Lecture 23

CS 4860

November 15, 2016

1 Announcements

Cornell's new president (appointed yesterday), Martha Pollack, is a CIS academic. She served as dean of Michigan's information science school, and her early career was in AI. I have worked with her and know her to be an excellent scientist and a visionary leader.

A week from today the class period will be devoted to a discussion of the major themes of the course and how to relate your projects to the themes.

2 Technical Topics for this Lecture

The "Great Schism" still exists but is more deeply understood, and signs of bridging it are emerging. The status is like this:

Constructive	Platonic ("Classical")
iFOL is well understood, but is not the foundational theory core. Type theory explains the computational meaning of iFOL and goes beyond, providing a broad "theory of computation." This theory is an adequate foundation for computer science and mathematics. Why do we trust it?	FOL is well understood, it supports ZFC set theory, the "default" foundation for "classical mathematics." How do we know it is consistent? What about Euclidean goemetry, is it consistent?

Hilbert proposed a bold program for "saving" classical mathematics by first *formalizing it* and then proving constructively that it is consistent. Gödel showed that this bold program could not work. It will not even work for the simple theory Q let alone PA or ZFC!

We will explore how to extend Heyting Arithmetic to provide an account of the constructive real numbers and relate that to our theory of constructive geometry. This will illustrate the methods of the constructive approach.

Arithmetic	Geometry
Real numbers $\rightarrow$	Computational model of axiomatic geometry

3 The Paradoxes

1. Set Theory

   - Russell's Paradox (1902)

     $$\text{Let } R = \{x : Set \mid x \notin x\}$$

     $$\text{Is } R \in R? \text{ Is } R \notin R?$$

   - Burali-Forti/Cantor (1897/1895) What is the cardinal number of the set of all sets? Let M be the *set of all sets*. $\overline{\overline{M}}$ is its cardinal number. $P(M)$ is the power set (also denoted **U**M, for the German untermenge). Cantor's theorem shows

     $$\overline{\overline{P(M)}} > \overline{\overline{M}} \text{ and also } P(M) \subset M, \text{ so } \overline{\overline{P(M)}} < \overline{\overline{M}}!$$

2. General

   - Richard Paradox: English sentences that define number theoretic functions of one varaible $f_0(n), f_1(n), f_2(n), \ldots$

     Consider $f_n(n) + 1$ expressed in English.

   - Berry Paradox (1906): "The least natural number not nameable in fewer than twenty-two syllables."

     In twenty-one syllables it names a number not nameable in fewer than 22 syllables.

   - This sentence is false.

     - if true, then it's false

     - if false, then it's true

   - Epimenides liar's paradox : "The statement I am now making is a lie."

4 Constructive Semantics

We explore a constructive semantics for Kleene's axioms for Heyting Arithmetic, HA, the constructive version of Peano Arithmetic, (PA).

We already gave an informal semantics for the propositional calculus in Lecture 3. If we leave out Kleene's "dreaded axiom 8," the one with the superscript, 8^0, the remaining axioms are constructively justified. We have discussed their computational meaning except for axiom 13, the induction principle. We see in the notes for the September 23^{rd} Nurpl demonstration one constructive account of induction. Essentially it is "realized" by primitive recursion.

- $A(0)$ is the *base case* for induction, and the value "computed" is the evidence for why $A(0)$ is constructively true. We discuss the nature of evidence for propositions below.
- $A(x) \Rightarrow A(s(x))$ is the *inductive case*, showing how to convert evidence $a(x)$ for $A(x)$ to evidence for $A(s(x))$, where Kleene writes x' for the successor of x, $s(x)$. Call this function f.

So the recursive function being used to justify and *implement induction* is this primitive recursive function:

$$ind(n) = \text{if } n = 0 \text{ then } a(0) \text{ else } f(ind(n-1)).$$

In Nuprl, we do not use the extremely inefficient *unary notation* for natural numbers, $0, s(0), s(s(0)), \cdots$. We use the "Big Nums" package in Lisp to represent the natural numbers and the integers (negative numbers as well). So the Nuprl induction rule can handle induction "going up" as well as going down, and it has clear computational meaning. We use these integers in defining the real numbers.

INT

formation

1. H >> Ui ext int by intro int
2. H >> int in Ui by intro

canonical

3. H >> int ext c by intro c
4. H >> c in int by intro

where c must be an integer constant.

noncanonical

5. H >> $-t$ in int
   >> t in int
6. H >> int ext m *op* n by intro *op*
   >> int ext m
   >> int ext n
7. H >> m *op* n in int by intro
   >> m in int
   >> n in int

where *op* must be one of +,-,*,/, or mod.

8. H,x:int,H' >> T ext ind($x;y,z.t_d;t_b;y,z.t_u$) by elim x new z[,y]
   y:int,y<0,z:T[y+1/x] >> T[y/x] ext t_d
   >> T[0/x] ext t_b
   y:int,0<y,z:T[y-1/x] >> T[y/x] ext t_u

The optional new name must be given if x occurs free in H'.

9. H >> ind($e;x,y.t_d;t_b;x,y.t_u$) in T[e/z]
   by intro [over $z.T$] [new u,v]
   >> e in int
   u:int,u<0,v:T[u+1/z] >> t_d[$u,v/x,y$] in T[u/z]
   >> t_b in T[0/z]
   u:int,0<u,v:T[u-1/z] >> t_u[$u,v/x,y$] in T[u/z]
10. H >> int_eq($a;b;t;t'$) in T by intro
    >> a in int
    >> b in int
    a=b in int >> t in T
    (a=b in int)->void >> t' in T

11. H >> less(a;b;t;t') in T by intro
    >> a in int
    >> b in int
    a<b >> t in T
    (a<b)->void >> t' in T

computation

12a. H >> ind(nt;x,$y.t_d$;t_b;x,$y.t_u$) = t in T by reduce 1 down
    >> t_d[nt,(ind(nt+1;x,$y.t_d$;t_b;x,$y.t_u$))/x,y] = t in T
    >> nt<0

12b. H >> ind(zt;x,y./t_d;t_b;x,$y.t_u$) = t in T by reduce 1 base
    >> t_b = t in T
    >> zt = 0 in int

12c. H >> ind(nt;x,$y.t_d$;t_b;x,$y.t_u$) = t in T by reduce 1 up
    >> t_u[nt,(ind(nt-1;x,$y.t_d$;t_b;x,$y.t_u$))/x,y] = t in T
    >> 0<nt

13a. H >> int_eq(a;a;t;t') = t'' in T by reduce 1
    >> t=t'' in T

13b. H >> int_eq(a;b;t;t') = t'' in T by reduce 1
    >> t' = t'' in T

where a and b are canonical int terms, and $a \neq b$.

