

Is $0.999\dots = 1$?

FRED RICHMAN
Florida Atlantic University
Boca Raton, FL 33431

Few mathematical structures have undergone as many revisions or have been presented in as many guises as the real numbers. Every generation reexamines the reals in the light of its values and mathematical objectives. [3]

Arguing whether $0.999\dots$ is equal to 1 is a popular sport on the newsgroup sci.math. It seems to me that people are often too quick to dismiss the idea that these two numbers might be different. The issues here are closely related to Zeno's paradox, and to the notion of potential infinity versus actual infinity. Also at stake is the orthodox view of the nature of real numbers.

One argument for the equality goes like this. Set $x = 0.999\dots$, multiply both sides by 10 to get $10x = 9.999\dots$, then subtract the first equation from the second. The result is $9x = 9$, so $x = 1$. Essentially you are observing that $9x + x = 9 + x$, which is true, and then concluding that $9x = 9$. That's a valid inference, *if* x is cancellable.

But one man's proof is another man's *reductio ad absurdum*. Although most everyone will agree that the above argument shows that if x is cancellable, then $x = 1$, the believer and the skeptic differ in their interpretation of what this means. The believer, quite reasonably, takes for granted that you can cancel x , and regards the argument as a proof of the equality. For the skeptic, who considers the equality to be false, the argument is a proof that you cannot cancel x (if you could, then the equality would hold). So the skeptic must adopt the position that subtraction of real numbers is not always possible.

The skeptic would say that $9x$ is equal to $8.999\dots$, not 9, using the usual algorithm for multiplication: In each digit position we multiply 9 times 9, and add to that a carry of 8 from the position to the right, except at the position before the decimal point, where we simply get the carry of 8. The skeptic considers the number $8.999\dots$ to be different from 9, just as $0.999\dots$ is different from 1, even though $8.999\dots + x = 9 + x$.

Here is an even simpler argument for the equality: Multiply the equation $1/3 = 0.333\dots$ by 3 to get $1 = 0.999\dots$. The multiplication step is pretty hard to fault, so a skeptic must challenge the first equation. This simple argument gets its force from the fact that most people have been trained to accept the first equation without thinking.

Yet a third kind of argument is that

$$0.999\dots = \frac{9}{10} + \frac{9}{100} + \frac{9}{1000} + \dots,$$

and the sum of the geometric series on the right is

$$\frac{9/10}{1 - 1/10} = 1.$$

A skeptic who accepts the series interpretation could say that $0.999\dots$ *converges* to 1, or that it is equal to 1 *in the limit*, but is not *equal* to 1. Standard usage is ambiguous as to whether the expression on the right stands for the series or for its limit. The fact that we use that notation whether or not the series converges argues in favor of the series interpretation. Also, we talk about the rate of convergence of such expressions.



Is $0.999 \dots = 1$?

Author(s): Fred Richman

Source: *Mathematics Magazine*, Vol. 72, No. 5 (Dec., 1999), pp. 396-400

Published by: Mathematical Association of America

Stable URL: <http://www.jstor.org/stable/2690798>

Accessed: 24/02/2011 08:58

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=maa>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Mathematical Association of America is collaborating with JSTOR to digitize, preserve and extend access to *Mathematics Magazine*.

<http://www.jstor.org>

So some distinction between convergence and equality in the present case might well be appropriate.

Perhaps the situation is that some real numbers, like $\sqrt{2}$, can only be approximated, whereas others, like 1, can be written exactly, but can also be approximated. So $0.999\dots$ is a series that approximates the exact number 1. Of course this dichotomy depends on what we allow for approximations. For some purposes we might allow any rational number, but for our present discussion the terminating decimals—the decimal fractions—are the natural candidates. These can only approximate $1/3$, for example, so we don't have an exact expression for $1/3$.

This might be a good point to say a word about infinity, potential and actual. What do those three dots in $0.999\dots$ signify? Simply the unending possibility of writing down more nines. That's what *potential infinity* is. *Actual infinity* is the idea that all those nines have been written down. Potential infinity invites us to consider $0.999\dots$ as a process. Actual infinity invites us to think of $0.999\dots$ as a completed object. People who think of $0.999\dots$ as a process—a series rather than a limit—are not so tempted to equate it with 1.

