Brouwerian Counterexamples

An 80-year-old but little-known method demonstrates the lack of numerical meaning in many classical theorems.

MARK MANDELKERN
New Mexico State University
Las Cruces, NM 88003

Not only did he launch a controversy which continues to the present day, but in his critique of classical mathematics early in the century, L. E. J. Brouwer also initiated a new mode of reasoning. In contrast to the idealistic thought common since Greek times, characterized by the notion that truth exists independently of humans, Brouwer realistically held as true only what was currently known. Thus truth changes from day to day, and also from person to person. Brouwer developed a type of counterexample which shows when a given statement is not true in this realistic sense. A recent paper [**32**] discussed, with no technical details, the general aspects of the controversy over constructivity which ensued from Brouwer's work. This paper considers the technical details involved in Brouwerian counterexamples; it tries to steer clear of polemics, leaving readers to decide for themselves whether or not the counterexamples indicate a need for constructive considerations in the practice of mathematics.

The term *classical mathematics*, as used here, refers to the sort of mathematics taught in virtually every school and college classroom in the world. Work in *constructive mathematics*, using only constructive methods, is (at least for the present) carried on by only a very small minority of mathematicians.

There is a certain danger in devoting an entire paper to negativistic results, which could give the erroneous impression that the purpose of constructive mathematics is to consider pathological counterexamples (see Appendix). The only reason for these counterexamples is to show that certain theorems in classical mathematics are not constructively valid, and thus to indicate the need for their replacement by positive constructive results. A few of these positive results are indicated here, but for a more complete exposition the reader must consult the literature; for a thorough introduction, see [**3**] or [**6**].

A new kind of counterexample Brouwer's critique of many classical theorems, his claim that they lacked numerical meaning, consisted of demonstrations showing that their truth would imply solutions to problems for which, in fact, no solutions were known. Thus he concluded that certain classical theorems could not be true, for if they were true, many people would be trying to use them to solve the unsolved problems. Such a demonstration is called a *Brouwerian counterexample*; it differs from an ordinary counterexample, which demonstrates that a given statement implies another statement which is known to be false.

Brouwer's examples were typically constructed on an *ad hoc* basis, using a variety of unsolved problems more or less at random—sometimes quite famous problems such as "Fermat's Last Theorem," and sometimes problems remarkable only for their insignificance, such as questions concerning the digits in the decimal expansion of π. More systematically, Errett Bishop [3] formulated several general *omniscience principles*, each of which implies the solution to a vast number of unsolved problems, or at least leads to certain information about such problems which in fact is not available. A Brouwerian counterexample then becomes a demonstration that a certain statement implies one of these omniscience principles. One advantage of this is that it shows more clearly the nonconstructivities in classical mathematics, since there will always be unsolved problems which the omniscience principles would solve. Here we first give a few of Brouwer's original *ad hoc* examples, and then develop the systematic formulation.

The least upper bound principle A characteristic and important theorem of classical mathematics is the *least upper bound principle*. To show its nonconstructive nature (its lack of numerical meaning), we'll show that if it were true, it would provide a finite method leading either to a proof of Fermat's Last "Theorem" or to an explicit counterexample. Since no one knows such a method, no one can claim that the least upper bound principle is true in a numerical sense. Fermat probably had no proof (see [42, Ch. 13] and [44]), but if he did, then for him it was a "theorem", while we can speak only of "Fermat's Last Problem."

Since Fermat's Last Problem concerns equations in integers, while the least upper bound principle concerns sets of real numbers, we must somehow represent the former by the latter. We'll construct a set S of real numbers by a sequence of steps, as follows. We start with step number 3. Using only positive integers up to and including 3, we look for a solution to the equation $x^n + y^n = z^n$, with $n = 3$. Finding no solution, we put the real number $1/3$ into our set S. (Even when any integers x, y, z are allowed, there is no solution with $n = 3$; see, for example [7, Th. 44.9].) If we had found a solution, we would have put the number $1 - 1/3$ into S. This process for constructing S continues; at the kth step we consider all quadruples (x, y, z, n) of positive integers up to k, with the restriction $n \geqslant 3$. When we find no solution to the Fermat equation, we put $1/k$ into S; when we do find a solution, we put $1 - 1/k$ into S. This defines the set S; see FIGURE 2. (There are many ways to form a suitable set; there is nothing special about the method used here.) Notice that the procedure at the kth step is always a finite one, although the whole process is infinite.

Now we apply the *hypothesis* that the least upper bound principle is true. We take this assertion in the numerical sense that we can calculate least upper bounds; that is, we can find explicit, arbitrarily close, rational approximations. Let t be the least upper bound of the set S. For example, someone might specify the number t using a decimal expansion. While such an expansion is infinite, only a finite process is required to calculate any given digit. We need calculate only one digit to tell whether t is less than 0.6 or more than 0.4. In the first case, it is easy to predict with absolute certainty that the Fermat equation will never be solved (with $n \geqslant 3$); you have proved Fermat's Last Theorem. In the second case, by the definition of least upper bound there is a number x in the set S that is more than 0.4. Now this number x must be of the form $1 - 1/k$. Looking at this integer k, you know exactly where to find, using a finite process, a solution to the Fermat equation, a counterexample to Fermat's Last Theorem. In either case you have solved Fermat's Last Problem. It is convenient to restate this counterexample concisely as follows.

Example 1. The *least upper bound principle* is nonconstructive; it would imply a solution to Fermat's Last Problem.

How do constructivists answer the classicist who (omnisciently) looks at Example 1

FIGURE 1
L. E. J. Brouwer (1881–1966). This article commemorates the 80th anniversary of Brouwer's seminal doctoral thesis.

and says, "Well, it's perfectly obvious that the least upper bound of the set S is either $1/3$ or 1." Their answer is, "If you could really tell which one of these alternatives actually holds, you wouldn't be here discussing it with us. You would be at your desk writing either your proof of Fermat's Last Theorem or your proof that you have a finite procedure for finding a counterexample." Saying that a number with certain properties, *if it exists*, must be equal to one or the other of two known numbers, is quite different from saying that such a number actually exists. When a constructivist

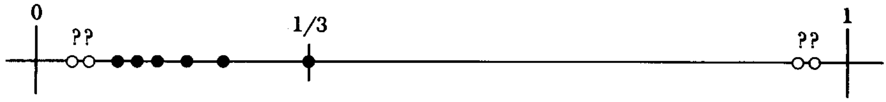

FIGURE 2.
Here are a few of the points, and possible points, of the set S constructed in Example 1 by searching for solutions of the Fermat equation. The points begin at $1/3$ and proceed to the left. However, if a solution is ever found, then points suddenly begin to appear at the far right of the interval. What is the least upper bound of S?

says that a number is either this or that, he is prepared to say which; or at least to give a *finite* procedure by which it may be determined.

Thus a person who really had a constructive proof of the least upper bound principle, a finite procedure for calculating rational approximations to the least upper bound, would immediately begin applying it to the set S to find a solution to Fermat's Last Problem, and similarly for hundreds of other unsolved problems in number theory and analysis.

How does constructive analysis proceed without the important least upper bound principle? By restricting to a certain class of sets, which suffice for all constructive applications, one obtains a constructive substitute; see Theorem 3 below.

Counterexamples such as the above are the basis of Brouwer's critique of classical mathematics, begun in 1907 [12], and the motivation behind modern Bishop-type constructive mathematics [3]. In the following sections we consider a few more informal counterexamples, and then, using these as a guide, adopt a precise formulation. The resulting analysis will eliminate the apparent dependence on specific unsolved problems. Thus, while Example 1 shows that the least upper bound principle is nonconstructive because it implies a solution to Fermat's Last Problem, it remains nonconstructive even if tomorrow somebody solves the problem.

Numerical meaning The general notion of numerical meaning was discussed in [32]; here we try to make this idea more precise. It is the strict constructive notion of numerical meaning which allows us to proceed from the existence of a least upper bound to a solution of Fermat's Last Problem. The least upper bound is a real number; thus we must formulate a precise definition of a constructive real number. This is nothing other than the same real number that is used throughout classical analysis, both elementary and advanced, but with a stricter interpretation. A real number is a Cauchy sequence of rational numbers. Each term of the sequence is a rational number, an explicit quotient of integers. There is no difficulty or controversy about the ordinary integers, their rational quotients, and the finite operations among these. But the idea of an infinite sequence is more difficult. The notion of infinity has had a long and interesting history, and has been discussed by thinkers from the ancient Greeks to modern astronomers. In mathematics there is a sharp contrast between the idea of a potential infinity and that of an actual infinity. This is, perhaps, the crux of the controversy between classical (idealistic) mathematics and constructive (realistic) mathematics. In the latter, only a potential infinity is considered. This means that while the definition of a real number allows you to calculate however many terms of the approximating sequence you want (and have time for), still you can never expect to calculate all the terms.