14a. H >> less(a;b;t;t') = t'' in T by reduce 1
    >> t=t'' in T

where a and b are canonical int terms such that $a < b$.

14b. H >> less(a;b;t;t') = t'' in T by reduce 1
    >> t'=t'' in T

where a and b are canonical int terms such that $a \geq b$.

5 Propositional Functions

Kleene's axiomatization of arithmetic uses first-order logic with equality. This is another pure logical theory that we could study right after FOL. We require that equality on the domain of disclosure D is an equivalence relation, and we also require that all functions and predicates *respect equality*. For arithmetic, we base that understanding on axioms 14 and 17.

Ax. 14: $a' = b' \Rightarrow a = b$ (we also write $s(a) = s(b) \Rightarrow a = b$.)

Ax. 17: $a = b \Rightarrow a' = b'$ $(a = b \Rightarrow s(a) = s(b))$

Kleene also defines these relations.

$$\begin{array}{lllll} \text{Def} & a < b & \textit{iff} & \exists c.\ (s(c) + a = b) \\ & a \leq b & \textit{iff} & a < b \ \vee \ a = b \\ & a < b < c & \textit{iff} & a < b \ \ \& \ \ b < c \end{array}$$

We can also prove that numerical equality ($\mathbb{N}$ equality, $x = y$ in $\mathbb{N}$) is decidable:

$$\forall x, y.\ (x = y \in \mathbb{N} \vee \ \sim (x = y \in \mathbb{N})).$$

Kleene also states a constructive version of the Least Number Principle :

$$\exists x.\ A(x) \Rightarrow \exists y.\ (A(y)\ \&\ \forall z.\ (z < y \Rightarrow \sim A(z))).$$

It is easy to prove the following basic facts.

$$a \cdot b = 0 \Rightarrow a = 0 \ \vee \ b = 0$$

$$a \cdot b = 1 \Rightarrow a = 1 \ \ \& \ \ b = 1$$

$$c \neq 0 \Rightarrow (a \cdot c = b \cdot c \Rightarrow a = b)$$

6 Constructive Real Numbers

L.E.J Brouwer became the "founder" of the most compelling alternative to the Platonic philosophy of mathematics. His insights as a PhD student were profound and merit an entire course. He deepened this theory over a long and increasingly influential career. Recently the Cornell research group in constructive type theory adopted the most difficult to grasp concepts in his theory. It took us years to fully understand these ideas and implement them. The deepest ideas only arise in the study of computable analysis. These ideas connect geometry and real analysis, and we use them to give a semantics for constructive Euclidean geometry.

One reason it took so long to absorb Brouwer's deepest ideas is that they were not compiled into an introductory account that could be taught in universities. They were applied to analyzing some of the deepest and most intractable ideas in real analysis. That changed when the American mathematician Errett Bishop wrote a book on the subject with the simple title, *Foundations of Constructive Analysis*, 1967. Here is what Bishop said.

Prolog

Most mathematicians would find it hard to believe that there could be any serious controversy about the foundations of mathematics, any controversy whose outcome could significantly affect their own mathematical activity. Their attitude well represents the actual state of affairs: during a half-century of splendid mathematical progress there has been no deviation from the norm The voices of dissent, never much heeded, have long been silent

Perhaps the times are not conducive to introspection Mathematics flourishes as never before, its scope is immense, its quality high. Mathematicians flourish as never before, their profession is respectable, their salaries good. Mathematical methods are more fashionable than ever before: witness the surge of interest in mathematical logic, mathematical biology, mathematical economics, mathematical psychology – in mathematical investigations of every sort. The extent to which many of these investigations are premature or unrealistic indicates the deep attraction mathematical exactitude holds for the contemporary mind.

And yet there is dissatisfaction in the mathematical community. The pure mathematician is isolated from the world, which has little need of his brilliant creations. He suffers from an alienation which is seemingly inevitable. he has followed the gleam and it has led him out of this world.

If every mathematician occasionally, perhaps only for an instant, feels an urge to move closer to reality, it is not because he believes that mathematics is lacking in meaning. He does not believe that mathematics consists in drawing brilliant conclusions from arbitrary axioms, of juggling concepts devoid of pragmatic content, of playing a meaningless game. On the other hand, many mathematical statements have a rather peculiar pragmatic content Consider the theorem that either every even integer greater than 2 is the sum of two primes, or else there exists an even integer greater than 2 that is not the sum of two primes. The pragmatic content of this theorem is not that if we go to the integers and observe, we shall see certain things happening

Rather, the pragmatic content of such a theorem, if it exists, lies in the circumstance that we are going to use it to help derive other theorems, themselves of peculiar pragmatic content, which in turn will be the basis for further developments.

It appears then that there are certain mathematical statements that are merely evocative, that make assertions without empirical validity. There are also mathematical statements of immediate empirical validity, which say that certain performable operations will produce certain observable results: for instance, the theorem that every positive integer is the sum of four squares. Mathematics is a mixture of the real and the ideal, sometimes one, sometimes the other, often so presented that it is hard to tell which is which. The realistic component of mathematics – the desire for pragmatic interpretation – supplies the control which determines the course of development and keeps mathematics from lapsing into meaningless formalism. The idealistic component permits simplifications, and opens possibilities which would otherwise be closed. The methods of proof and the objects of investigation have been idealized to form a game, but the actual conduct of the game is ultimately motivated by pragmatic considerations.

For 50 years now there have been no significant changes in the rules of this game. Mathematicians unanimously agree on how mathematics should be played. Accepted standards of performance suffice to regulate the course of mathematical activity, and there is no prospect that these standards will be changed in any significant respect by a revision of the idealistic code. In fact, no efforts are being made to impose such a revision.

There have been, however, attempts to constructivize mathematics, to purge it completely of its idealistic content. The most sustained attempt was made by L.E.J. Brouwer, beginning in 1907. The movement he founded has long been dead, killed partly by compromises of Brouwer's disciples with the viewpoint of idealism, partly by extraneous peculiarities of Brouwer's system which made it vague and even ridiculous to practising mathematicians, but chiefly by the failure of Brouwer and his followers to convince the mathematical public that abandonment of the idealistic viewpoint would not sterilize or cripple the development of mathematics. Brouwer and other constructivists were much more successful in their criticisms of classical mathematics than in their efforts to replace it with something better. Many mathematicians familiar with Brouwer's objections to classical mathematics concede their validity but remain unconvinced that there is any satisfactory alternative.

This book is a piece of constructivist propaganda, designed to

show that there does exist a satisfactory alternative. To this end, we develop a large portion of abstract analysis within a constructive framework.

