To XXXX, a student entering Junior High School
in response to a letter sent to my Math. Dept.
from not very far away not very long ago.

Dear XXXX,

This responds to your letter of (*date*) expressing doubts about your teacher's assertion that blades of grass, pencils, and grains of sugar could be infinite in number. Your doubts are justified. Perhaps your teacher learned about infinity from the same text as did my younger son's teacher; she told him the leaves on a tree could be infinite because "some leaves might fall off and others spring from bud before all could be counted." And she said the grains of sand on a beach were "surely more than anyone could count" and therefore infinite. Not so.

Over 22 centuries ago, Archimedes wrote a letter to his patron Gelon, elder son of King Hieron of Syracuse (then a Greek colony on Sicily), praising the adoption of Ionian Greek numeration in place of Attic. Called Ψαμμιτεσ (*the Sand Reckoner*), this letter uttered a number Archimedes proved bigger than the number of grains of fine sand sufficient to fill the universe as it was thought to be then. It was smaller then than we believe it to be now, but still far bigger than the Earth, so the number of grains of sand on an Earthly beach must be finite, far less than a number like 10^{51} written in Ionian Greek. The same goes for grains of sugar.

You and your teacher could do for sugar what Archimedes did for sand. Buy a bag of coarse granulated sugar, not powdered "Confectioners' Sugar" because its grains are so microscopic that counting them would take longer than educational value can justify. Coarse grains can be counted easily with the aid of a magnifying glass. Count out some number n of them big enough to sample the grains' different sizes adequately; one or two thousand grains should be enough, and will take less than an hour to count if you don't blow them away. Make sure that your count, if slightly wrong, errs on the high side so that your little pile of sugar will have at most a known number n of grains. Next, take your pile to a chemist or pharmacist with an accurate weigh-scale to find out how much n grains of sugar weigh; if slightly wrong, make sure the alleged weight w errs on the low side, so your pile will weigh at least w grams, say. The bag's weight W is printed on it; for instance, a five-pound bag weighs less than 2300 grams. If you don't trust what's printed on the bag, your obliging chemist or pharmacist may weigh it for you to find a number W slightly bigger than the bag's weight. Then you will know the bag can hold at most $N = n \cdot W/w$ grains; this N is finite, and therefore so is the exact number of grains in the bag. And now you know it to be so without having to go to the trouble of actually counting the grains.

What makes a set finite is the existence of an integer provably bigger than every integer needed to label each member of that set with a positive integer different from every other member's label.

Such a set's members may well be too numerous for anyone to label or count in a human's life-span, even too numerous to be counted by the fastest electronic computer before it runs out of electric power; yet all members can be counted in principle so long as there are finitely many.

An *infinite* set is just a set that is not finite. (Whether it is *uncountable* too remains to be seen.)

For example, the integers constitute an infinite set because no biggest integer exists.

The *Primes*, positive integers each divisible by no other than itself and 1 , constitute an infinite set because no biggest prime exists; Euclid of Alexandria proved this about 23 centuries ago.

On the other hand, the number of atoms in the visible universe (within a vast sphere outside which stars recede from us so fast that their light will never reach us) must be finite though the number is ever changing and too gargantuan for you or me to count exactly.

Depending on its context, "infinite" is a word standing for one of many longer explanations.

Outside mathematics, it has been invoked with words like "eternal" to inspire awe, often to distinguish mortal Man and his works from those of a divine Creator unhindered by limitations of time and space. Also there is poetic and dramatic licence: Shakespeare has Juliet tell Romeo

"My bounty is as boundless as the sea,
My love as deep; the more I give to thee,
The more I have, for both are infinite."

An exaggeration, perhaps, but no worse than when someone says "On a clear night away from city lights, an infinite number of stars twinkle in the sky."

Strictly speaking, mathematically, there is no "infinite number", not like other numbers.

If ∞ were merely a number, $\infty+1$ would be bigger, but it is not. Since ∞ is not zero, $\infty+\infty$ would be bigger than ∞ , but it is not, nor is $\infty\cdot\infty$. Can we suppose ∞ = (anything nonzero)/0 ? If so, $-\infty = -3/0 = 3/(-0) = 3/(+0) = +\infty$. As you see, infinity could participate in a kind of arithmetic so unlike ordinary arithmetic that it should no more persuade us that ∞ is a number than should the biblical injunction "Go forth and multiply" persuade us that people are numbers.