A real number might be defined by a sequence of approximating rational numbers expressed as finite decimals. For example, $\sqrt{2}$ can be defined by the sequence

$$1, 1.4, 1.41, 1.414, 1.4142, 1.41421, 1.414213, \ldots.$$

This short list of approximations is the result of only the first few applications of the well-known rule which explicitly defines each term of the sequence. A sequence is a function, or rule, rather than an infinite list; the latter has only a potential existence. We will focus attention on such rules. Thus the least upper bound principle (if constructive) must provide a *finite procedure* by which any desired approximation to the least upper bound may be calculated. Example 1 shows that the classical theorem on least upper bounds provides no such procedure. The notion of constructive procedure requires that it be finite only *in principle*. Errett Bishop expressed this as follows. "How do you know whether a proof is constructive? Try to write a computer program. If you can program a computer to do it, it should be constructive. Notice I said write the program. Don't necessarily run it on the computer and wait around for the result." [4]

Trichotomy of real numbers This is a precept of classical mathematics so ingrained in the thought of working mathematicians and students that assaults on its validity are usually met not with rigorous defensive measures, but (even more effectively) with complete disregard. That any given real number is either positive, negative, or zero may seem so familiar an idea as to be intuitively true, but nevertheless it is nonconstructive. To show this, it suffices to show that the trichotomy principle, interpreted so as to have numerical meaning, would imply the solution to an unsolved problem. For the latter we take the *perfect number problem*. A positive integer is *perfect* if it is the sum of its proper divisors, for example $6 = 1 + 2 + 3$, $28 = 1 + 2 + 4 + 7 + 14$, 496, etc.; see [7; 11.25–11.26]. Nobody knows whether or not there exists an odd perfect number. Rather than constructing a set to establish a connection between the principle under test and the unsolved problem, as in the last example, in this case we need a real number. It requires only a finite procedure to test an integer for perfection; the results are used to define a sequence $\{\alpha_k\}$. At the kth step, if no odd perfect number $\leqslant k$ is found, define $\alpha_k = 0$, but if one is found, define $\alpha_k = 1/2^k$. Now define

$$\alpha = \sum_{k=1}^{\infty} \alpha_k.$$

Clearly $\alpha \geqslant 0$; applying the hypothesis that the trichotomy principle has numerical meaning, either $\alpha = 0$ or $\alpha > 0$. In the first case it follows that each α_k is 0; thus we have a theorem, every perfect number is even. In the second case at least one of the terms α_k must be positive, and this leads us straight to an explicit construction of an odd perfect number. (We have used constructive notions of positive real numbers and convergent sequences and series, which are discussed below.) Thus we have proved the following.

Example 2. The *principle of trichotomy of real numbers* is nonconstructive; it would imply a solution to the perfect number problem.

Discontinuous functions It is a common practice in elementary analysis courses to demonstrate the importance of continuity conditions in certain theorems by giving examples of discontinuous functions, showing what can go wrong when continuity is not present. However, there are problems in the construction of such discontinuous functions.

Consider a typical function often mentioned in calculus classes; it is defined on the unit interval $[0,1]$ by

$$f(0)=0$$
$$f(x)=1 \quad \text{whenever } x>0.$$

This definition presents serious constructive difficulties. As we saw in Example 2, there is no general finite procedure for deciding whether a given real number in the interval is zero or positive. Thus there is no general finite procedure for deciding, for a given point x in the interval, what value f assigns to x. This means that while the definition above does define a function, this function is not defined on the entire interval, but only at those points for which one knows one or the other of the two alternatives, zero or positive. Example 2 constructs a number in the interval for which neither alternative is known.

There is nothing nonconstructive about the definition above, it defines a function which may even be sketched as shown in FIGURE 3, but this function is not defined on the entire unit interval. Thus a Brouwerian counterexample on this topic will concern the following statement.

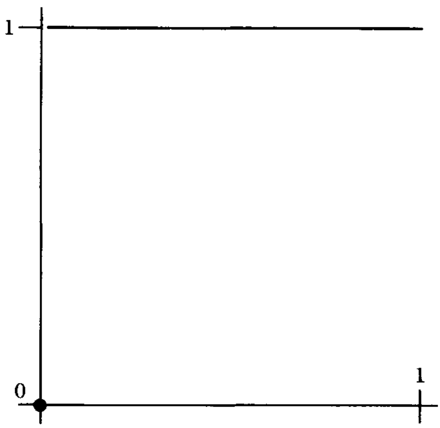

FIGURE 3.
A good try at defining a discontinuous function on the closed unit interval, but it doesn't work. There are points x for which we don't know whether $x=0$ or $x>0$; at such points the function is not defined.

DISCONTINUOUS FUNCTION PRINCIPLE (DFP). *There exists a function f defined on the closed unit interval* $[0,1]$ *such that* $f(0)=0$, *and* $f(x)=1$ *whenever* $x>0$.

This statement, although it uses some of the same phrases, is distinctly weaker than the statement that the definition above defines a function on the entire interval; the latter statement has already been shown to be nonconstructive by Example 2. The statement DFP does not state that the conditions $f(0)=0$, and $f(x)=1$ for $x>0$, define a function on the entire interval, but rather that some function exists, defined somehow on all of $[0,1]$, which has these values at these points, whatever other values it may have at whatever other points.

For a Brouwerian counterexample to DFP, we will use another unsolved problem from number theory, the Goldbach Conjecture. This conjecture says that every even number greater than 2 can be represented as the sum of two primes, such as $4=2+2$, $6=3+3,\ldots,$ $1928=1201+727$, etc. Goldbach stated this in Moscow, in

1742 (= 1429 + 313), but to this day no one knows whether it is true or false; see [**17**, p. 30] or [**41**, p. 47]. Use the method of Example 2; test each even number in succession to obtain a sequence $\{\alpha_k\}$ of nonnegative rational numbers, and a real number $\alpha = \Sigma\alpha_k$, where α_k represents the result of testing the kth even number, beginning with 4. For any given even number, it requires only a finite calculation to test the Goldbach Conjecture in that instance. If for every even number > 2 there do exist two prime summands, then all α_k will be 0 and thus $\alpha = 0$. But if the Goldbach Conjecture ever fails, then some α_k will be $1/2^k$, and $\alpha > 0$. Thus we have constructed a real number α in the unit interval such that the Goldbach Conjecture is true if and only if $\alpha = 0$.

Now apply the hypothesis that DFP is true, and use it to calculate $f(\alpha)$. The resulting information about the Goldbach Conjecture is a bit different from that obtained in the first two examples. This means that the nonconstructivity of DFP is different from that of the least upper bound principle and trichotomy; DFP is a less powerful hypothesis. Calculate a rational approximation to $f(\alpha)$ sufficient to determine whether $f(\alpha) < 1$ or $f(\alpha) > 0$. In the first case α cannot be positive and thus from the elementary constructive properties of real numbers it follows that $\alpha = 0$; the Goldbach Conjecture is true. The second case is different; we can conclude that α cannot be zero, but it does not follow that $\alpha > 0$. (The constructive ordering of the real line is discussed below. The difference arises because $\alpha \leqslant 0$ is an essentially negativistic statement (equivalent to the impossibility of $\alpha > 0$), whereas $\alpha > 0$ is an affirmative statement which requires a construction that is not available in this situation.) So in the second case we can only conclude that the Goldbach Conjecture cannot be true, without, however, finding an explicit counterexample. This strange sort of situation has not been known to arise in number theory or analysis. Nevertheless, such a partial solution to an unsolved problem would certainly be interesting, and of pragmatic significance in focusing further research. Thus our derivation of such unavailable information about the Goldbach Conjecture serves sufficiently well to establish the nonconstructivity of DFP; we have proved the following.

Example 3. The *discontinuous function principle* is nonconstructive; it would imply either a proof of the Goldbach Conjecture or a proof of its falsity (without an explicit counterexample).

The weaker conclusion in this example, compared with the first two, is not due to properties of the Goldbach Conjecture, but rather of continuity. We could equally well have used the Goldbach Conjecture in the first two examples. The systematic formulation of Brouwerian counterexamples will clarify these matters by separating the principles under test from the unsolved problems.

The intermediate value theorem Before systematizing the method of Brouwerian counterexamples, we'll discuss one which subjects the intermediate value theorem to a test for numerical meaning, relating it to the decimal expansion of π. The intermediate value theorem says that a continuous function which is positive at the left end of an interval, and negative at the right, must be zero at some point between. See, for example [**22**, p. 190]. This theorem was first proved rigorously (in the classical sense) by Bernard Bolzano, in Bohemia, in 1817 [**8**]. Bolzano was one of the first modern mathematicians who tried to eliminate geometric intuition from proofs; he might well have appreciated the constructive approach of this century; see [**1**, p. 15], [**21**], and [**26**]. The nonconstructivity of the intermediate value theorem was discussed informally in [**32**].

Since there is no connection between intermediate values and questions about digits in the expansion of π, the need for a systematic analysis becomes even more pressing. Questions about these digits were a favorite of Brouwer's; for example, see [**15**, p. 6]. They are of no significance; their only value is that it is unlikely that anyone will ever try to solve them, and thereby necessitate a revision of all the *ad hoc* examples in which they are used.