This development is carried through with an absolute minimum of philosophical prejudice concerning the nature of constructive mathematics. There are no dogmas to which we must conform. Our program is simple· to give numerical meaning to as much as possible of classical abstract analysis. Our motivation is the well-known scandal, exposed by Brouwer (and others) in great detail, that classical mathematics is deficient in numerical meaning.

Some familiarity with Brouwer's critique is essential. Following Brouwer, Chapter 1 is primarily devoted to an examination of the defects of classical mathematics, and a presentation of the thesis that all mathematics should have numerical meaning. Chapter 3 presents constructive versions of the fundamental concepts of sets and functions, and examines some of the obstacles to the constructivization of general topology

The remaining chapters are primarily technical, and constitute a course in abstract analysis from the constructive point of view. Very little formal preparation is required of the reader, although a certain level of mathematical sophistication is probably indispensible. Every effort has been made to follow the classical development as closely as possible; digressions have been relegated to notes at the ends of the various chapters.

The task of making analysis constructive is guided by three basic principles. First, to make every concept affirmative. (Even the concept of inequality is affirmative.) Second, to avoid definitions that are not relevant. (The concept of a pointwise continuous function is not relevant; a continuous function is one that is uniformly continuous on compact intervals.) Third, to avoid pseudogenerality. (Separability hypotheses are freely employed.)

The book has a threefold purpose: to present the constructive point of view, to show that the constructive program can succeed, and to lay a foundation for further work. These immediate ends tend to an ultimate goal – to hasten the inevitable day when constructive mathematics will be the accepted norm

We are not contending that idealistic mathematics is worthless from the constructive point of view. This would be as silly as contending that unrigorous mathematics is worthless from the classical point of view. Every theorem proved with idealistic methods presents a challenge: to find a constructive version, and to give it a constructive proof.

Errett Bishop

Chapter 1. A Constructivist Manifesto

1. The Descriptive Basis of Mathematics

Mathematics is that portion of our intellectual activity which transcends our biology and our environment The principles of biology as we know them may apply to life forms on other worlds, yet there is no necessity for this to be so. The principles of physics should be more universal, yet it is easy to imagine another universe governed by different physical laws. Mathematics, a creation of mind, is less arbitrary than biology or physics, creations of nature, the creatures we imagine inhabiting another world in another universe, with another biology and another physics, will develop a mathematics which in essence is the same as ours. In believing this we may be falling into a trap. Mathematics being a creation of our mind, it is, of course, difficult to imagine how mathematics could be other than it is without our actually making it so, but perhaps we should not presume to predict the course of the mathematical activities of all possible types of intelligence. On the other hand, the pragmatic content of our belief in the transcendence of mathematics has nothing to do with alien forms of life. Rather, it serves to give a direction to mathematical investigation, resulting from the insistence that mathematics be born of an inner necessity.

The primary concern of mathematics is number, and this means the positive integers. We feel about number the way Kant felt about space. The positive integers and their arithmetic are presupposed by the very nature of our intelligence and, we are tempted to believe, by the very nature of intelligence in general. The development of the theory of the positive integers from the primitive concept of the unit, the concept of adjoining a unit, and the process of mathematical induction carries complete conviction. In the words of Kronecker, the positive integers were created by God. Kronecker would have expressed it even better if he had said that the positive integers were created by God for the benefit of man (and other finite beings). Mathematics belongs to man, not to God. We are not interested in properties of the

positive integers that have no descriptive meaning for finite man. When a man proves a positive integer to exist, he should show how to find it. If God has mathematics of his own that needs to be done, let him do it himself.

Almost equal in importance to number are the constructions by which we ascend from number to the higher levels of mathematical existence. These constructions involve the discovery of relationships among mathematical entities already constructed, in the process of which new mathematical entities are created. The relations which form the point of departure are the order and arithmetical relations of the positive integers. From these we construct various rules for pairing integers with one another, for separating out certain integers from the rest, and for associating one integer with another. Rules of this sort give rise to the notions of set and function.

A set is not an entity which has an ideal existence. a set exists only when it has been defined. To define a set we prescribe, at least implicitly, what we (the constructing intelligence) must do in order to construct an element of the set, and what we must do to show that two elements of the set are equal. A similar remark applies to the definition of a function: in order to define a function from a set A to a set B, we prescribe a finite routine which leads from an element of A to an element of B, and show that equal elements of A give rise to equal elements of B.

Building on the positive integers, weaving a web of ever more sets and more functions, we get the basic structures of mathematics: the rational number system, the real number system, the euclidean spaces, the complex number system, the algebraic number fields, Hilbert space, the classical groups, and so forth. Within the framework of these structures most mathematics is done. Everything attaches itself to number, and every mathematical statement ultimately expresses the fact that if we perform certain computations within the set of positive integers, we shall get certain results.

Mathematics takes another leap, from the entity which is constructed in fact to the entity whose construction is hypothetical To some extent hypothetical entities are present from the start whenever we assert that every positive integer has a certain property, in essence we are considering a positive integer whose construction is hypothetical. But now we become bolder and consider a hypothetical set, endowed with hypothetical operations subject to certain axioms In this way we introduce such structures as topological spaces, groups, and manifolds. The motivation for doing this comes from the study of concretely constructed examples, and the justification comes from the possibility of applying the theory of the hypothetical structure to the

study of more than one specific example. Recently it has become fashionable to take another leap and study, as it were, a hypothetical hypothetical structure – a hypothetical structure qua hypothetical structure. Again the motivations and justifications attach themselves to particular examples, and the examples attach themselves to numbers in the ultimate analysis. Thus even the most abstract mathematical statement has a computational basis.

The transcendence of mathematics demands that it should not be confined to computations that I can perform, or you can perform, or 100 men working 100 years with 100 digital computers can perform. Any computation that can be performed by a finite intelligence – any computation that has a finite number of steps – is permissible. This does not mean that no value is to be placed on the efficiency of a computation An applied mathematician will prize a computation for its efficiency above all else, whereas in formal mathematics much attention is paid to elegance and little to efficiency. Mathematics should and must concern itself with efficiency, perhaps to the detriment of elegance, but these matters will come to the fore only when realism has begun to prevail. Until then our first concern will be to put as much mathematics as possible on a realistic basis without close attention to questions of efficiency.

2. The Idealistic Component of Mathematics

Geometry was highly idealistic from the time of Euclid and the ancients until the time of Descartes, unfolding from axioms taken either to be self-evident or to reflect properties of the real world. Descartes reduced geometry to the theory of the real numbers, and in the nineteenth century Dedekind, Weierstrass, and others, by the arithmetization of the real number system, brought space into the concrete realm of objects constructed by pure thought.