The most famous exponent of potential infinity, as opposed to actual infinity, was Aristotle, who said, in Book III of his *Physics* [6], "In no other sense, then, does the infinite exist; but it does exist in this sense, namely potentially," and, "my argument ... denies that the infinite can exist in such a way as to be actually infinite." Gauss [4] considered "the use of an infinite quantity as a completed one" as "never permissible in mathematics." On the other hand, Georg Cantor [1], who set the tone for twentieth century mathematics, believed that "each potential infinite ... presupposes an actual infinite."

Zeno's paradox of dichotomy has to do with the idea that, when you approach an object, you repeatedly halve your distance from it, with the result that you never reach it. Here is a decimal version of the paradox appropriate to the issue at hand: If you travel from 0 to 1, you must go successively through the points 0.9, 0.99, 0.999, 0.9999, and so on. You can never reach 1 because you never finish visiting all those intermediate points. Of course this paradox has been refuted many times, but that is an essential part of being a paradox: there are compelling arguments both for and against it. As Benson Mates said [7], "it is possible to have impeccable arguments for both sides of a contradiction." Zeno's paradox has survived for thousands of years. The controversy about $0.999\dots = 1$ is one of its aspects.

Decimal numbers What kind of setting would support the skeptic's view? By a (nonnegative) *decimal number*, I mean an infinite string of digits with a decimal point, like 1247.4215347528... As usual, we don't allow the string to start with a 0 unless the decimal point comes immediately after. There are a couple of standard notational conventions. An infinite string of 6's (without a decimal point) is denoted by $\bar{6}$. So the number at issue, $0.999\dots$, is denoted by $0.\bar{9}$. The number 120.45 $\bar{0}$ is said to *terminate*, and is denoted simply by 120.45, while 120. $\bar{0}$ is denoted by 120 (with no decimal point).

The decimal numbers are ordered in the standard way. Line up the decimal points and compare corresponding digits. At the first place where the digits differ, the number with the bigger digit is the bigger number. So $999.999\dots$ is less than $1247.421\dots$, because the initial 1 of the latter number is the first place where the digits differ, while $1247.430\dots$ is bigger than $1247.421\dots$, because the 3 in the former number is bigger than the corresponding digit 2 in the latter, and that is the first place they differ. In particular, $0 = 0.000\dots$ is the smallest decimal number, and $0.999\dots < 1$.

How do you add two decimal numbers? There is a problem because carries can propagate over arbitrarily long stretches, and we can't start adding at the far right! But the carry can never be bigger than 1, so if the sum of the two digits in a given place is less than 9, or if it is greater than 9, then we can compute the digits in the sum up to, but not necessarily including, that place. If the sum of the digits is *exactly* 9 from some place on, then there will be no carry at, or past, that place.

Digression The question as to whether there is a carry into a given place cannot be decided by a finite computation. That means that you can't necessarily *compute* the decimal expansion of a sum from the decimal expansions of its addends. For example, suppose we have a number x whose decimal expansion starts out $0.05555\dots$. So x is close to $0.0\bar{5}$, but we can't be sure it is equal to $0.0\bar{5}$ because we only know as many digits in its expansion as we care to compute. It may be that $x \leq 0.0\bar{5}$, or it may be that $x > 0.0\bar{5}$. What is the first digit after the decimal point in the expansion of $0.0\bar{4} + x$? It is 0 if $x \leq 0.0\bar{5}$, and it is 1 if $x > 0.0\bar{5}$. We may not have enough information, even after computing the first million places, to determine which of these alternatives holds.

The fact that you can't compute the decimal expansion of a sum from the decimal expansions of its addends is a well known phenomenon that was noticed by Turing. In a fully constructive treatment of the real numbers, this is often stated by saying (informally) that not every positive real number has a decimal expansion. More precisely, there is no known constructive proof that every positive real number has a decimal expansion.

What is the nature of this algebraic structure—decimal numbers under addition—that we have defined? You can check that the addition is commutative and associative, and that there is an identity, 0. The cancellable elements are precisely the terminating decimals, because $0.9 + x = 1 + x$ for all nonterminating x .