Here we'll show that the intermediate value theorem, if constructively valid, would lead to an answer to the question "If the sequence of digits 123456789 ever appears in the decimal expansion of π, and the digit 9 in the first such sequence occurs at the nth place, will n then be even or odd?" Notice that no integer n is actually defined by this question. Only if someday someone finds a sequence 123456789 will n then be defined. An answer to the question would be only a prediction whether, in this event, n will be even or odd. This is a much weaker question than whether or not such a sequence of digits does occur. The use of the weaker question reflects nothing about the number π, but rather about the intermediate value theorem under test, which is not enough to answer the stronger question.

To apply the intermediate value theorem, we must define a continuous function, and to define this function, we first define a real number β. This number is defined by an infinite series, and this in turn is based on the decimal expansion of π. Define

$$\beta = \sum_{k=1}^{\infty} \frac{\alpha_k}{10^k},$$

where the factors α_k are integers to be defined in a moment. This is almost like giving a decimal expansion, but with an important difference: the integers α_k may be positive *or* negative. The rule for α_k is as follows. If, in the decimal expansion of π, the kth digit is at the end of a sequence 123456789, and this is the *first* such sequence, then α_k is 1 when k is even, but -1 when k is odd; otherwise, α_k is 0. Because the sequence $\{\alpha_k\}$ is bounded, the series converges and defines the real number β. Since we know that no sequence 123456789 occurs in at least the first few thousand digits, β is a very small number; but we don't know whether it is positive, negative, or zero.

To define a continuous function f on the closed unit interval $[0, 1]$, we first give f the values

$$f(0) = 1$$
$$f(1/3) = f(2/3) = \beta$$
$$f(1) = -1$$

and then complete the definition of f by using straight lines between these points. It must be shown that it is indeed possible to define constructively a continuous function in this way, but we postpone this until a later section. This is no minor point to be glossed over, for some similar constructions, such as the one discussed above in connection with DFP, are *not* constructively valid, and one must distinguish carefully.

FIGURE 4 shows three views of this function, corresponding to three possibilities for the number β, but it is dangerously misleading. We cannot say that one of the curves shown represents f. Example 2 shows that we do *not*, in general, know which possibility holds for a given real number β. In fact, the figure represents precisely those cases in which we are not interested, where there is no unsolved problem. When constructing Brouwerian counterexamples, the figures typically only remind us of what we don't know.

Now we apply the intermediate value theorem to the function f in order to "solve"

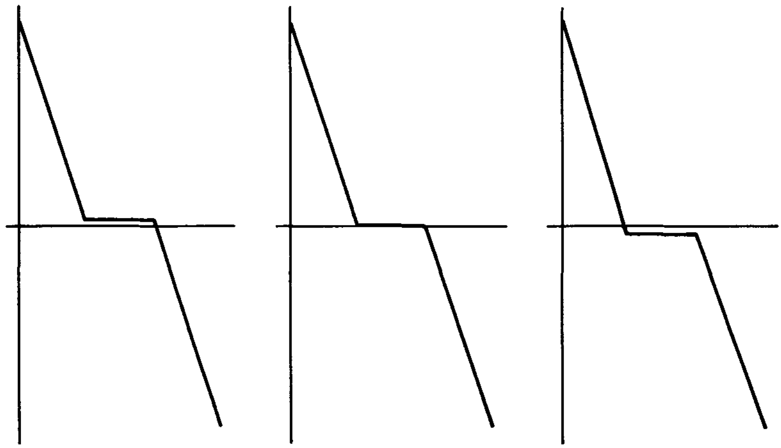

FIGURE 4
Three dangerously misleading views of the function f used in Example 4. Where f crosses the axis, nobody knows.

the "digits of π" problem. Since the function f has values both above and below the axis, the intermediate value theorem claims to provide a specific crossing point, a number x for which $f(x) = 0$. To investigate the numerical meaning of the intermediate value theorem, we interpret this claim in a strict constructive sense—we suppose that an explicit rule is given for the construction of x. If x is given by a decimal expansion, we then calculate the first digit. (We do not mean to imply that all real numbers have constructively defined decimal expansions, for there are some problems about this; see [36]. Nevertheless, using these expansions serves sufficiently well to demonstrate the method, which is essentially the same no matter how the real number x is approximated.) If the first digit is 5 or less, then it is clear that β cannot be positive, and it follows that the number n in the "digits of π" question (were it to exist) would be odd. On the other hand, if this digit is more than 5, then n would be even. Note that FIGURE 4 is useful in visualizing the proof, but this is only because we are under the spell of the intermediate value hypothesis.

Thus the intermediate value theorem, if constructive, would lead to a solution to the "digits of π" problem. Since, in fact, we have no such solution, we conclude that the intermediate value theorem is constructively invalid. This result is recorded as follows:

Example 4. The *intermediate value theorem* is nonconstructive; it would imply a solution to the "digits of π" problem.

Using sequences to encode unsolved problems To analyze the nature of a Brouwerian counterexample, notice that Example 1 can be broken down into two distinct parts. First we construct a set S which represents the (unsolved) Fermat Problem. Then we use the "theorem" under investigation, in this case the least upper bound principle, to obtain information about this set, thereby "solving" the unsolved problem. The set S plays merely an intermediary role; for other examples other mathematical objects may be used. A set is used in Example 1 because the least upper bound principle is about sets.

It is convenient and helpful to use only a few standard objects for these examples. Surprisingly, they can almost all be handled with a (seemingly) very simple object, an ordinary infinite sequence of integers, and even then using only two integers, 0 and 1. In fact, a sequence is really not so simple an object; the catch is in the (potentially) infinite process which defines the sequence. Since each of us is a merely finite being, no one can actually carry out the infinite process to see what happens. Only occasionally can we predict what would happen; that is the great accomplishment of mathematics, transcending the finiteness of human existence to obtain accurate predictions for infinite processes.

Consider a typical sequence each of whose terms is either 0 or 1. A proof that all the terms are zero constitutes a proof of some theorem, while the existence of a term equal to 1 would be a counterexample (in the usual sense). By mere calculation using the rule defining the sequence, we might find the first million terms to be all 0, but we have no way to tell what might happen in the next million, or further. The essence of a typical theorem in number theory or analysis is *prediction*. No one has ever proved anything by calculating all the terms of an infinite sequence, only by predicting the outcome of such potentially infinite calculations.

The significant fact about these simple sequences is that most of the unsolved problems of number theory and analysis can be encoded using them. To encode Fermat's Last Problem using a sequence, we define the term a_n as 0 in the first case indicated in Example 1, and as 1 in the second case. If you can prove that all the terms of the sequence are 0 (that is, *predict* in a convincing way that each term calculated, no matter how far into the future, will be 0), then you will have proved Fermat's Last Theorem. On the other hand, if someone ever calculates a term that turns out to be 1, then he or she will have a counterexample to Fermat's Last Theorem (in the ordinary sense), and can say that Fermat's Last Theorem is false. It is also conceivable (although less likely) that someone could prove that Fermat's Last Theorem is contradictory; then again it would be false, but in a different sense, as discussed above in connection with Example 3, and below in connection with WLPO.

Decision sequences These sequences of 0's and 1's are so useful for investigating the numerical meaning of mathematical statements, and for analyzing the nature of Brouwerian counterexamples, that it is convenient to adopt a few conventions for their use; in this way we'll see the similarity and relationship between different counterexamples. While it is possible to use sequences with any integers as terms, positive, negative, or zero, we will find it much simpler to use only sequences of 0's and 1's. These suffice for all presently known counterexamples, although one should be prepared for unforeseen complications which might arise in the future. This choice of only two values represents the situation where we use some finite process to look for something (as in the examples above) and either we find it or we don't; the 0 or 1 simply records our results. There are counterexamples in which more values may be convenient; for example, 0, -1, and 1 (which we used in the definition of β in Example 4). However, the systematic use of a single type of sequence has advantages, such as in the comparison of counterexamples, which outweighs this convenience. In situations where a term -1 in a sequence might be handy, we can use instead a factor such as $(-1)^n$ elsewhere (as in Example 4* below). Thus our decision sequences will consist only of 0's and 1's. Similarly, sequences with at most one term equal to 1 have often been useful for counterexamples, but again these are easily converted to examples using decision sequences, by the device used in Example 9 below.

Example 1 was typical in that the searches were cumulative; at each step the search included all previous searches. In practice, there would be no need to repeat all the previous work, but it was convenient to express the results that way, because if one search is successful (the result recorded as a 1), then all succeeding searches will be also. Thus the sequence of recorded results consists of an initial segment of 0's, and then (sometimes) a 1 occurs, after which all the remaining terms are also 1's. In Example 1, Fermat's Last Theorem is true if and only if the sequence consists only of 0's, and it is false (in the strong sense of an explicit counterexample) if and only if there exists a 1 in the sequence. In view of these considerations we adopt the following.

Definition. A *decision sequence* is a nondecreasing sequence $\{a_n\}$ of 0's and 1's.

By "nondecreasing" we mean that $a_n \leqslant a_{n+1}$ for all n. The only question of interest for a given decision sequence is whether or not 1's begin to appear somewhere. It is convenient to assume that decision sequences under consideration begin with a 0, for otherwise there is no problem.

