, we explain (in a hand-wavy way) how to construct Hilbert's space-filling curve (with a tiny tweak).<sup>1</sup>

We must define a function,  $h$ , as the limit of a sequence of functions  $h_1, h_2, h_3, \dots$ . We first describe the construction. Then we show it is space-filling. Then we show it is a curve.

We will take  $h$ 's range to be the unit square,  $S$ . Here is our first approximation to  $h$ , i.e.,  $h_1$ :



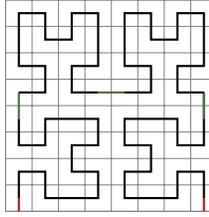
To keep track of things, we have imposed a  $2 \times 2$  grid on the square. We can think of the curve starting in the bottom left quarter, moving to the top left, then to the top right, then finally to the bottom right. Here is the second stage in the construction, i.e.,  $h_2$ :



The different colours will help explain how  $h_2$  was constructed. We first place scaled-down copies of the non-red bit of  $h_1$  into the bottom left, top left, top right, and bottom right of our square (drawn in black). We then connect these four figures (with green lines). Finally, we connect our figure to the boundary of the square (with red lines).

Now to  $h_3$ . Just as  $h_2$  was made from four connected, scaled-down copies of the non-red bit of  $h_1$ , so  $h_3$  is made up of four scaled-down copies of the non-red bit of  $h_2$  (drawn in black), which are then joined together (with green lines) and finally connected to the boundary of the square (with red lines).

<sup>1</sup>For a more rigorous explanation, see [Rose \(2010\)](#). The tweak amounts to the inclusion of the red parts of the curves below. This makes it slightly easier to check that the curve is continuous.



And now we see the general pattern for defining  $h_{n+1}$  from  $h_n$ . At last we define the curve  $h$  itself by considering the point-by-point limit of these successive functions  $h_1, h_2, \dots$ . That is, for each  $x \in S$ :

$$h(x) = \lim_{n \rightarrow \infty} h_n(x)$$

We now show that this curve fills space. When we draw the curve  $h_n$ , we impose a  $2^n \times 2^n$  grid onto  $S$ . By Pythagoras's Theorem, the diagonal of each grid-location is of length:

$$\sqrt{(1/2^n)^2 + (1/2^n)^2} = 2^{(\frac{1}{2}-n)}$$

and evidently  $h_n$  passes through every grid-location. So each point in  $S$  is at most  $2^{(\frac{1}{2}-n)}$  distance away from some point on  $h_n$ . Now,  $h$  is defined as the limit of the functions  $h_1, h_2, h_3, \dots$ . So the maximum distance of any point from  $h$  is given by:

$$\lim_{n \rightarrow \infty} 2^{(\frac{1}{2}-n)} = 0.$$

That is: every point in  $S$  is 0 distance from  $h$ . In other words, every point of  $S$  lies on the curve. So  $h$  fills space!

It remains to show that  $h$  is, indeed, a curve. To show this, we must define the notion. The modern definition builds on one given by Jordan in 1887 (i.e., only a few years before the first space-filling curve was provided):

**Definition set.1.** A curve is a continuous map from  $L$  to  $\mathbb{R}^2$ .

This is fairly intuitive: a curve is, intuitively, a “smooth” map which takes a canonical line onto the plane  $\mathbb{R}^2$ . Our function,  $h$ , is indeed a map from  $L$  to  $\mathbb{R}^2$ . So, we just need to show that  $h$  is continuous. We defined continuity in ?? using  $\varepsilon/\delta$  notation. In the vernacular, we want to establish the following: *If you specify a point  $p$  in  $S$ , together with any desired level of precision  $\varepsilon$ , we can find an open section of  $L$  such that, given any  $x$  in that open section,  $h(x)$  is within  $\varepsilon$  of  $p$ .*

So: assume that you have specified  $p$  and  $\varepsilon$ . This is, in effect, to draw a circle with centre  $p$  and radius  $\varepsilon$  on  $S$ . (The circle might spill off the edge of  $S$ , but that doesn't matter.) Now, recall that, when describing the function  $h_n$ , we drew a  $2^n \times 2^n$  grid upon  $S$ . It is obvious that, no matter how small  $\varepsilon$  is, there is some  $n$  such that some individual grid-location of the  $2^n \times 2^n$  grid on  $S$  lies wholly within the circle with centre  $p$  and radius  $\varepsilon$ .

So, take that  $n$ , and let  $I$  be the largest open part of  $L$  which  $h_n$  maps wholly into the relevant grid location. (It is clear that  $(a, b)$  exists, since we already noted that  $h_n$  passes through every grid-location in the  $2^n \times 2^n$  grid.) It now suffices to show to show that, whenever  $x \in I$  the point  $h(x)$  lies in that same grid-location. And to do *this*, it suffices to show that  $h_m(x)$  lies in that same grid location, for any  $m > n$ . But this is obvious. If we consider what happens with  $h_m$  for  $m > n$ , we see that exactly the “same part” of the unit interval is mapped into the same grid-location; we just map it into that region in an increasingly stretched-out, wiggly fashion.

## Photo Credits

## Bibliography

Rose, Nicholas J. 2010. Hilbert-type space-filling curves. URL <https://web.archive.org/web/20151010184939/http://www4.ncsu.edu/~njrose/pdfFiles/HilbertCurve.pdf>.