Unfortunately, the promise held out to mathematics by the arithmetization of space was not fulfilled, largely due to the intervention, around the turn of the century, of the formalist program. The successful formalization of mathematics helped keep mathematics on a wrong course. The fact that space has been arithmetized loses much of its significance if space, number, and everything else are fitted into a matrix of idealism where even the positive integers have an ambiguous computational existence. Mathematics becomes the game of sets, which is a fine game as far as it goes, with rules that are admirably precise. The game becomes its own justification, and the

fact that it represents a highly idealized version of mathematical existence is universally ignored.

Of course, idealistic tendencies have been present, if not dominant, in mathematics since the Greeks, but it took the full flowering of formalism to kill the insight into the nature of mathematics which its arithmetization could have given.

To see how some of the most basic results of classical analysis lack computational meaning, take the assertion that every bounded non-void set A of real numbers has a least upper bound. (The real number b is the *least upper bound* of A if $a \leqq b$ for all a in A, and if there exist elements of A that are arbitrarily close to b.) To avoid unnecessary complications, we actually consider the somewhat less general assertion that every bounded sequence (x_k) of rational numbers has a least upper bound b (in the set of real numbers). If this assertion were constructively valid, we could compute b, in the sense of computing a rational number approximating b to within any desired accuracy; in fact, we could program a digital computer to compute the approximations for us. For instance, the computer could be programmed to produce, one by one, a sequence $((b_k, m_k))$ of ordered pairs, where each b_k is a rational number and each m_k is a positive integer, such that $x_j \leqq b_k + k^{-1}$ for all positive integers j and k, and $x_{m_k} \geqq b_k - k^{-1}$ for all positive integers k. Unless there exists a general method M that produces such a computer program corresponding to each bounded, constructively given sequence (x_k) of rational numbers, we are not justified, by constructive standards, in asserting that each of the sequences (x_k) has a least upper bound. To see the scope such a method M would have, consider a constructively given sequence (n_k) of integers, each of which is either 0 or 1. Using the method M, we compute a rational number b_3 and a positive integer $N \equiv m_3$ such that (i) $n_j \leqq b_3 + \frac{1}{3}$ for all positive integers j, and (ii) $n_N \geqq b_3 - \frac{1}{3}$. Either $n_N = 0$ or $n_N = 1$. If $n_N = 0$, then (i) and (ii) imply that

$$n_j \leqq b_3 + \tfrac{1}{3} \leqq n_N + \tfrac{2}{3} = \tfrac{2}{3}$$

for all j. Since each n_j is either 0 or 1, it follows that $n_j = 0$ for all j Thus for each of the sequences (n_k) being considered, the method M either produces a proof that the n_k are all equal to 0, or produces a positive integer N such that $n_N = 1$. Of course, such a method M does not exist, and nobody expects that one will ever be found Such a method would solve most of the famous unsolved problems of mathematics – in particular, Fermat's last theorem, the Goldbach conjecture, and the Riemann hypothesis, since each of these problems can be reduced to finding, for a certain sequence (n_k) of the type being

considered, either a proof that $n_k=0$ for all k or a proof that $n_k=1$ for some k.

For another instance, consider the intuitively appealing theorem that every continuous real-valued function f on the closed interval $[0, 1]$, with $f(0)<0$ and $f(1)>0$, vanishes at some point x_0. This theorem can be derived from the least-upper-bound principle: take x_0 to be the least upper bound of the set of all x for which $f(x)<0$ The fact that we make use of the least-upper-bound principle does not mean our theorem is nonconstructive, it only means the given proof is nonconstructive. A closer examination demonstrates that our theorem itself is nonconstructive. This demonstration, which we now give, uses facts from the constructive theory of continuous functions, with which the reader is probably not familiar; nevertheless, it should provide some insight. Let (n_k) be any constructively given sequence of integers, each of which is either -1, 0, or 1. Define the real number a by $a \equiv \sum_{k=1}^{\infty} 3^{-k} n_k$. There exists a unique constructively given continuous function f on $[0, 1]$ such that $f(0)=-1$, $f(1)=1$, $f(\frac{1}{3})=f(\frac{2}{3})=a$, and f is linear on each of the intervals $[0, \frac{1}{3}]$, $[\frac{1}{3}, \frac{2}{3}]$, and $[\frac{2}{3}, 1]$. If our theorem is valid, there exists a point x_0 with $f(x_0)=0$ By computing a sufficiently close rational approximation to x_0, we show that either $x_0<\frac{2}{3}$ or $x_0>\frac{1}{3}$ In the first case, $a \geqq 0$, and therefore the first nonzero term of the sequence (n_k), if one exists, equals 1. Similarly, in the second case, the first nonzero term, if one exists, equals -1. Thus our theorem gives a method, which, applied to each of the sequences (n_k) being considered, either (i) proves that any term that equals 1 is preceded by a term that equals -1, or (ii) proves that any term that equals -1 is preceded by a term that equals 1. Nobody believes that such a method will ever be found.

Brouwer fought the advance of formalism and undertook the disengagement of mathematics from logic He wanted to strengthen mathematics by associating with every theorem and every proof a pragmatically meaningful interpretation. His program failed to gain support He was an indifferent expositor and an inflexible advocate, contending against the great prestige of Hilbert and the undeniable fact that idealistic mathematics produced the most general results with the least effort. More important, Brouwer's system itself had traces of idealism and, worse, of metaphysical speculation. There was a preoccupation with the philosophical aspects of constructivism at the expense of concrete mathematical activity A calculus of negation was developed which became a crutch to avoid the necessity of getting precise constructive results. It is not surprising that some of Brouwer's precepts were then formalized, giving rise to so-called intuitionistic

number theory, and that the formal system so obtained turned out not to be of any constructive value. In fairness to Brouwer it should be said that he did not associate himself with these efforts to formalize reality, it is the fault of the logicians that many mathematicians who think they know something of the constructive point of view have in mind a dinky formal system or, just as bad, confuse constructivism with recursive function theory

Brouwer became involved in metaphysical speculation by his desire to improve the theory of the continuum. A bugaboo of both Brouwer and the logicians has been compulsive speculation about the nature of the continuum In the case of the logicians this leads to contortions in which various formal systems, all detached from reality, are interpreted within one another in the hope that the nature of the continuum will somehow emerge. In Brouwer's case there seems to have been a nagging suspicion that unless he personally intervened to prevent it, the continuum would turn out to be discrete. He therefore introduced the method of free-choice sequences for constructing the continuum, as a consequence of which the continuum cannot be discrete because it is not well enough defined This makes mathematics so bizarre it becomes unpalatable to mathematicians, and foredooms the whole of Brouwer's program. This is a pity, because Brouwer had a remarkable insight into the defects of classical mathematics, and he made a heroic attempt to set things right.