Omniscience principles Decision sequences form the connecting link between nonconstructive classical theorems and unsolved problems. This link is made precise by formulating general statements about decision sequences. In Example 1, for example, a solution to an unsolved problem results if we can tell whether or not a certain decision sequence contains a 1. On the other hand, the least upper bound principle, if true in a numerical sense, would provide precisely that information for any decision sequence. Thus we formulate the following:

LIMITED PRINCIPLE OF OMNISCIENCE (LPO). *Given any decision sequence* $\{a_n\}$, *there is a finite procedure which results either in a proof that* $a_n = 0$ *for all* n, *or in the construction of an integer* n *such that* $a_n = 1$.

One usually says only "either all $a_n = 0$ or some $a_n = 1$". The explicit finite procedure, and the proof or construction, are implied. From a classical point of view, it is obvious that either all the terms are zero, or there is a 1. How can anyone imagine a situation in which neither of these alternatives is true? Such a situation would be one instance of what is referred to as the "Middle" in the *Principle of Excluded Middle*. This principle, which we'll refer to as EM, concerns not only decision sequences, but states that any (meaningful) statement is either true or false. Aristotle formulated this principle, but used it only in finite situations, in which it is constructively valid. When it is (inappropriately) applied to the mathematics of the infinite, it leads to results such as the least upper bound principle, the trichotomy principle, the discontinuous function principle, and the intermediate value principle, which, as we have seen, are nonconstructive. The principle of excluded middle is discussed further in [32, pp. 275–276]. Bishop referred to EM (with a reformulation) as the *Principle of Omniscience* ([3, p. 9] or [6, p. 11]); LPO is a special case which applies only to denumerable problems. To interpret LPO as stating that some statements are true or false requires a bit of care. LPO says that for any decision sequence, the statement "some term has value 1" is either true or false. "True" must be taken in the strict sense that an integer n is constructed and a proof is given that $a_n = 1$; then "false" leads to a proof that all $a_n = 0$. On the other hand, the statement "every term has value 0" does not produce the same results, because its falsity only involves a certain contradiction, from which the construction of an integer n, such that $a_n = 1$, does not follow. This is a good example of the difference between the existential and universal quantifiers, when used constructively.

Though it may seem obvious from a classical viewpoint, LPO appears in an entirely different light when viewed constructively. Classically, since according to EM every statement is either true or false, the middle alternative may be described as saying that both possibilities mentioned in LPO are false, and that indeed is unthinkable. But the constructive interpretation of the middle alternative is simply that we do not know, and this is not only possible, but is actually the present situation regarding many unsolved problems.

The term "omniscience" is used in naming these principles to remind us that we are not omniscient! The position of LPO in Example 1 is now clear; we (effectively) showed first that the least upper bound principle implies LPO, and then we showed that LPO implies a solution to Fermat's Last Problem. It is convenient to state and prove explicitly the first part of Example 1 as follows:

Example 1*. The *least upper bound principle* is nonconstructive; it implies LPO.

Proof. Let $\{a_n\}$ be any decision sequence, and let S denote the set of values of its terms. The set S has at least one element, and at most two, but in general we do not know exactly how many; in any event it is a nonvoid bounded set of real numbers. Using the least upper bound principle as an hypothesis, let t denote the least upper bound of S. We need only calculate a rational approximation within $1/6$ to tell whether the real number t is less than $2/3$ or more than $1/3$. (For more details, see the Constructive Dichotomy Lemma below.) In the first case, it is clear that each term in the given decision sequence is zero. (This does not mean that we actually calculate all the infinitely many terms of the sequence and check each one, but rather that we are able to *predict*, with absolute certainty, that no matter how many terms may be calculated, no matter by whom, and no matter how far into the future, each term will turn out to be 0.) In the second case, using the definition of least upper bound, we construct a number x in S that is more than $1/3$. Since this number x must be a term of the given decision sequence, it must be equal to 1. Thus we have arrived at one or the other of the two alternatives stated in LPO.

Similarly, we restate Example 2 as follows; the proof is left as an exercise.

Example 2*. The *principle of trichotomy of real numbers* is nonconstructive; it implies LPO.

The power of LPO It is not the situation that if Fermat's Last Theorem is proved tomorrow, then the least upper bound principle would suddenly be constructive. Example 1* shows not only that the least upper bound principle implies a solution to Fermat's Last Problem, as shown in Example 1, but that it implies LPO, which would yield solutions to *hundreds* of unsolved problems. Here we'll give only a few examples.

Example 5. LPO implies solutions to each of the following problems:
(a) Fermat's Last Problem.
(b) The Perfect Number Problem.
(c) The Goldbach Conjecture.
(d) The Riemann Hypothesis.

Proof. (a) We showed above how Fermat's Last Problem may be encoded as a decision sequence $\{a_n\}$, such that if all $a_n = 0$, then Fermat's Last Theorem is true, but if some term $a_k = 1$ is ever calculated, it will lead to a counterexample. Thus LPO solves the problem.

(b) and (c) are left as exercises.

(d) The Riemann Hypothesis is a long-standing, unsolved problem involving the *Riemann zeta function* $\zeta(s)$ of the complex variable s; this function plays an important role in the theory of prime numbers [39, pp. 424–431], [43]. The hypothesis states that, aside from a sequence of "trivial" roots, each root $s = \sigma + it$ of $\zeta(s)$ lies on the vertical line $\sigma = 1/2$. For each positive integer n, a finite calculation allows one to determine either that $|\sigma - 1/2| < 1/n$ for all nontrivial roots with $|t| < n$, or that

$|\sigma - 1/2| > 0$ for some such root. Define $a_n = 0$ or $a_n = 1$, accordingly. If all $a_n = 0$, then for each nontrivial root $s = \sigma + it$ we have $|\sigma - 1/2| < 1/n$ for all $n > |t|$, and therefore $\sigma = 1/2$; this proves the Riemann Hypothesis! On the other hand, if some $a_n = 1$, then we have a counterexample.

These few examples should suffice. Most problems in number theory and analysis can be encoded as decision sequences to which LPO applies. For some problems it is a bit more difficult. For example, the question of whether or not there are infinitely many twin primes leads us to construct a sequence of 0's and 1's (not nondecreasing) and to ask whether there are infinitely many 1's. LPO answers this question also; see the section below on the Bolzano-Weierstrass Principle.

Discontinuous functions and WLPO A close look at Example 3 leads us to formulate another omniscience principle, and to restate the example.

WEAK LIMITED PRINCIPLE OF OMNISCIENCE (WLPO). Given any decision sequence $\{a_n\}$, there is a finite procedure which produces either a proof that $a_n = 0$ for all n, or a proof that "$a_n = 0$ for all n" is contradictory.

Example 3*. The *discontinuous function principle* is nonconstructive; it implies WLPO.

The power of WLPO is less than that of LPO, although it is enough to establish the nonconstructivity of certain classical theorems. The conjectures listed in Example 5 can each be encoded into a decision sequence $\{a_n\}$ such that the conjecture is true if and only if all $a_n = 0$. For any one of these, WLPO would provide a finite procedure leading either to a proof of the conjecture or to a proof of its falsity, but without a counterexample. Although this would not settle the problem completely, such an application of WLPO would be of great pragmatic value. It would either give you a proof, or show that a proof was impossible, in which case you could give up trying to find a proof, and concentrate further efforts on the search for a counterexample. Thus Brouwerian counterexamples using WLPO are sufficient to indicate the nonconstructivity of certain classical theorems and "constructions."

The intermediate value theorem and LLPO Considering Example 4, one might reasonably ask "What does it matter whether the first sequence 123456789 (if any) in the digits of π ends at an even or an odd place?" This is another good reason to free Brouwerian counterexamples from such *ad hoc* considerations. A close look at Example 4 shows that the intermediate value theorem leads to another omniscience principle.

LESSER LIMITED PRINCIPLE OF OMNISCIENCE (LLPO). Given any decision sequence $\{a_n\}$, there is a finite procedure which predicts whether the first integer k (if any), such that $a_k = 1$, is even or odd.

Thus the first part of Example 4 may be expressed as follows.

Example 4*. The *intermediate value theorem* is nonconstructive; it implies LLPO.

Proof. Let $\{a_n\}$ be a decision sequence. This sequence has no -1's, as did the sequence used in Example 4, and thus we use here a different definition for β:

$$\beta = \sum_{n=1}^{\infty} \frac{(-1)^n a_n}{10^n}.$$

The definition of the function f is the same as before; FIGURE 4 gives us an omniscient view of its properties. The intermediate value theorem, if constructive, would give an explicit procedure for the construction of a point x with $f(x) = 0$. It is convenient to use the constructive dichotomy lemma (see the next section below); thus either $x < 2/3$ or $x > 1/3$. If $x < 2/3$, then it is clear that β cannot be positive, and it follows that the first integer k such that $a_k = 1$ (if any) would be odd. On the other hand, if $x > 1/3$, then k would be even.

The power of LLPO is even less than that of WLPO, but still sufficient to demonstrate the nonconstructivity of certain classical theorems, because there is in fact no such finite procedure available to us. The relevance of LLPO for Brouwerian counterexamples depends on the fact that people agree that the discovery of such a general finite procedure seems extremely unlikely, even impossible.

Constructive constructions It is time to discuss the basic constructive properties of real numbers and functions which were used in the above counterexamples. The title of this section reminds us of the fact that in classical mathematics the term "construction" appears frequently, but rarely in the sense used here.

