THE REALS ARE UNCOUNTABLE ... AND MORE ...

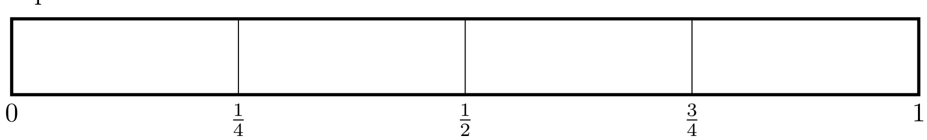
THOMAS SHORES

First of all, what does that even mean? Just what is a "countable" set of objects? It's simply a set whose elements can be enumerated or listed, one at a time, with each element finitely far down the list. For example, the set of primary colors of the RGB color model is countable since we can list them as "red, green, blue." So is your pocket change. These sets are also finite. But the infinite set of all rational numbers between 0 and 1 is countable as well. Here's a start on the list:

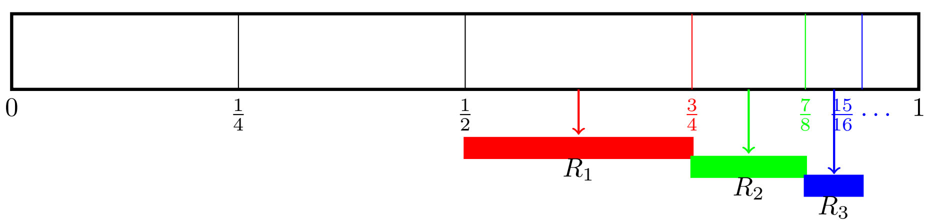
$$\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \ldots.$$

Of course, I can't write down the whole list since it is infinite, hence the "...". I could write down a formula, but you can see the pattern, and though numbers will be repeated, every fraction between 0 and 1 is in the list, so that set is clearly countable. So how is it that the set of all real numbers in the unit interval is not countable? Could it be that there is some way to see that it is countable, but we're just not being clever enough? We'll see that the answer is "no", the set of real numbers in the unit interval (never mind all of them!) is really uncountable.

Let's first lay down a ruler giving us an "unit interval" of length one, and mark a few points on it:

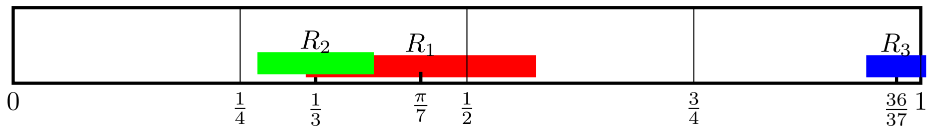


OK, Now we're going to create a bunch of copies of subintervals of the unit interval as follows: Copy a section starting at 1/2 and getting halfway to 1, so ending at 3/4. Call it R_1. Copy a section starting at 3/4 and going halfway to 1, so ending at 7/8 and call it R_2. Next, copy a section starting at 7/8 and going halfway to 1, so ending at 15/16. Call it R_3. Keep going ...



Clearly, we can carry this process on forever, getting an infinite sequence of copies $R_1, R_2, R_3, \ldots$ of subintervals, each of positive width. So what is the total width of all these copies? Well, the subintervals from which they're copied are laying side by side, stretching from 1/2 to 1, so the total width is clearly 1/2.

OK, so now let's suppose that we have a countable sequence of real numbers between 0 and 1, say $r_1, r_2, r_3, \ldots$. Let's cover these numbers with our copies of the subintervals, placing R_1 over number r_1 in such a way that r_1 is at the midpoint of the copy. Do likewise for r_2 with the second copy R_2, for r_3 with the third copy R_3, and so forth. Here's a picture of what our covering would look like after three steps with, say, a sequence that started with $r_1 = \pi/7$, $r_2 = 1/3$, $r_3 = 36/37$.



In general, some of these copies will overlap, others will extend past the unit interval, but that doesn't matter. The upshot is that we can cover every number $r_1, r_2, r_3, \ldots$ with a copy of an interval of positive width and such that the total width of all these copies laying on the unit interval is no more than 1/2.

Conclusion? The sequence of numbers $r_1, r_2, r_3, \ldots$ cannot possibly be the set of all real numbers on the unit interval, because in that case the intervals we constructed would cover the whole unit interval and hence have total width at least one. Therefore, the real numbers on the unit interval are an uncountable set, ie, they cannot be listed in the form $r_1, r_2, r_3, \ldots$.

But wait! There's more!! Discard the endpoints of each of the intervals we constructed. Each r_n is still in the interior of its interval. So what we are covering the sequence of r_n's with is a collection of "open intervals", and their union is called an "open set" in topology. Now observe: There is no reason that we had to start constructing our subintervals at 1/2. We could have started at 3/4 or 7/8 or 15/16, etc., yielding an infinite collection of open intervals of positive lengths such that their total length is 1/4 or 1/8 or 1/16, or other arbitrarily small sizes. You get the idea: Any countable sequence of real numbers can be covered by an open set of arbitrarily small total length, ie, size. This non-negative number is termed the total "measure" of the set in the language of measure theory. Here, it is a basic rule that if one "measurable" set A is a subset of another "measurable" set B, then the measure of A is less than or equal to that of B. It follows that any sequence of real numbers is contained in sets of arbitrarily small measure, hence itself must be of measure zero. This is a bit trickier than it looks: Sure, in a sequence of numbers like $0, 1, 2, \ldots$ the points are isolated and obviously of "width" zero, so the measure the whole set is clearly zero. But consider the rational numbers in the unit interval: They're countable, so have measure zero, yet they are very "thick" in the interval in the sense that any subinterval of positive width must contain a rational, so their zero measure is a bit less intuitive.

So there you have it — your introduction to an area of abstract mathematics called "measure theory".