Is $1/3 = 0.\bar{3}$? Clearly, if a sum is cancellable, then each addend is cancellable, so there is no decimal number x such that $x + x + x = 1$. That is, $1/3$ is not a decimal number. More generally, no nonterminating decimal number x can satisfy an equation of the form $mx = n$ with m and n positive integers.

What about multiplication of decimal numbers? This is more complicated than addition, but can also be described as a natural extension of the way we multiply decimal fractions. We saw a little of that in the description of how to multiply 9 times $0.999\dots$ to get $8.999\dots$. However, it is convenient to define multiplication in terms of cuts, and we will want to look at cuts in any case for the insight they give into the controversy.

Dedekind cuts Dedekind cuts are usually defined in the ring of rational numbers, but if we are interested in decimal numbers, we will want to work with a different ring. Let D be any dense subring of the rational numbers. That is, D is any subring of the rational numbers other than the ring of integers. We have in mind the ring of *decimal fractions*, those rational numbers that can be expressed with denominator a power of 10. A *Dedekind cut* in D may be defined as a nonempty proper subset S of D such that if $x < y$ and $y \in S$, then $x \in S$.

This is essentially Dedekind's definition in [2]. Dedekind then identified the cut $\{x \in D : x < r\}$ with the cut $\{x \in D : x \leq r\}$, for each r in D , saying they were "only unessentially different." A similar move, made for example in [8, Definition 1.4], is to restrict ourselves to Dedekind cuts that do not have a greatest element, so $\{x \in D : x \leq r\}$ is not considered to be a cut. Why do that? Precisely to rule out the existence of distinct numbers $0.\bar{9}$ and 1. Indeed, $0.\bar{9}$ corresponds to the cut

$\{x \in D : x < 1\}$ while 1 corresponds to the cut $\{x \in D : x \leq 1\}$. In general, we may identify an element d in D with the cut $\{x \in D : x \leq d\}$ (we call these *principal cuts*). So we see that in the traditional definition of the real numbers, the equation $0.\bar{9} = 1$ is built in from the beginning. That is why anyone who challenges that equation is, in fact, challenging the traditional formal view of the real numbers.

If D is the ring of decimal fractions, then each decimal number ξ gives rise to a Dedekind cut

$$\{x \in D : x \leq \xi\}$$

in D . Note that this cut contains 0. Conversely, any Dedekind cut S in the ring of decimal fractions, that contains 0, is associated with a unique decimal number ξ as follows. For fixed m , the largest element of S that is of the form $n/10^m$ gives the digits in ξ up to the m -th place to the right of the decimal point. For example, the nonterminating decimal number 3.1415926535... corresponds to the cut consisting of all those decimal fractions r such that $r < 3$, or $r < 3.1$, or $r < 3.14$, or $r < 3.141$, or ...

Let cut D denote the set of all Dedekind cuts in D . Define the sum of two cuts in the usual way

$$u + v = \{x + y : x \in u \text{ and } y \in v\}.$$

It is easily shown that the commutative and associate laws hold, and that the principal cut $\{x \in D : x \leq 0\}$ is an additive identity. So cut D is a *monoid*. The elements of D , in the guise of principal cuts, form a subgroup of this monoid. In fact, D consists precisely of the (additively) cancellable elements of cut D . This is because

$$u + d = \{x + d : x \in u\}$$

while if $1^- = \{x \in D : x < 1\}$, then $u + 1^- = u + 1$ for any cut u that is not principal. However, the nonprincipal cuts are cancellable among themselves, and are closed under addition, so they also form a subgroup of cut D . This group may be identified with the traditional real numbers, as Rudin [8] does with cuts in the rational numbers. Recall that any traditional positive real number has a unique *nonterminating* decimal expansion. Note that $0^- = \{x \in D : x < 0\}$ is the identity element of the group of nonprincipal cuts.

The order on cut D is given by inclusion of cuts. The *weakly positive* cuts are those that contain the rational number 0. These correspond exactly to the decimal numbers if D is the ring of decimal fractions. The *product* of two weakly positive cuts u and v is defined to be $\{st : s \in u \text{ and } t \in v\}$. This multiplication on weakly positive cuts shows how to multiply any two decimal numbers. It's straightforward to show that the associative, commutative, and distributive laws hold. So the decimal numbers form a *positive, totally ordered, commutative semiring* in the sense of [5].