For a complete description of the construction of the real numbers, one must refer to Chapter 2 of [3] or [6], and for the construction of the extended real numbers, to [29]. Here we consider briefly only a few of the most important concepts. A real number is a Cauchy sequence of rational numbers. Thus, given any real number, arbitrarily close rational approximations are always available. The notion of *positive real number* is crucial, and closely connected with the idea of constructive existence. When we say that a real number x is *positive*, we mean that we have explicitly constructed a positive integer k and a rational approximation q within $1/k$ of x, and have proved that $q > 1/k$. (Since q is a quotient of integers, this is a good example of an application of the fundamental constructivist thesis, that all concepts should be reduced to elementary calculations with the integers.) Thus, to prove a real number x is positive requires a concrete construction; a proof that $x \leqslant 0$ is contradictory will not suffice. (On the other hand, the definition of $x \leqslant 0$, while given in an affirmative manner, is equivalent to the statement that $x > 0$ is contradictory; see [3, Lm. 5, p. 24] or [6, Ch. 2, 2.18].)

With this definition of *positive*, and the resultant notion of strict inequality, we consider now one of the most frequently used constructive properties of real numbers. It is the main constructive substitute for trichotomy, and reflects the essence of a real number, given only by approximations.

CONSTRUCTIVE DICHOTOMY LEMMA. *If a and b are real numbers with $a < b$, then for any real number x, either $x < b$ or $x > a$.*

Proof. The given condition $a < b$ means that $b - a > 0$; thus a positive integer k can be constructed with $b - a > 2/k$. It follows that there is a rational number s such that $a + 1/k < s < b - 1/k$. Choose a rational approximation r to within $1/k$ of the given real number x. Since trichotomy does hold for the rational numbers, we have either $r \leqslant s$, in which case $x < b$, or we have $r > s$, in which case $x > a$. For more details, see [3, Cor., p. 24] or [6, Ch. 2, 2.17]. The lemma is illustrated in FIGURE 5.

Constructive definitions for convergence of sequences and series are also straightforward. One says that a sequence $\{x_n\}$ of real numbers *converges* to a real number x if one has constructed a sequence $\{N_k\}$ of *convergence parameters* with the property that $|x - x_n| < 1/k$ whenever $n \geqslant N_k$. This is exactly the same as the classical definition, except that classically one says something like "for all k there exists N_k

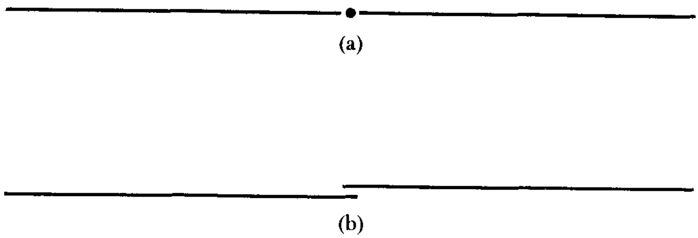

FIGURE 5
Classical trichotomy vs. constructive dichotomy of the real numbers. In the omniscient view (a), the real line is divided precisely into three distinct nonoverlapping parts. The constructive view (b) reveals only two cases. They are not mutually exclusive; however, the region of overlap may be as small as desired, according to the degree of precision required.

such that ..." without being careful to explain in what sense "there exists" is to be understood. Classically one allows convergence to be proved by assuming that such parameters N_k do not exist and deriving a contradiction, while constructively these parameters must be explicitly constructed by means of a finite procedure. The convergence of an infinite series reduces in the usual way to convergence of the sequence of partial sums.

We can now justify the definition of a continuous function by means of straight lines, as in Examples 4 and 4*. We have the function f defined clearly enough on the three subintervals $[0,1/3]$, $[1/3,2/3]$, and $[2/3,1]$. However, these subintervals do not constitute the entire interval $[0,1]$, because we have no finite procedure which determines in which subinterval a given point lies. (A slight modification of Example 2* shows that such a procedure would imply LPO.) We give only a brief sketch of the definition of f. The important conditions, which do hold here, are that uniformly continuous functions are used on each subinterval, and that they connect properly. To define f at an arbitrary point x of $[0,1]$ means to give an approximation to $f(x)$ to within ε, for any $\varepsilon > 0$. We may assume that $x < 2/3$. The other case given by the dichotomy lemma, when $x > 1/3$, is similar. Using the dichotomy lemma again, we have either $x < 1/3$ or $x > 1/3 - \varepsilon/6$. In the first case we have the value of f already defined, while in the second case it suffices to use β as the required approximation to $f(x)$. We leave the remaining details, including the continuity of f, as an exercise for the reader (who may wish to consult Chapter 2 of [3] or [6]). An alternative method for constructing such functions is given in Example 14.28 of [31].

Using the constructive dichotomy lemma, we can replace many nonconstructive classical theorems by constructive substitutes which are fully adequate for the constructive development of analysis. The intermediate value theorem is replaced by a constructive theorem which, given any small positive number ε, constructs a point x at which $|f(x)| < \varepsilon$. This is [3, Ch. 2, ex. 11], or [6, Ch. 2, 4.8]; we give only a brief sketch of the proof, leaving the reader to fill in the details. Use the uniform continuity of f to construct $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$, and divide the interval into n subintervals, each of length less than δ. At each subdivision point the dichotomy lemma determines either that f has value less than ε or that f has value more than 0. At the ends of at least one subinterval we must obtain opposite determinations; at the left end of the first such subinterval we find a suitable point x. Thus we have the following.

THEOREM 1. *If f is a uniformly continuous function on a closed bounded interval $[a, b]$, with $f(a) > 0$ and $f(b) < 0$, then for any $\varepsilon > 0$ there exists a point x in the interval such that $|f(x)| < \varepsilon$.*

As for any theorem in constructive mathematics, the phrase "there exists" appearing above is to be understood in the strict constructive sense; in this case, from the definition of f and the other data given, an explicit finite procedure is obtained which constructs the point x.

Sometimes the intermediate value theorem is also expressed in the classically equivalent form of its contrapositive, as, for example, in [20, p. 80]. Constructively, the contrapositive of a statement follows from the statement, but is not an equivalent. One form of contrapositive of Theorem 1 states that a function bounded away from zero cannot assume opposite signs at the endpoints. We prove a stronger version of this in the next theorem, which seems not to have appeared previously. A function f is said to be *never zero* on a set S if $|f(x)| > 0$ for every x in S; it is said to *have constant sign* on S if $f(x) \cdot f(y) \geqslant 0$ for all x and y in S.

THEOREM 2. *If a function f is continuous and never 0 on an interval $[a, b]$, then f has constant sign on $[a, b]$.*

Proof. Since $|f(a)| > 0$, it follows that either $f(a) > 0$ or $f(a) < 0$ (use a rational approximation). We need consider only the case $f(a) > 0$; the other case is similar. Let y be any point of the interval; either $f(y) > 0$ or $f(y) < 0$. Suppose the second case occurs; then the sets $U = \{x : f(x) < 0\}$ and $V = \{x : f(x) > 0\}$ are nonvoid open subsets of $[a, b]$ which cover the entire interval. By the constructive connectivity theorem [27, Thm. 2], these sets have a common point, which is absurd. Thus the first case must obtain, and $f(y) > 0$.

Other constructive forms of the intermediate value theorem are also available; some are listed in [3, p. 59] and [6, p. 63]. This multiplicity of constructive forms is typical. After the fracturing of a classical theorem by a Brouwerian counterexample, constructive workers pick up the pieces and remold them into a number of different constructively valid theorems, each displaying a different aspect of the situation.

The least upper bound principle also has a powerful constructive substitute. We restrict the theorem to totally bounded sets. A set S is said to be *totally bounded* if, given any $\varepsilon > 0$, we can construct a finite subset F such that every point of S lies within ε of some point of F. The maximum of the points in F gives us an approximation to the least upper bound of S to within ε. Since giving arbitrarily close approximations to a real number is equivalent to defining it, we say that the least upper bound of S exists. We state the result as follows.

THEOREM 3. *Every nonvoid totally bounded set of real numbers has a least upper bound and a greatest lower bound.*

The details are in [3, Thm. 3, p. 34] or [6, Ch. 4, 4.3]. For least upper bounds and greatest lower bounds in the extended real number system, see [28] and [29]; for alternative constructive notions, see [31, sec. 4].

The procedures produced by the above theorems might be quite lengthy; for some comments on this, see [3, p. 3] or [6, p. 6], and [5]. In an important application, rather than giving a solution on a hand calculator in a half a minute, the procedures could lead instead to years of work trying to write programs efficient enough to produce the solution on a large computer in only a month. Nevertheless, we maintain the distinction between an infinite calculation, which we have absolutely no hope of actually carrying out, and a finite process, however long. The important questions on the efficiency of procedures belong to the second phase of the constructivization of mathematics. It is too soon to demand progress on this problem from the very few present-day constructivists. More help is needed—invited!—urgently awaited!—perhaps it will come from among the readers of this MAGAZINE.