3. The Constructivization of Mathematics

A set is defined by describing exactly what must be done in order to construct an element of the set, and what must be done in order to show that two elements are equal. There is no guarantee that the description will be understood; it may be that an author thinks he has described a set with sufficient clarity but a reader does not understand. For an illustration, consider the set of all sequences (n_k) of integers. To construct such a sequence we must give a rule which associates an integer n_k with each positive integer k in such a way that for each value of k the associated integer n_k can be determined in a finite number of steps by an entirely routine process Now this definition could perhaps be interpreted to admit sequences (n_k) in which n_k is constructed by a search, the proof that the search actually produces a value of n_k after a finite number of steps being given in some formal system Of course, we do not have this interpretation in mind, but it is impossible to consider every possible interpretation of our definition

and say whether that is what we have in mind. There is always ambiguity, but it becomes less and less as the reader continues to read and discovers more and more of the author's intent, modifying his interpretations if necessary to fit the intentions of the author as they continue to unfold. At any stage of the exposition the reader should be content if he can give a reasonable interpretation to account for everything the author has said. The expositor himself can never fully know all the possible ramifications of his definitions, and he is subject to the same necessity of modifying his interpretations, and sometimes his definitions as well, to conform to the dictates of experience

The constructive interpretations of the mathematical connectives and quantifiers have been established by Brouwer.

To prove the statement (P and Q) we must prove the statement P and prove the statement Q, just as in classical mathematics. To prove the statement (P or Q) we must either prove the statement P or prove the statement Q, whereas in classical mathematics it is possible to prove (P or Q) without proving either the statement P or the statement Q.

The connective "implies" is more complicated. To prove (P implies Q) we must show that P necessarily entails Q, or that Q is true whenever P is true. The validity of the computational facts implicit in the statement P must ensure the validity of the computational facts implicit in the statement Q, but the way this actually happens can only be seen by looking at the proof of the statement (P implies Q). Statements formed with this connective – for example, statements of the type ((P implies Q) implies R) – have a less immediate meaning than the statements from which they are formed, although in actual practice this does not seem to lead to difficulties in interpretation.

The negation (not P) of a statement P is the statement (P implies $(0=1)$) Classical mathematics makes no distinction between the content of the statements P and (not (not P)), whereas constructively the latter is a weaker statement.

Brouwer's system makes essential use of negation in defining, for instance, inequality and set complementation. Thus two elements of a set A are unequal according to Brouwer if the assumption of their equality somehow allows us to compute that $0=1$. It is natural to want to replace this negativistic definition by something more affirmative, phrased as much as possible in terms of specific computations leading to specific results. Brouwer himself does just this for the real number system, introducing an affirmative and stronger relation of inequality in addition to the negativistic relation already defined. Experience shows that it is not necessary to define inequality in terms of negation. For those cases in which an inequality relation in needed,

it is better to introduce it affirmatively; the same remarks apply to set complementation.

Van Dantzig and others have gone so far as to propose that negation could be entirely avoided in constructive mathematics Experience bears this out: in many cases where we seem to be using negation – for instance, in the assertion that either a given integer is even or it is not – we are really asserting that one of two finitely distinguishable alternatives obtains. Without intending to establish a dogma, we may continue to employ the language of negation but reserve it for situations of this sort (at least until experience changes our minds) and for counterexamples and purposes of motivation This will have the advantage of making mathematics more immediate, and in certain situations forcing us to sharpen our results

Proofs by contradiction are constructively justified in finite situations. When we have proved that one of finitely many alternatives holds at a certain stage in the proof of a theorem, to finish the proof of the theorem it is enough to show that the theorem is a consequence of each of the alternatives. Should one of the alternatives lead to a contradiction – that is, imply $(0=1)$ – either we may say that the alternative in question is ruled out and pass on to the consideration of the other alternatives, or we may be more meticulous and prove that the theorem is a consequence of the equality $0=1$.

A universal statement, to the effect that every element of a certain set A has a certain property P, has the same meaning in constructive as in classical mathematics. To prove such a statement we must show by some general argument that if x is any element of A, then x has property P.

Constructive existence is much more restrictive than the ideal existence of classical mathematics. The only way to show that an object exists is to give a finite routine for finding it, whereas in classical mathematics other methods can be used. In fact, the following principle is valid in classical mathematics: *Either all elements of A have property P or there exists an element of A with property (not P)* This principle, which we shall call the *principle of omniscience*, lies at the root of most nonconstructivity in classical mathematics. This is already true of the principle of omniscience in its simplest form. if (n_k) is a sequence of integers, then either $n_k=0$ for some k or $n_k\neq 0$ for all k. We shall call this the *limited principle of omniscience*. Theorem after theorem of classical mathematics depends in an essential way on the limited principle of omniscience, and is therefore not constructively valid. Some instances of this are: the theorem that a continuous real-valued function on a closed, bounded interval attains its maximum; the fixed-point theorem for a continuous map of a closed cell into

itself, the ergodic theorem, and the Hahn-Banach theorem Nevertheless these theorems are not lost to constructive mathematics: each of these theorems P has a constructive substitute Q which is a constructively valid theorem Q implying P in the classical system by a more or less simple argument based on the limited principle of omniscience. For example, the statement "every continuous function from a closed cell in euclidean space into itself admits a fixed point" finds a constructive substitute in the theorem that such a function admits a point which is arbitrarily near to its image.

The extent to which good constructive substitutes exist for the theorems of classical mathematics can be regarded as a demonstration that classical mathematics has a substantial underpinning of constructive truth

When a classical mathematician claims he is constructivist, he probably means he avoids the axiom of choice. This axiom is unique in its ability to trouble the conscience of the classical mathematician, but in fact it is not a real source of nonconstructivity in classical mathematics. A choice function exists in constructive mathematics, because a choice is *implied by the very meaning of existence.* Applications of the axiom of choice in classical mathematics either are irrelevant or are combined with a sweeping use of the principle of omniscience. The axiom of choice is used to extract elements from equivalence classes where they should never have been put in the first place. For example, a real number should not be defined as an equivalence class of Cauchy sequences of rational numbers; there is no need to drag in the equivalence classes. The proof that the real numbers can be well ordered is an instance of a proof in which a sweeping use of the principle of omniscience is combined with an appeal to the axiom of choice. Such proofs offer little hope of constructivization: it is not likely that the theorem "the real numbers can be well ordered" will be given a constructive version consonant with the intuitive interpretation of the classical result.