"An infinite number" is a misnomer arising from loose use of language. Linguistic details that usually don't matter in ordinary discourse matter a lot to mathematicians and contract lawyers, who must seem fussy in the practice of their professions. For instance, someone who says "I saw a number of shoes scattered on the floor" is presumed to have seen some shoes, not none; "a number" cannot mean zero in that context. Does this imply that zero is not a number? Almost everybody in Europe used to think so until the fifteenth century. Nowadays a fussy speaker would have to say "I saw a positive number of shoes scattered on the floor." More than 1 ? Until the sixteenth century, many scholars continued to believe that 1 was not a number; they were misled by the plural form of the word "units" in Euclid's definition: "A number is an aggregate composed of units." He was not fussy enough; he should have said "A (positive) number is either the unit 1 or an aggregate composed of units" as if a number were a length measured in some small unit like a millimeter. Instead he sowed confusion among mediæval scholars trying to extract wisdom from ancient manuscripts as if they were all written like the Bible.

The meaning of a word can depend considerably upon its context. We distinguish verb "peer" from noun "peer" by its grammatical context, and noun "peer" from "pier" by its semantic context, though all three words sound alike and we know many kinds of pier. Nowadays we know many kinds of number too. Besides the *natural* numbers (positive integers) 1, 2, 3, …, there are 0 and negative integers –1, –2, –3, … . There are *rational* numbers (fractions) like 5930/47619 = 0.12453012453012453… (repeated forever). And there are *irrational* numbers like $\sqrt{2}$ = 1.41421356… and π = 3.14159265358979… . All these are called "real" numbers to distinguish them from *imaginary* numbers like $\sqrt{-4}$; and sums of real and imaginary numbers are *complex* numbers that can be added, subtracted, multiplied and divided very much like real numbers. Evidently the noun "number" has accumulated a number of meanings, and so has the verb (not to mention the adjective), depending upon context. But, except figuratively, there is no infinite number, not like these other numbers; they are all finite.

Mathematicians came to our current understandings of infinity late in the nineteenth century. Now "∞" stands for one of many lengthier explanations, depending upon context, always about something beyond finite. For example, mathematicians know that integers n have this property:

" $1/n \to 0$ as $n \to \infty$ ", pronounced "one-over-n tends to zero as n tends to infinity."

However, n never gets to ∞ and $1/n$ never gets to 0 . Stated without ∞ , that quoted assertion means this:

"Every tiny positive real number, no matter how tiny, is
bigger than $1/n$ for all but finitely many positive integers n ."

Generally, "∞" seems preferable to the lengthier circumlocutions we would need without it.

"Infinity" can stand for "somewhere beyond any finite place", belying a toy spaceman's cry

"To Infinity, and beyond!"

Mathematicians use infinity in this sense to imagine boundaries for regions that seem at first to have no boundaries. For instance, the real number line looks like an infinitely long tape-measure with its numbers hidden, something like this: ···——————··· It has no ends because there is no biggest number. However, we can attach ends called "$-\infty$" and "$+\infty$" to the real number line like this: •—————• . Then every real number x lies between the ends; $-\infty < x < +\infty$.

If that were the only way to *close* the line, everybody would do it and we would teach it in all grade-schools. But there is another way: Wrap the real number line around a circle and attach one point at infinity to its end(s) so that it looks like the letter "Q" with infinity at the bottom. This makes $-\infty = +\infty$, which may seem unreasonable at first but does let rational arithmetic with ∞ work better; now $-\infty = -3/0 = 3/(-0) = 3/(+0) = +\infty$ without a contradiction, though neither $\infty+\infty$ nor $\infty-\infty$ are numbers. Although putting just one end at infinity on the real number line runs counter to the ways most of us construe numbers, it was understood briefly by so eminently a non-mathematician as Winston Spencer Churchill, Prime Minister of Great Britain during its heroic stand alone against Nazi Germany early in World War II. Here is how he is quoted:

"I had a feeling once about Mathematics, that I saw it all —
Depth beyond depth was revealed to me — the Byss and the Abyss.
I saw, as one might see the transit of Venus — or even the Lord Mayor's show,
a quantity passing through infinity and changing its sign from plus to minus.
I saw exactly how it happened and why the tergiversation was inevitable;
and how the one step involved all the others. It was like politics.
But it was after dinner and I let it go!"