Twin primes and the Bolzano-Weierstrass theorem A famous unsolved problem in number theory is whether or not there are infinitely many twin primes of the form p and $p+2$, such as 3 and 5, 5 and 7, 11 and 13,..., 209267 and 209269, etc. See, for example [17, p. 31] and [42, Ch. 1]. Although one might wish to utilize this problem in a Brouwerian counterexample, a difficulty arises. At first glance neither LPO nor the other principles mentioned above seem to imply a solution to this problem. The *Bolzano-Weierstrass principle*, although unrelated to any problem in number theory, presents some similar difficulties; although it clearly implies LPO, the converse is not immediate. The difficulties, which stem from the need to consider arbitrary sequences of positive integers, are resolved by the following.

THEOREM 4. *The following are equivalent.*

(a) LPO. Limited Principle of Omniscience. *For any decision sequence* $\{a_n\}$, *either all* $a_n = 0$ *or some* $a_n = 1$.

(b) BSP. Bounded Sequence Principle. *Any sequence of positive integers is either bounded or unbounded.*

(c) KSP. Constant Subsequence Principle. *Any bounded sequence of positive integers has a constant subsequence.*

(d) BWP. Bolzano-Weierstrass Principle. *Any bounded sequence of real numbers has a convergent subsequence.*

(e) MSP. Monotone Sequence Principle. *Any bounded monotone sequence of real numbers converges.*

We give here only a sketch of the proof that the *Constant Subsequence Principle* implies the *Bolzano-Weierstrass Principle*. The proof of BWP proceeds in essentially the same manner as in most elementary analysis texts, by interval-halving. Using the constructive dichotomy lemma, we may assume that the "halves" of the interval have a small overlap, and thus one can tell in which of these halves each term of the given sequence $\{x_n\}$ lies. Constructive danger is first sighted at the point of deciding which half contains infinitely many terms of the sequence. Here KSP navigates an omniscient course. Define $p_n = 1$ or $p_n = 2$ according as x_n lies in the left or right half. Since the sequence $\{p_n\}$ is bounded, KSP provides a constant subsequence; this corresponds to a subsequence of $\{x_n\}$ which lies wholly in one half of the interval, towards which we should steer. The reader can try the rest of the proof as an exercise, or refer to [33].

We now consider the twin prime problem; extending Example 5.

Example 6. LPO implies a solution to the twin prime problem.

Proof. Construct a sequence $\{p_n\}$ of positive integers as follows; if both n and $n+2$ are prime, define $p_n = n$, and otherwise $p_n = 1$. There are infinitely many twin primes if and only if the sequence $\{p_n\}$ is unbounded. Thus the Bounded Sequence Principle would yield a solution to the twin prime problem.

Because of this example, one might say that the Bolzano-Weierstrass Principle and the Monotone Sequence Principle are nonconstructive because they each imply a solution to the Twin Prime Problem. The Bounded Sequence Principle may also be used in connection with other questions in number theory, such as whether or not there exist infinitely many even numbers which are sums of two primes.

The limited principle of existence Brouwerian counterexamples utilizing the nonconstructive omniscience principles LPO, WLPO, and LLPO quite clearly demonstrate the nonconstructivity of classical theorems, because there are unsolved problems in analysis and number theory for which these principles would yield solutions or information not actually available. There are other classical theorems, however, which imply principles involving decision sequences which certainly seem nonconstructive, and for which no proofs are at hand, and yet for which there are no known unsolved problems whose solution they would provide. Some of these principles relate to very central problems in constructive mathematics. In this section we discuss one of these, LPE, which involves fundamental properties of the real numbers, and in a later section another, WLPE, which is related to the constructivity of continuity theorems.

THE LIMITED PRINCIPLE OF EXISTENCE (LPE). Given any decision sequence $\{a_n\}$, for which it is contradictory that $a_n = 0$ for all n, there is a finite procedure which results in the construction of an integer n such that $a_n = 1$.

The contrast of conditions in LPE evokes the sharp antithesis between classical and constructive mathematics; between pseudoexistence, derived from a proof by contradiction, and constructive existence, derived from an explicit finite process. However, we have no example of an unsolved problem whose solution would be given by LPE. Thus a counterexample involving LPE, such as in the next section below, provides less conclusive evidence of nonconstructivity than one involving the other omniscience principles. Still, a finite procedure as specified in LPE seems, from a constructive viewpoint, very unlikely. In any event, a counterexample involving LPE has pragmatic value in that it tends to limit further efforts to prove the conjecture and intensifies efforts to find a Brouwerian counterexample in the strict sense. LPE is sometimes referred to as Markov's Principle. In recursive function theory, in contrast to the strict Bishop-type constructive mathematics discussed in this paper, arguments are made for the plausibility of LPE, and it is often used as an axiom.

Consider a decision sequence $\{a_n\}$ of the sort considered in LPE: it is contradictory that $a_n = 0$ for all n; yet we have no proof that there exists an integer n such that $a_n = 1$. Using the method of the above counterexamples, we obtain a real number

$$\alpha = \sum_{n=1}^{\infty} \frac{a_n}{10^n},$$

which is clearly $\geqslant 0$ but cannot be 0, because then all the terms a_n in the decision sequence would be necessarily 0. On the other hand, we have no proof that $\alpha > 0$, for this would mean that we had constructed some term $a_n = 1$. Since α bears to 0 a relation not covered by the conventional terminology and symbols, we adopt the following.

Definition. A real number α will be said to be *almost positive* when $\alpha \geqslant 0$ and it is contradictory that $\alpha = 0$. This condition will be written $\alpha \cdot\!> 0$.

With this terminology, LPE has a simple expression: *every almost positive real number is positive.*

Irrational numbers When forming definitions for the constructive development of mathematics, one has a fairly wide choice. This is because the classical definitions typically have a variety of classically equivalent, but constructively quite distinct, formulations, and one must exercise great care in choosing a definition with useful numerical meaning. One example, in a sense the first to arise following the generation of the real number system out of the rationals, is the definition of irrational number. Consider a real number x, and two forms of the classical definition:

(i) For every rational number q, the equality $x = q$ is contradictory.
(ii) For every rational number q, the inequality $|x - q| > 0$ holds.

Although these two conditions are classically equivalent, they are constructively quite different, as the following example shows.

Example 7. The statement "*If x is a real number such that the equality $x = q$ is contradictory for every rational number q, then $|x - q| > 0$ for every rational q*" is nonconstructive; it implies LPE.

Proof. Assuming the statement in quotes, we will obtain a proof of LPE. Let $\{a_n\}$ be a decision sequence such that it is contradictory that all $a_n = 0$, and let x be the real number defined by

$$x = \sum_{n=1}^{\infty} \frac{a_n\sqrt{2}}{2^n}.$$

First we show that x satisfies condition (i). Let q be a rational number and suppose that $x = q$. Since trichotomy holds for the rationals, we have either $x = 0$ or $x > 0$. In the first case it follows that all $a_n = 0$, a contradiction. In the second case, the decision sequence becomes constantly 1 from some point on, and the sum of the series has the form $x = \sqrt{2}/2^k$ where k is the largest integer such that $a_k = 0$. But such a value for x cannot be rational, so again we have a contradiction. Thus condition (i) is satisfied. By hypothesis, x satisfies condition (ii). Since $q = 0$ is rational, it follows that $x > 0$, and there exists an integer n such that $a_n = 1$. This proves LPE.

Nevertheless, even with the strict definition (ii), one has no difficulty finding a plentiful supply of irrational numbers. For example, the usual proof that $\sqrt{2}$ is irrational, while not constructive, becomes so with only a little more care. The familiar classical proof, that for any integer p and any integer $q \neq 0$, the numbers p^2/q^2 and 2 are distinct, is constructively valid; it involves only integers, about which there are no constructive complications. The definition of $\sqrt{2}$ presents no difficulties; for example, decimal approximations suffice, since the precise rule for the determination of each approximation involves only finite decimals. To show that $\sqrt{2}$ is irrational in the strong sense of condition (ii) above, we must show that for any p and q, the inequality $|p/q - \sqrt{2}| > 0$ holds; this means we must construct a positive integer k with $|p/q - \sqrt{2}| > 1/k$. It suffices to consider the case $1 < p/q < 2$, for otherwise one simply takes $k = 3$. What the traditional proof actually shows, using Euclid's *Fundamental Theorem of Arithmetic*, is that p^2 and $2q^2$ are distinct integers (since there are an even number of factors 2 in the unique prime decomposition of p^2, but an odd number in the decomposition of $2q^2$). Thus these integers differ by at least 1, and we have

$$|p/q - \sqrt{2}| \cdot |p/q + \sqrt{2}| \cdot q^2 = |p^2 - 2q^2| \geqslant 1.$$

Since $0 < |p/q + \sqrt{2}| < 4$ it follows that

$$|p/q - \sqrt{2}| > 1/4q^2$$

and we may take $k = 4q^2$. This is typical of much of classical mathematics, which, along with many nonconstructivities, does contain a vast amount of numerical meaning which merely needs to be made explicit, although this usually requires more effort than in this example.

Omniscience principles and real numbers Each of the four omniscience principles so far discussed has an equivalent formulation involving the ordering of the real numbers. It is convenient to have these formulations available, for ease in constructing counterexamples, and for finding relationships between the various omniscience principles.