Almost every conceivable type of resistance has been offered to a straightforward realistic treatment of mathematics, even by constructivists. Brouwer, who has done more for constructive mathematics than anyone else, though it necessary to introduce a revolutionary, semimystical theory of the continuum Weyl, a great mathematician who in practice suppressed his constructivist convictions, expressed the opinion that idealistic mathematics finds its justification in its applications to physics. Hilbert, who insisted on constructivity in metamathematics but believed that the price of a constructive mathematics was too great, was willing to settle for consistency. Brouwer's disciples joined forces with the logicians in attempts to formalize

constructive mathematics Others seek constructive truth in the framework of recursive function theory Still others look for a short cut to reality, a point of vantage which will suddenly reveal classical mathematics in a constructive light. None of these substitutes for a straightforward realistic approach has worked It is no exaggeration to say that a straightforward realistic approach to mathematics has yet to be tried It is time to make the attempt.

Notes

Errett Bishop was never happy with the standard constructive interpretation of implication (the one given in Section 3). Among the alternatives he felt worthy of serious investigation is "Gödel implication", as discussed in [9]. Bishop also worked on a deeper study of implication, but unfortunately he left only fragmentary notes on his ideas

At first sight, Bishop' remark, "A choice function exists in constructive mathematics, because a choice is *implied by the very meaning of existence*", appears to be contradicted by counterexamples of the sort discussed in connection with the least-upper-bound principle. In fact, there is no contradiction here. To see this, consider a paraphrase of Bishop's remark: if to each x in a set A there corresponds an element y of a set B such that a given property $P(x, y)$ holds, then it is implied by the very meaning of existence in constructive mathematics that there is a finite routine for computing an appropriate y in B from a given x in A; although this routine may not be a function relative to the given equality relation on A, it *is* a function relative to the equality relation of identity (intensional equality) on A, in which two elements are equal if and only if they are given as identically the same object

Chapter 2. Calculus and the Real Numbers

Section 1 establishes some conventions about sets and functions. The next three sections are devoted to constructing the real numbers as certain Cauchy sequences of rational numbers, and investigating their order and arithmetic. The rest of the chapter deals with the basic ideas of the calculus of one variable. Topics covered include continuity, the convergence of sequences and series of continuous functions, differentiation, integration, Taylor's theorem, and the basic properties of the exponential and trigonometric functions and their inverses Most of the material is a routine constructivization of the corresponding part of classical mathematics; for this reason it affords a good introduction to the constructive approach.

We assume that the reader is familiar with the order and arithmetic of the integers and the rational numbers. For us, a *rational number* will be an expression of the form p/q, where p and q are integers with $q \neq 0$. Two rational numbers p/q and p'/q' are *equal* if $pq' = p'q$. The integer n is identified with the rational number $n/1$.

There are geometric magnitudes which are not represented by rational numbers, and which can only be described by a sequence of rational approximations. Certain such approximating sequences are called real numbers. In this chapter we construct the real numbers and study their basic properties. Then we develop the fundamental ideas of the calculus.

1. Sets and Functions

Before constructing the real numbers, we introduce some notions which are basic to much of mathematics.

The totality of all mathematical objects constructed in accordance with certain requirements is called a *set*. The requirements of the

construction, which vary with the set under consideration, determine the set. Thus the integers form a set, the rational numbers form a set, and (we anticipate here the formal definition of 'sequence') the collection of all sequences of integers is a set.

Each set will be endowed with a binary relation $=$ of *equality*. This relation is a matter of convention, except that it must be an *equivalence relation*, in other words, the following conditions must hold for all objects x, y, and z in the set:

(1 1) (i) $x=x$
(ii) If $x=y$, then $y=x$
(iii) If $x=y$ and $y=z$, then $x=z$.

The relation of equality given above for rational numbers is an equivalence relation. In this example there is a finite, mechanical procedure for deciding whether or not two given objects in the set are equal Such a procedure will not exist in general: there are instances in which we are unable to decide whether or not two given elements of a set are equal; such an instance, in the theory of real numbers, will be given later.

We use the standard notation $a\in A$ to denote that a is an *element*, or *member*, of the set A, or that the construction defining a satisfies the requirements a construction must satisfy in order to define an object of A. We also use the notation $\{a_1, a_2, \ldots\}$ for a set whose elements can be written in a (possibly finite) list.

The dependence of one quantity on another is expressed by the basic notion of an operation An *operation* from a set A into a set B is a finite routine f which assigns an element $f(a)$ of B to each given element a of A. This routine must afford an explicit, finite, mechanical reduction of the procedure for constructing $f(a)$ to the procedure for constructing a. If it is clear from the context what the sets A and B are, we sometimes denote f by $a\mapsto f(a)$, in order to bring out the form of $f(a)$ for a given element a of A. The set A is called the *domain* of the operation, and is denoted by $\mathrm{dmn} f$. In the most important case, we have $f(a)=f(a')$ whenever $a, a'\in A$ and $a=a'$; the operation f is then called a *function*, or a *mapping* of A into B, or a *map* of A into B For two functions f, g from A into B, $f=g$ means that $f(a)=g(a)$ for each element a of A. Taken with this equality relation, the collection of all functions from A into B becomes a set.

The notation $f\colon A\to B$ indicates that f is a function from the set A to the set B.

A function x whose domain is the set $\mathbb{Z}^+$ of positive integers is called a *sequence*. The object $x_n\equiv x(n)$ is called the n^{th} *term* of the

sequence. The finite routine x can be given explicitly, or it can be left to inference: for example, by writing the terms of the sequence in order

$$(x_1, x_2, \ldots)$$

until the rule of their formation becomes clear. Different notations for the sequence whose n^{th} term is x_n are: $n \mapsto x_n$, $(x_1, x_2, \ldots)$, $(x_n)_{n=1}^{\infty}$, and (x_n). Thus the sequence whose n^{th} term is n^2 can be written $n \mapsto n^2$, or $(1, 4, 9, \ldots)$, or $(n^2)_{n=1}^{\infty}$, or simply (n^2).

A *subsequence* of a sequence (x_n) consists of the sequence (x_n) and a sequence $(n_k)_{k=1}^{\infty}$ of positive integers such that $n_1 < n_2 < \ldots$. We identify such a subsequence with the sequence whose k^{th} term is x_{n_k}.

Sometimes we shall speak of sequences whose domain is some set of integers other than $\mathbb{Z}^+$. For example, we shall write $(x_n)_{n=0}^{\infty}$ to denote a mapping x from the set of nonnegative integers, where $x_n \equiv x(n)$ for each n.

Another example arises as follows. If n is a positive integer, then a *finite sequence* of *length* n is a function from the set $\{1, 2, \ldots, n\}$ into a set B.

