Math 108
Irrational, transcendental, indescribable

We have already seen that, since there are countably many rational numbers and uncountably many real numbers:

Corollary. *There are uncountably many irrational numbers.* □

In other words, there are infinitely more irrational numbers than rational numbers. Put still another way, in a sense that can be made precise in real analysis, almost all real numbers are irrational.

The purpose of this handout is to explain that while the above corollary is somewhat mind-blowing, the real numbers are *even more bizarre* than the above results might indicate. To explain how and why, we begin by recalling the following theorem.

Theorem. *If A_n is a countable set for each $n \in \mathbf{Z}^+$, then $\bigcup_{n \in \mathbf{Z}^+} A_n$ is countable.* □

Definition. We say that a real number is *algebraic* if it is a root of some polynomial $p(x)$ with rational coefficients. A real number that is not algebraic is said to be *transcendental*.

For example, any rational number a is algebraic, because a is a root of the rational polynomial $x - a$; $\sqrt{2}$ is algebraic because it is a root of $x^2 - 2$; and so on. It can also be shown (though this is far from obvious) that sums, differences, products, and quotients of nonzero algebraic numbers are also algebraic. On the other hand, while e, π, and similar numbers have been known to be transcendental for over a century, as of right now (2016), for most specific real numbers r, it seems to be very difficult to prove that r is transcendental.

Nevertheless, we have the following theorem.

Theorem. *There are countably many algebraic numbers. Consequently, there are uncountably many transcendental numbers, i.e., most real numbers are transcendental.*

Proof. Let P_k be the set of all rational polynomials of degree k. We first observe that a rational polynomial $a_k x^k + \cdots + a_0$ is defined uniquely by its $k + 1$ coefficients $a_k, \ldots, a_0$ ($a_k \neq 0$). It follows that P_k is equivalent to the Cartesian product $(\mathbf{Q} \setminus \{0\}) \times \mathbf{Q} \times \cdots \times \mathbf{Q}$, where there are k copies of $\mathbf{Q}$ (details of the bijection left as an exercise). Therefore, since a finite Cartesian product of countable sets is countable, P_k is countable.

Next, for each $k \in \mathbf{Z}^+$, since P_k is countable, let $P_k = \{p_{k,i}(x) \mid i \in \mathbf{Z}^+\}$, let $R_{k,i}$ be the set of all roots of the polynomial $p_{k,i}(x)$, and let $R_k = \bigcup_{i \in \mathbf{Z}^+} R_{k,i}$ be the set of all roots of any polynomial of degree k. Since any polynomial of degree k has at most k roots, each $R_{k,i}$ is finite, which means that R_k is a countable union of countable sets, and therefore countable.

Finally, since the set of algebraic numbers A is precisely the union of the roots of some rational polynomial of some degree, we see that $A = \bigcup_{k \in \mathbf{Z}^+} R_k$. Therefore, A is a countable union of countable sets, which means that A is countable. □

So now, consider the following definitions, throughout which we fix a finite **alphabet** A. This is a very mild restriction; for example, we can take A to be the union of all characters used in any language on Earth, human, computer, or otherwise. (In fact, we can even take A to be countable, and the following results still hold.)

Definition. We define the *descriptions of length k* to be $A^k = A \times \cdots \times A$, where there are k copies of A, and we define the set of all *descriptions D* to be $D = \bigcup_{k \in \mathbf{Z}^+} A^k$. We define an *interpretation I* of D to be a function $I : D \to \mathbf{R}$.

In other words, the descriptions of length k are the strings of length k in the alphabet A, and an interpretation of those strings assigns a real number to each possible description. Of course, this is on the optimistic side of things; in reality, most descriptions (such as the random typing "lksdhjfdslh" or the text of *Hamlet* in Hebrew) will not specify any particular real number. If this makes you uneasy, assign some arbitrary real number, like 0, to any non-numeric or nonsensical description.

Definition. Given a fixed interpretation I, we define the set of *describable* numbers to be the range of I (a subset of $\mathbf{R}$). Any number that is not describable is said to be *indescribable*.

For example, taking A to be the "universal alphabet" described above, and taking I to be the usual interpretation of strings in that alphabet, any rational number is describable as "p/q", any algebraic number is describable (e.g., "the root of $x^5 + 16x - 3$ such that..."), any number given by any finite formula is describable (e.g., $e = \sum_{n=0}^{\infty}(1/n!)$), any number that is the output of any finite computer program is describable, and so on.

Nevertheless, we still have the following theorem.

Theorem. *For a fixed alphabet A and an interpretation I, there are only countably many describable numbers. Consequently, there are uncountably many indescribable numbers, i.e., most real numbers are indescribable.*

Proof. Since our fixed language A is countable, the finite Cartesian product A^k is countable, and the countable union of countable sets $D = \bigcup_{k \in \mathbf{Z}^+} A^k$ is countable. Therefore, the range of I, which is

$$I(D) = \{I(d) \mid d \in D\} = \bigcup_{d \in D} \{I(d)\},$$

is a countable union of countable sets. It follows that the set of describable numbers, which is the range of I, is countable. $\qquad\square$

In other words, we have shown that no matter what we do, no matter how clever we are, and no matter how hard we try, given any real number r, we will almost certainly not be able to describe r individually. Who would have guessed that the number line, something you probably first saw back in grade school, is so fundamentally unknowable!