THEOREM 5. *Each of the omniscience principles listed below has the two equivalent formulations indicated.*

(a) The Limited Principle of Omniscience (LPO).
 (i) *For any decision sequence* $\{a_n\}$, *either all* $a_n = 0$ *or some* $a_n = 1$.
 (ii) *For any real number* x, *either* $x \leqslant 0$ *or* $x > 0$.

(b) The Weak Limited Principle of Omniscience (WLPO).
 (i) *For any decision sequence* $\{a_n\}$, *either all* $a_n = 0$ *or it is contradictory that all* $a_n = 0$.
 (ii) *For any real number* x, *either* $x \leqslant 0$ *or* $x \cdot > 0$.

(c) The Lesser Limited Principle of Omniscience (LLPO).
 (i) *For any decision sequence* $\{a_n\}$, *either the first integer* k (*if any*), *such that* $a_k = 1$, *is even, or it is odd.*
 (ii) *For any real number* x, *either* $x \leqslant 0$ *or* $x \geqslant 0$.

(d) The Limited Principle of Existence (LPE).
 (i) *For any decision sequence* $\{a_n\}$, *if it is contradictory that all* $a_n = 0$, *then some* $a_n = 1$.
 (ii) *For any real number* x, *if* $x \cdot > 0$, *then* $x > 0$.

From this theorem it follows that LPO implies WLPO, that WLPO implies LLPO, and that LPO is equivalent to WLPO and LPE combined. The theorem is easily verified using the following two lemmata connecting real numbers and decision sequences.

LEMMA 1. *For any real number* x *there exists a corresponding decision sequence* $\{a_n\}$ *such that*
 (i) $x \leqslant 0$ *if and only if all* $a_n = 0$
 (ii) $x > 0$ *if and only if some* $a_n = 1$.

Conversely, for any decision sequence $\{a_n\}$ *there exists a corresponding real number* x *satisfying these two conditions.*

Proof. Let x be a given real number. For each positive integer n, the constructive dichotomy lemma provides a finite procedure which results in one of two conclusions, either $x < 1/n$, or $x > 0$. Define $a_n = 0$ or $a_n = 1$ accordingly, continuing with the later choice once it occurs; this defines a decision sequence $\{a_n\}$. If all $a_n = 0$, the dichotomy lemma always leads to the first alternative; thus $x < 1/n$ for all n, and it follows that $x \leqslant 0$. The converse is clear. If $x > 0$, then there exists an integer n such that $x > 1/n$. At the nth step in the construction of the decision sequence, the second alternative is necessitated; thus some $a_n = 1$. The converse to this is also clear.

Conversely, given a decision sequence $\{a_n\}$, define

$$x = \sum_{n=1}^{\infty} \frac{a_n}{2^n}.$$

The two conditions are easily verified.

The next lemma is useful in connection with LLPO. The strangely hypothetical condition "the first integer k (if any), such that $a_k = 1$, is even" is more conveniently expressed by the straightforward affirmative condition "$a_n = a_{n+1}$ for all even n".

LEMMA 2. *For any real number* x *there exists a corresponding decision sequence* $\{a_n\}$ *such that*

(i) $x \leqslant 0$ *if and only if* $a_n = a_{n+1}$ *for all even* n
(ii) $x \geqslant 0$ *if and only if* $a_n = a_{n+1}$ *for all odd* n.
Conversely, for any decision sequence $\{a_n\}$ *there exists a corresponding real number* x *satisfying these two conditions.*

Proof. For a given real number x, the construction of a suitable decision sequence is left as an exercise; see [31, sec. 2.6]. Conversely, given a decision sequence $\{a_n\}$, define

$$x = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} a_n}{n}.$$

The elementary topology of the real line A typical theorem found in elementary analysis courses is the following; see, for example [18, 3.19.6].

The Open Interval Hypothesis. Any nonvoid open set on the real line is the union of an at most countable family of disjoint open intervals.

Example 8. The *open interval hypothesis* is nonconstructive, it implies LPO.

Proof. The hypothesis is nonconstructive even if limited to *bounded* sets, and even if *any* family of disjoint open intervals, not necessarily countable, is allowed. We assume the open interval hypothesis and derive a proof of LPO. It is convenient to use Theorem 5; thus for any real number x we must show either $x \leqslant 0$ or $x > 0$. Define $y = |x|$; then $y \geqslant 0$ and it will suffice to show that either $y > 0$ or $y = 0$. Using the dichotomy lemma, we may assume $y < 1$, for if $y > 0$ then there is nothing more to prove. Consider the set

$$U = (-1, y) \cup (0, 1).$$

Clearly U is an open set. (In fact, it is a countable, even finite, union of open intervals, but they are not necessarily disjoint. Disjointness is the crucial part of the statement being tested; it determines the connected components of a set.) By hypothesis, U has a decomposition $U = \bigcup_\alpha I_\alpha$ into disjoint open intervals. Since the point $1/2$ lies in U, it must lie in one of the intervals I_α; this interval has the form (a, b). Applying the constructive dichotomy lemma, either $a < 0$ or $a > -1$. In the first case the point 0 lies in U, and it follows that $y > 0$, while in the second case it follows that $y = 0$. This proves LPO. The set U is illustrated in Figure 6.

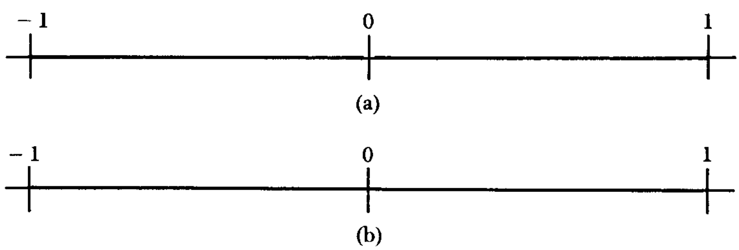

FIGURE 6.
Two views of the open set $U = (-1, y) \cup (0, 1)$ used in Example 8. If the number y is 0, then U consists of two disjoint open intervals, but if y is positive, then U is the entire interval $(-1, 1)$. We have no way to determine that one of these cases applies. What are the connected components of U?

However, most constructively important open sets on the line *can* be resolved into countably many disjoint open intervals; see [**29**] and [**30**]. For other open problems on the topology of the line in need of constructive work, see [**10**] and [**31**].

Counterexamples in algebra A well-known classical theorem in elementary abstract algebra states that every field has a characteristic which is either 0 or a prime p. (The characteristic is the least positive integer m such that $m \cdot e = 0$, where e is the identity of the field, if such an integer exists, and 0 when no such positive integer m exists. See, for example [**2**, pp. 386–392] or [**45**, pp. 91–93].)

Example 9. The statement "*Every field has a characteristic which is either* 0 *or a prime*" is nonconstructive; it implies LPO.

Proof. Let $\{a_n\}$ be a decision sequence, let $\{p_n\}$ be the ordered sequence of all positive primes, and define

$$A = \{0\} \cup \{p_k : a_{k-1} < a_k\}.$$

Note that A has at least one element and at most two, but we do not know which. (The inequality is a device used with decision sequences to indicate conveniently "the first k such that $a_k = 1$ (if any)" and, for the sake of uniformity in counterexamples, to alleviate the need for "sequences of 0's and 1's with at most one term equal to 1".) In the domain $\mathbb{Z}$ of integers, let P be the subring generated by A, and let D be the quotient ring $\mathbb{Z}/P$. Then D is an integral domain; the proof (a rocky shallows) is left for the reader. Thus D has a quotient field F. Apply now the hypothesis that F has a characteristic. If this characteristic is 0, then in the decision sequence all $a_n = 0$. But if this characteristic is a prime, with position k in the sequence $\{p_n\}$, then $a_k = 1$. Thus the hypothesis implies LPO.

Whether or not D is itself a field is also a mystery. For more details, and some related positive constructive results, see [**25**] and [**35**]. Using, in place of $\{p_n\}$, the sequence of primes having residue 1 modulo 4, and considering the polynomial $x^2 + 1$, one may prove the following. (Hint: see [7, Thm. 13.2].)

Example 10. The statement "*Every polynomial over a field is either irreducible or can be factored into irreducible polynomials*" is nonconstructive; it implies LPO.

Continuity One of the oldest constructivity problems is whether or not every real-valued function on the closed unit interval is continuous. We formulate this problem as follows.

Continuity Principle (CP). Every real-valued function on the closed unit interval is continuous.

The typical classical counterexample to this principle is constructively invalid, as shown above in the section on discontinuous functions. Brouwer [**13**] proved CP. His proof, however, was not constructive in the strict sense; it used methods of questionable constructivity (which are still used in intuitionistic mathematics). Expositions of Brouwer's proof may be found in [**19**, Ch. 3] and [**23**, Ch. 3].

Since CP is classically false, we consider only the following (classically true) weak mutation of CP:

Limited Continuity Principle (LCP). Every real-valued function on the closed unit interval, which is nondecreasing and approximates intermediate values, is continuous.