The *cartesian product*, or simply the *product*, of sets $X_1, \ldots, X_n$ is defined to be the set

$$X \equiv X_1 \times X_2 \times \ldots \times X_n$$

of all ordered n-tuples $(x_1, \ldots, x_n)$ with $x_1 \in X_1$, $x_2 \in X_2, \ldots$, and $x_n \in X_n$. Elements $(x_1, \ldots, x_n)$ and $(x_1', \ldots, x_n')$ of the cartesian product are *equal* if the *coordinates* (or *components*) x_k and x_k' are equal elements of X_k for each k.

If x is a finite sequence of elements of a set B, then x can be identified with the element $(x(1), \ldots, x(n))$ of the cartesian product

$$B^n \equiv B \times B \times \ldots \times B,$$

where n is the length of x

Returning to functions in general, we say that a function $f: A \to B$ maps A *onto* B if to each element b of B there corresponds an element a of A with $f(a) = b$. In other words, f maps A onto B if there is an operation g from B into A such that $f(g(b)) = b$ for each b in B. A set A is *countable* if there exists a mapping of $\mathbb{Z}^+$ onto A, intuitively, this means that the elements of A can be arranged in a sequence with possible duplications.

The elements of the cartesian product $\mathbb{Z} \times \mathbb{Z}$ of the set $\mathbb{Z}$ of integers with itself can be arranged in a sequence as follows. We order the elements (m, n) of $\mathbb{Z} \times \mathbb{Z}$, first according to the value of $|m| + |n|$, then according to the value of m, and finally according to the value of

n. This produces the sequence

$$(1\,2)\quad ((0,0),(-1,0),(0,-1),(0,1),(1,0),(-2,0),(-1,-1),\ldots),$$

in which each element of $\mathbb{Z}\times\mathbb{Z}$ occurs exactly once. In the sequence (1.2), omit every term (m,n) with $n=0$, and replace each term (m,n) with $n\neq 0$ by m/n; this produces the sequence

$$(1.3)\qquad (0/-1, 0/1, -1/-1, \ldots),$$

in which every expression p/q, with p and q integers and $q\neq 0$, occurs exactly once Keeping only the term 0/1 of (1.3), and those terms for which $q>0$, $p\neq 0$, and p is relatively prime to q, we obtain a sequence

$$(1\,4)\qquad (0/1, -1/1, 1/1, \quad)$$

which has the property that for any given rational number r there exists exactly one term equal to r.

For each positive integer n, let $\mathbb{Z}_n$ be the set $\{0,1,\ldots,n-1\}$ If there is a mapping of $\mathbb{Z}_n$ onto the set A, then we say that A has *at most n elements*. A set with at most n elements for some n is said to be *subfinite*, or *finitely enumerable*. Note that every subfinite or countable set has at least one element.

Before we introduce stronger notions than countability and subfiniteness, we must discuss the composition of functions The *composition* of two functions $f\colon A\to B$ and $g\colon B\to C$ is the function $g\circ f\colon A\to C$ defined by

$$(g\circ f)(a)\equiv g(f(a))\qquad (a\in A).$$

Composition is associative:

$$h\circ(g\circ f)=(h\circ g)\circ f$$

whenever the compositions are defined

If $f.\ A\to B$, $g\colon B\to A$, and $g(f(a))=a$ for all a in A, then the function g is called a *left inverse* of f, and the function f is called a *right inverse* of g (Note that f has a left inverse if and only if it is *one-one*, in the sense that $a=a'$ for all elements a,a' of A with $f(a)=f(a')$) When g is both a left and a right inverse of f, then it is simply called an *inverse* of f; f is then called a *one-one correspondence*, or a *bijection*, and the sets A and B are said to be *in one-one correspondence* with each other

A set which is in one-one correspondence with the set $\mathbb{Z}^+$ of positive integers is said to be *countably infinite*. For example, let f be the sequence (1.4), and define a function g from the set $\mathbb{Q}$ of rational numbers to $\mathbb{Z}^+$ by writing $g(r)\equiv n$, where n is the unique positive integer for which $f(n)=r$. Then g is an inverse of f; so that the set $\mathbb{Q}$

is countably infinite. A similar proof using (1.2) shows that $\mathbb{Z}\times\mathbb{Z}$ is countably infinite

A set which is in one-one correspondence with $\mathbb{Z}_n$ is said to *have n elements*, and to be *finite*. Every finite set is countable.

It is not true that every countable set is either countably infinite or subfinite. For example, let A consist of all positive integers n such that both n and $n+2$ are prime; then A is countable, but we do not know if it is either countably infinite or subfinite. This does not rule out the possibility that at some time in the future A will have become countably infinite or subfinite; it is possible that tomorrow someone will show that A is subfinite. This set A has the property that if it is subfinite, then it is finite. Not all sets have this property.

2. The Real Number System

The following definition is basic to everything that follows.

(2.1) **Definition.** A sequence (x_n) of rational numbers is *regular* if

$$|x_m - x_n| \leqq m^{-1} + n^{-1} \qquad (m, n \in \mathbb{Z}^+). \tag{2.1.1}$$

A *real number* is a regular sequence of rational numbers. Two real numbers $x \equiv (x_n)$ and $y \equiv (y_n)$ are *equal* if

$$|x_n - y_n| \leqq 2n^{-1} \qquad (n \in \mathbb{Z}^+). \tag{2.1.2}$$

The set of real numbers is denoted by $\mathbb{R}$.

(2.2) **Proposition.** *Equality of real numbers is an equivalence relation.*

Proof: Parts (i) and (ii) of (1.1) are obvious. Part (iii) is a consequence of the following lemma.

(2.3) **Lemma.** *The real numbers $x \equiv (x_n)$ and $y \equiv (y_n)$ are equal if and only if for each positive integer j there exists a positive integer N_j such that*

$$|x_n - y_n| \leqq j^{-1} \qquad (n \geqq N_j). \tag{2.3.1}$$

Proof: If $x = y$, then (2.3.1) holds with $N_j \equiv 2j$.

Assume conversely that for each j in $\mathbb{Z}^+$ there exists N_j satisfying (2.3.1). Consider a positive integer n. If m and j are any positive integers with $m \geqq \max\{j, N_j\}$, then

$$\begin{aligned}|x_n - y_n| &\leqq |x_n - x_m| + |x_m - y_m| + |y_m - y_n| \\ &\leqq (n^{-1} + m^{-1}) + j^{-1} + (n^{-1} + m^{-1}) < 2n^{-1} + 3j^{-1}.\end{aligned}$$

Since this holds for all j in $\mathbb{Z}^+$, (2.1.2) is valid. □

Notice that the proof of Lemma (2.3) singles out a specific N_j satisfying (2.3.1). This situation is typical: every proof of a theorem which asserts the existence of an object must embody, at least implicitly, a finite routine for the construction of the object.