The picture here is the traditional real numbers, in the form of nonprincipal cuts, living uneasily together with the ring D , in the form of principal cuts. For each element d of D , there is a traditional real number d^- just below it, and $u + d^- = u + d$ for each traditional real number u . That, for traditionalists, is a complete description of the additive structure of cut D . Note that $d^- = d + 0^-$.

Clearly $0.\bar{9} = 1 + 0^-$, so 0^- is sort of a negative infinitesimal. On the other hand, you can't solve the equation $0.\bar{9} + X = 1$ because, in cut D , the sum of a traditional real with any real is a traditional real.

Another point of view We looked at Dedekind cuts in order to describe multiplication of decimal numbers, and to see another way of describing the decimal numbers

themselves. Because we looked at cuts in the ring of decimal fractions, both positive and negative, we got some numbers in addition to the decimal numbers. While the number 1^- corresponds to the decimal number $0.\bar{9}$, there is no decimal number corresponding to $(-1)^-$, which is the cut $\{x \in D : x < -1\}$. Nor is there a decimal number corresponding to -1 itself.

However, cut D can be characterized in the following way (for D the nonnegative decimal fractions). It contains the decimal numbers, and all the decimal fractions, both positive and negative. Each element of cut D can be written as a difference $u - d$ of a decimal number u and a (nonnegative) decimal fraction d . We may think of cut D as being obtained from the decimal numbers by adjoining the negative decimal fractions, and taking sums. This construction is legitimate because the decimal fractions are cancellable in D .

Instead of extending the decimal numbers to include additive inverses of those decimal numbers that are cancellable under addition, we could extend them to include multiplicative inverses of those decimal numbers that are cancellable under multiplication. These are exactly the positive decimal fractions, because $(0.\bar{9})x = x$ whenever x is a nonterminating decimal number. This construction is an instance of forming a semiring of fractions; see [5]. It is not hard to verify that the result is (isomorphic to) the weakly positive elements of cut \mathbb{Q} , where \mathbb{Q} is the ring of rational numbers.

Open problems Although we can introduce negative decimal fractions, negative numbers in general present a serious problem because we don't have cancellation in cut D . We can't simply write them as additive inverses of positive numbers. Moreover, we have no interpretation for the number $-3.14159265\dots$ because this represents a process of approximation from *above*, -3.14159 being greater than $-3.14159265\dots$, whereas in cut D all real numbers are approximated from *below*. Of course we could just introduce symbols like $-3.14159265\dots$, but it's not clear how to get a satisfactory coherent system that incorporates them.

Because of this, multiplication of *arbitrary* real numbers is also a serious problem, if for no other reason than that we don't know how to multiply -1 by $3.14159265\dots$. Even in the traditional approach, multiplication is awkward. The elegant treatment of addition is replaced by an ugly division into cases: one defines how to multiply positive numbers, and extends to negative numbers according to the usual rules [8, pp. 7–8].

REFERENCES

1. Cantor, Georg, *Über die verschiedenen Ansichten in Bezug auf die actualunendlichen Zahlen*, 1886, quoted in Hallett, Michael, *Cantorian Set Theory and Limitation of Size*, Oxford University Press, Oxford, UK, 1988.
2. Dedekind, Richard, Continuity and irrational numbers, (1872), in *Essays on the Theory of Numbers*, Dover, New York, NY, 1963.
3. Falting, F., N. Metropolis, B. Ross, and G.-C. Rota, The real numbers as a wreath product, *Advances in Math.*, 16 (1975), 278–304.
4. Gauss, Carl Friedrich, *Briefwechsel zwischen C. F. Gauss und H. C. Schumacher*, 1860, quoted in Dauben, Joseph Warren, *Georg Cantor, His Mathematics and Philosophy of the Infinite*, Princeton University Press, Princeton, NJ, 1990.
5. Golan, Jonathan S., *The Theory of Semirings with Applications in Mathematics and Theoretical Computer Science*, Wiley, New York, NY, 1991.
6. Heath, Thomas, *Mathematics in Aristotle*, Oxford University Press, Oxford, UK, 1970.
7. Mates, Benson, *Skeptical Essays*, University of Chicago Press, Chicago, IL, 1981.
8. Rudin, Walter, *Principles of Mathematical Analysis*, McGraw-Hill, New York, NY, 1964.