… … …

Like the line, the plane can be closed at infinity in more than one way. Imagine standing in a flat prairie with nothing to obstruct your view of the horizon; it would look like a circle around you. In fact, it is (very nearly) a circle because the earth is not flat; what you see of the prairie is (very nearly) a circular cap cut from the (very nearly) spherical surface of the earth by a plane.

How would your view differ if the prairie were perfectly flat? Instead of seeing a distant wagon rise above the horizon as it approached, you would see the wagon grow from a tiny speck always in view. The horizon around you would still look like a circle, namely *the circle at infinity.*

Next imagine a snapshot of the wagon taken while it is still far away. In the photograph, the horizon appears as a straight line, namely *the line at infinity.* If the wagon maintains a straight course as it approaches, its wheels leave ruts that are parallel straight lines; and all straight lines parallel to these ruts in the flat prairie appear on the photograph to intersect at the same point on the line at infinity. On this line each point is where its own family of parallel straight lines in the prairie appear to come together in the photograph.

The plane can also be closed at just one *point at infinity.* In our mind's eye, replace "plane" by the surface of a sphere like the earth, and replace "straight lines" by circles on the sphere through the North pole. This is the point at infinity, the one point where parallel straight lines in the plane intersect. Our identification of circles with lines is called "Stereographic Projection" because it can be accomplished by putting a lantern at the North pole and a large flat screen tangent to the sphere at its South pole; then circles through the North pole cast shadows upon the screen that are straight lines. The lantern has no shadow on the screen. This projection helps complex arithmetic work better with one complex ∞ much as the circular closure of the real line helps real arithmetic work better with one real ∞ . The details are a story for another day.

Over time "infinite" has come to mean many things, but always "beyond everything finite." At first sight, this characterization of "infinite" seems to require a prior characterization of "finite". There is a tricky way to distinguish "infinite" from "finite" without defining either of them first:

Recall using positive integers to label all members of a finite set in such a way that each member's label is different from any other member's, and some integer is known bigger than every label. Instead of integers, the members of some other set can serve as labels; and then the distinction between label and thing labelled can be discarded. Thus, we say a *bijection* exists between two sets S and T just when all their elements are paired, one element of each set per pair, and no element appears in more than one pair. For example, in the absence of polygamy, polyandry, same-sex marriages and divorces, marriage establishes a bijection between the set of all husbands and the set of all wives. (Death diminishes both sets, turning a wife into a widow or a husband into a widower if either survives.) Another example is the bijection between the positive integers and the even positive integers; the pairs are $\{n, 2n\}$ for $n = 1, 2, 3, \dots$. A bijection exists between any set of half a dozen oranges and the set $[1, 2, 3, 4, 5, 6]$.

We say T is a *proper subset* of set S just when set T consists of some but not all nor none of the members of S . Married mothers constitute a proper subset of the set of wives, some of whom are childless. The even integers constitute a proper subset of the set of integers. Except when S has very few elements, it has lots of proper subsets.

Now, a set S is *infinite* just when a bijection exists between S and some proper subset of S ; otherwise, when no such subset and bijection exist, S is *finite*. Consequently every set is either finite or infinite, and in principle we can distinguish them without first knowing how to count. This is why the definition "An infinite set is a set whose members cannot be counted" would be misleading even if it were not ambiguous; it should not be taught to teachers. The definitions of *finite* and *infinite* on page 1 are safe because they are deduced scrupulously from bijections.

Since "infinite" appears with different meanings in so many contexts, most people may feel tempted to use the word carelessly. Resist that temptation, and not merely because choosing the right word when you can is the morally right thing to do. Mark Twain put it this way:

"The difference between the right word and the almost right word
is the difference between lightning and the lightning bug."

Besides, words chosen carefully often reveal insights obscured by careless descriptions. For example, the word "incomputable" usually means "too big to compute" and is treated in my dictionary as a near-synonym for "incalculable", "countless", "immeasurable", "inestimable", "infinite", "innumerable" and "measureless". But "uncomputable" is not in my dictionary.

What are numbers for if not to be computed? And yet, there are *uncomputable* real numbers!