By *nondecreasing* we mean that $f(x) \leqslant f(y)$ whenever $x \leqslant y$. *Approximates intermediate values* means that if $f(x) < \lambda < f(y)$ and $\varepsilon > 0$, then there exists a point z in the interval such that $f(z)$ is within ε of λ; (cf. Theorem 1). Classically, for nondecreasing functions, this is equivalent to attaining intermediate values exactly. Thus LCP is a partial converse to the intermediate value theorem; classical proofs are found in some calculus texts, for example [**22**, p. 192].

A Brouwerian counterexample to LCP is available, but, as for the counterexample above concerning the irrationals, not in the strict sense. It does not relate LCP to an unsolved problem, but to a principle, which, in the fashion of LPE, only "seems" nonconstructive. Thus we enter here into recent results about which there is room for difference of opinion. The principle involved is a weak form of LPE, and a converse to a weak form of constructive dichotomy. An equivalent form of the constructive dichotomy theorem is: if $c > 0$, then for any real number x, either $x > 0$ or $x < c$. A weak form of this is: if $c > 0$, then for any real number x, either $x \cdot > 0$ or $x < \cdot c$ (for this notation see the section on LPE above). The converse to this is the following.

The Weak Limited Principle of Existence (WLPE). If c is a real number such that, for any real number x either $x \cdot > 0$ or $x < \cdot c$, then $c > 0$.

Any real number c which satisfies the condition "for any real number x either $x \cdot > 0$ or $x < \cdot c$" is said to be *pseudo-positive*. Thus WLPE has the simple expression "every pseudo-positive real number is positive." Clearly, any pseudo-positive real number is almost positive; thus LPE implies WLPE. For a more complete discussion of LPE and its weaker versions, see [**34**].

The structure of WLPE strongly suggests that it is nonconstructive; it purports to derive quite affirmative information, the construction of specific integers which demonstrate that c is positive, from a mere dichotomy of negativistic conditions. However, this is somewhat speculative and WLPE will require more time for a definitive evaluation. Nevertheless, taking the notion of *nonconstructive* in the broad sense, that a classical theorem is reduced to an elementary, simply expressed, classical property of the real numbers which, from our extensive experience with similar properties, seems to preclude all possibility of constructive proof, we have the following counterexample to LCP. For the proof, see [**31**, Thm. 16.5].

Example 11. The *limited continuity principle* is nonconstructive; it is equivalent to WLPE.

Appendix. Negativistic counterexamples vs. positive constructive developments

There is a danger in devoting an entire paper to negativistic counterexamples. As Errett Bishop has written, "The counterexamples are deceitful. The reader is asked not to form the impression that the purpose of constructive mathematics is to consider pathological numbers. The only reason for discussing such numbers is to show that certain statements are not constructively valid" [**3**, p. 60] [**6**, p. 65]. The danger is less today, however, than in 1967 when the above quote appeared. Prior to that time one had only Brouwer's critique (providing the crucial motivation for the constructivization of mathematics), certain intuitionistic results (often mixed with nonconstructive elements such as free choice sequences), some idealistic logical considerations using formal systems (an approach diametrically opposed to Bishop's strict constructivist thesis), and results in recursive function theory (extensive, but only semi-constructive, because of restricted concepts of number and function, and some use of nonconstructive reasoning). But until 1967 there were few systematic, strictly constructive advances. Thus it was important that Bishop try to correct the prevalent misunderstandings. At this time, however, we have available Bishop's monumental work [**3**], which constructivizes a large portion of analysis, and indicates the direction for further positive constructive work.

The purpose of this article is to describe Brouwer's critique of 80 years ago, showing the nonconstructivities in classical mathematics. This critique must be evaluated in the crucially different present-day context, in which there are available not only powerful methods for the constructive development of mathematics, but also sufficient examples of their application. The references below include only a few recent constructive advances; their bibliographies provide more extensive references.

REFERENCES

1. G. Birkhoff, ed., *A Source Book in Classical Analysis*, Harvard University Press, Cambridge, 1973.
2. G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, rev. ed., MacMillan, New York, 1953.
3. E. Bishop, *Foundations of Constructive Analysis*, McGraw-Hill, New York, 1967. (To be reprinted by Springer Verlag, Berlin, with an introduction by G. Stolzenberg.)
4. ______, *Aspects of constructivism*, Notes on lectures given at New Mexico State University, Las Cruces, 1972.
5. ______, Schizophrenia in contemporary mathematics (1973), in *Errett Bishop; Reflections on Him and His Research*, ed. by M. Rosenblatt, Amer. Math. Soc., Providence, 1985.
6. E. Bishop and D. Bridges, *Constructive Analysis*, (Grundlehren der mathematischen Wissenschaften, 279), Springer, Berlin, 1985.
7. E. D. Bolker, *Elementary Number Theory, an algebraic approach*, Benjamin, New York, 1970.
8. B. Bolzano, Rein analytischer Beweis des Lehrsatzes, dass zwischen je zwey Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege, Abh. der Kgl. Ges. der Wiss. Prague, 1817, 1–60.
9. D. Bridges, *Constructive Functional Analysis*, Pitman, London, 1979.
10. ______, Connectivity properties of metric spaces, *Pacific J. Math.* 80 (1979), 325–331.
11. D. Bridges and R. Mines, What is constructive mathematics?, *Math. Intelligencer* 6 (1984), no. 4, 32–38.
12. L. E. J. Brouwer, *Over de grondslagen der wiskunde*, Thesis, Amsterdam, 1907.
13. ______, Beweis, daß jede volle Funktion gleichmäßig ist, *Proc. Akad. Amsterdam* 27 (1924), 189–194.
14. ______, *Collected Works, vol. 1, Philosophy and Foundations of Mathematics*, ed. A. Heyting, North-Holland, Amsterdam, 1975.
15. ______, *Cambridge Lectures on Intuitionism*, Cambridge University Press, 1981.
16. Y. K. Chan, Notes on constructive probability theory, *Ann. Probability* 2 (1974), 51–75.
17. R. Courant and H. Robbins, *What is Mathematics?*, Oxford University Press, London, 1941.
18. J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York, 1960.
19. M. Dummett, *Elements of Intuitionism*, Clarendon Press, Oxford, 1977.
20. L. Gillman and R. H. McDowell, *Calculus*, Norton, New York, 1973.
21. I. Grattan-Guinness, *The Development of the Foundations of Mathematical Analysis from Euler to Riemann*, The MIT Press, Cambridge, 1970.
22. G. H. Hardy, *A Course of Pure Mathematics*, 10th ed., University Press, Cambridge, 1958.
23. A. Heyting, *Intuitionism, an introduction*, 3rd ed., North-Holland, Amsterdam, 1971.
24. W. Julian, ε-continuity and monotone operations, *Pacific J. Math.* 115 (1984), 385–390.
25. W. Julian, R. Mines, and F. Richman, Algebraic numbers, a constructive development, *Pacific J. Math.* 74 (1978), 91–102.
26. D. R. Kurepa, Around Bolzano's approach to real numbers, *Czech. Math. J.* 32 (1982), 655–666.
27. M. Mandelkern, Connectivity of an interval, *Proc. Amer. Math. Soc.* 54 (1976), 170–172.
28. ______, Suprema of located sets, *J. London Math. Soc.* (2), 20 (1979), 161–164.
29. ______, Resolutions on the line, *Pacific J. Math.* 88 (1980), 91–100.
30. ______, Components of an open set, *J. Austral. Math. Soc.* (*Series A*) 33 (1982), 249–261.
31. ______, Constructive Continuity, *Memoirs Amer. Math. Soc.* 277, (1983).
32. ______, Constructive Mathematics, this MAGAZINE 58 (1985), 272–280.
33. ______, Limited omniscience and the Bolzano-Weierstrass principle, *Bull. London Math. Soc.*, in press.
34. ______, Constructively complete finite sets, *Z. Math. Logik Grundlag. Math.*, in press.
35. R. Mines, Algebraic number theory, a survey, *The L. E. J. Brouwer Centenary Symposium*, North-Holland, Amsterdam, 1982, pp. 337–358.
36. J. Myhill, What is a real number?, *Amer. Math. Monthly* 79 (1972), 748–754.

37. F. Richman, Meaning and information in constructive mathematics, *Amer. Math. Monthly* 89 (1982), 385–388.
38. F. Richman, G. Berg, W. Julian, and R. Mines, The constructive Jordan curve theorem, *Rocky Mountain J. Math.* 5 (1975), 215–236.
39. S. Saks and A. Zygmund, *Analytic Functions*, Drukarnia Naukowa, Warsaw, 1952.
40. G. Stolzenberg, Review: Foundations of Constructive Analysis by E. Bishop, *Bull. Amer. Math. Soc.* 76 (1970), 301–323.
41. D. J. Struik, *A Source Book in Mathematics, 1200-1800*, Cambridge, 1969 (reprint, Princeton, 1986).
42. H. Tietze, *Famous Problems of Mathematics*, Graylock Press, Baltimore, 1965.
43. E. C. Titchmarsh, *The Theory of the Riemann zeta-function*, Oxford, 1951.
44. H. S. Vandiver, Fermat's Last Theorem, *Amer. Math. Monthly* 53 (1946), 555–578.
45. B. L. van der Waerden, *Modern Algebra*, vol. 1, Ungar, New York, 1949. (Original; *Moderne Algebra*, Springer, Berlin, 1931.)