We can compute any number that we are told or figure out how to compute. The instructions for computing a number constitute a *program*; it may be executed by either a human or an electronic computer if presented in a suitable language. For example, here is a program to compute $\sqrt{2}$:

- Start with two positive integers m and n whose ratio m/n approximates $\sqrt{2}$, perhaps poorly.
- Replace m and n respectively by $M := m^2 + 2n^2$ and $N := 2mn$;
  now M/N approximates $\sqrt{2}$ rather better than m/n did.
- Repeat the foregoing replacement process until $\sqrt{2}$ is approximated as closely as you like.

Here is a sample of the results from the program's execution:

m = 3	n = 2	m/n = 1.*5*
m = 17	n = 12	m/n = 1.41*66*…
m = 577	n = 408	m/n = 1.41421*57*…
m = 665857	n = 470832	m/n = 1.41421356237*47*…
…	…	…
		$\sqrt{2}$ = 1.414213562373095…

In general a number x is deemed computable just when a program exists that will deliver at least a few digits of x soon, and as many more digits of x as we like if we follow the program long enough. The program's text must have finite length; otherwise no way would exist to ascertain whether all relevant instructions in the program were being followed since, at any moment, most of the program would remain to be seen. How many programs each finitely long can exist?

Right now only finitely many programs have ever been written. However, at least in principle, we can imagine a collection of infinitely many programs each able to compute a different number. Therefore infinitely many numbers are computable. Because each such program's text is finitely long, each program can be labelled with a positive integer without using any label more than once. This amounts to labelling all the computable numbers each with its own integer. But the real numbers cannot be labelled that way; over a century ago the set of real numbers was proved *uncountably infinite* in the sense that no bijection can exist between the real numbers and the integers even though both sets are infinite. Therefore infinitely many more real numbers exist than programs can exist to compute them. Since we can't compute these uncomputable numbers, we can't know which they are although they are by far in the majority. The thought is humbling.

As you can see now, a careful distinction between two kinds of infinite sets, *countable* like the integers and *uncountable* like the real numbers, has revealed an interesting limitation upon our computers, perhaps our minds too. You can see also why mathematicians and computer scientists bridle at careless definitions like "An infinite set is one whose members are uncountable." For over a century we have tried to teach such matters carefully, trying to get them right and trying also to convey to you now the imaginations and insights from great minds of past centuries.

Alas, in every army large enough, there is always somebody
who doesn't get the message, or gets it wrong, or forgets it.

I have written a very long answer to your letter in the hope of liberating you and your teacher (for whom a copy of this letter is enclosed) from a widespread fear of infinity described plaintively by an influential French poet, playwright and novelist Alfred de Musset (1810-57):

Malgré moi l'infini me tourmente.

(I can't help being tormented by the idea of the infinite.)

Why was your letter passed to me to answer? Partly because I persuaded the computer industry to include hardware support for ∞ along with the *floating-point* numbers used throughout science and engineering for fast approximate computation. Infinity has turned out to be a mixed blessing.

For further heavy reading …

About counting finite sets, and the distinction between countably infinite and uncountably infinite sets, see books about *Discrete Mathematics* like K.H. Rosen's. There are chapters about infinite sets in *Introduction to the Foundations of Mathematics* by R.L. Wilder, but it is not really introductory.

About the evolution of mathematical ideas, see books on the *History of Mathematics* like C.B. Boyer's or D.J. Struik's. An excellent old survey of mathematics, including a short chapter about infinity, is *What is Mathematics?* by R. Courant and H. Robbins (but they say "denumerable" instead of "countable").

About connections between numbers and language, see a marvelous book *Number Words and Number Symbols* by K. Menninger.

About points, lines and circles at infinity, see texts on *Geometry* like H.S.M. Coxeter's or R. Hartshorne's. Stereographic projection is discussed in books like H. Schwerdtfeger's *Geometry of Complex Numbers*.

About how computers handle ∞ , see my web page `http://www.cs.berkeley.edu/~wkahan` .

Churchill's reminiscence is quoted widely. It comes from his book *My Early Life: A Roving Commission* (1930).

Your librarian may be able to direct you to lighter reading about those subjects; but always stay skeptical of mathematical statements (including mine) that you don't understand. Remembering them is O.K. You don't have to believe them until you have figured them out for yourself. That way you reduce the risk of being misled by someone else's misunderstandings which, in my long experience, are foremost among the reasons why people find Mathematics hard at school.

Yours sincerely,

Prof. W. Kahan
Mathematics Dept. #3840
University of California
Berkeley CA 94720-3840

`<wkahan@eecs.berkeley.edu>`