The rational number x_n is called the n^{th} *rational approximation* to the real number $x \equiv (x_n)$. Note that the operation from $\mathbb{R}$ to $\mathbb{Q}$ which takes the real number x into its n^{th} rational approximation is not a function.

For later use we wish to associate with each real number $x \equiv (x_n)$ an integer K_x such that

$$|x_n| < K_x \qquad (n \in \mathbb{Z}^+).$$

This is done by letting K_x be the least integer which is greater than $|x_1| + 2$. We call K_x the *canonical bound* for x.

The development of the arithmetic of the real numbers offers no surprises: we operate with real numbers by operating with their rational approximations.

(2.4) **Definition.** Let $x \equiv (x_n)$ and $y \equiv (y_n)$ be real numbers with respective canonical bounds K_x and K_y. Write

$$k \equiv \max\{K_x, K_y\}.$$

Let α be any rational number. We define

(a) $x + y \equiv (x_{2n} + y_{2n})_{n=1}^{\infty}$

(b) $xy \equiv (x_{2kn} y_{2kn})_{n=1}^{\infty}$

(c) $\max\{x, y\} \equiv (\max\{x_n, y_n\})_{n=1}^{\infty}$

(d) $-x \equiv (-x_n)_{n=1}^{\infty}$

(e) $\alpha^* \equiv (\alpha, \alpha, \alpha, \ldots)$.

(2.5) **Proposition.** *The sequences* $x + y$, xy, $\max\{x, y\}$, $-x$, *and* α^* *of Definition* (2.4) *are real numbers.*

Proof (a) Write $z_n \equiv x_{2n} + y_{2n}$ Then $x+y \equiv (z_n)$. For all positive integers m and n,

$$\begin{aligned} |z_m - z_n| &\leqq |x_{2m} - x_{2n}| + |y_{2m} - y_{2n}| \\ &\leqq (2n)^{-1} + (2m)^{-1} + (2n)^{-1} + (2m)^{-1} = n^{-1} + m^{-1}. \end{aligned}$$

Thus $x+y$ is a real number.

(b) Write $z_n \equiv x_{2kn} y_{2kn}$. Then $xy \equiv (z_n)$. For all positive integers m and n,

$$\begin{aligned} |z_m - z_n| &= |x_{2km}(y_{2km} - y_{2kn}) + y_{2kn}(x_{2km} - x_{2kn})| \\ &\leqq k|y_{2km} - y_{2kn}| + k|x_{2km} - x_{2kn}| \\ &\leqq k((2km)^{-1} + (2kn)^{-1} + (2km)^{-1} + (2kn)^{-1}) = n^{-1} + m^{-1}. \end{aligned}$$

Thus xy is a real number.

(c) Write $z_n \equiv \max\{x_n, y_n\}$ Then $\max\{x, y\} \equiv (z_n)$. Consider positive integers m and n. For simplicity assume that

$$x_m = \max\{x_m, x_n, y_m, y_n\}.$$

Then

$$\begin{aligned} |z_m - z_n| &= |x_m - \max\{x_n, y_n\}| \\ &= x_m - \max\{x_n, y_n\} \leqq x_m - x_n \leqq n^{-1} + m^{-1}. \end{aligned}$$

Thus $\max\{x, y\}$ is a real number.

(d) For all positive integers m and n,

$$|-x_m - (-x_n)| = |x_m - x_n| \leqq m^{-1} + n^{-1}.$$

Thus $-x$ is a real number

(e) This is obvious $\square$

There is no trouble in proving that $(x, y) \mapsto x+y$, $(x, y) \mapsto xy$, and $(x, y) \mapsto \max\{x, y\}$ are functions from $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$, that $x \mapsto -x$ is a function from $\mathbb{R}$ to $\mathbb{R}$; and that $\alpha \mapsto \alpha^*$ is a function from $\mathbb{Q}$ to $\mathbb{R}$.

The operation

$$x \mapsto |x| \equiv \max\{x, -x\}$$

is therefore a function from $\mathbb{R}$ to $\mathbb{R}$, and the operation

$$(x, y) \mapsto \min\{x, y\} \equiv -\max\{-x, -y\}$$

is a function from $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$

The next proposition states that the real numbers obey the same rules of arithmetic as the rational numbers.

(2.6) **Proposition.** *For arbitrary real numbers x, y, and z and rational numbers α and β,*

(a) $x+y = y+x$. $xy = yx$

(b) $(x+y)+z=x+(y+z)$, $x(yz)=(xy)z$

(c) $x(y+z)=xy+xz$

(d) $0^*+x=x$, $1^*x=x$

(e) $x-x=0^*$

(f) $|xy|=|x||y|$

(g) $(\alpha+\beta)^*=\alpha^*+\beta^*$, $(\alpha\beta)^*=\alpha^*\beta^*$, *and* $(-\alpha)^*=-\alpha^*$.

We omit the simple proofs of these results.

We shall use standard notations, such as $x+y+z$ and $\max\{x, y, z\}$, without further comment.

There are three basic relations defined on the set of real numbers. The first of these, the equality relation, has already been defined. The remaining relations, which pertain to order, are best introduced in terms of certain subsets $\mathbb{R}^+$ and $\mathbb{R}^{0+}$ of $\mathbb{R}$.

(2.7) **Definition.** A real number $x\equiv(x_n)$ is *positive*, or $x\in\mathbb{R}^+$, if

$$x_n>n^{-1} \tag{2.7.1}$$

for some n in $\mathbb{Z}^+$. A real number $x\equiv(x_n)$ is *nonnegative*, or $x\in\mathbb{R}^{0+}$, if

$$x_n\geqq -n^{-1} \quad (n\in\mathbb{Z}^+). \tag{2.7.2}$$

The following criteria are often useful

(2.8) **Lemma.** *A real number $x\equiv(x_n)$ is positive if and only if there exists a positive integer N such that*

$$x_m\geqq N^{-1} \quad (m\geqq N). \tag{2.8.1}$$

A real number $x\equiv(x_n)$ is nonnegative if and only if for each n in $\mathbb{Z}^+$ there exists N_n in $\mathbb{Z}^+$ such that

$$x_m\geqq -n^{-1} \quad (m\geqq N_n). \tag{2.8.2}$$

Proof: Assume that $x\in\mathbb{R}^+$. Then $x_n>n^{-1}$ for some n in $\mathbb{Z}^+$. Choose N in $\mathbb{Z}^+$ with

$$2N^{-1}\leqq x_n-n^{-1}.$$

Then

$$\begin{aligned} x_m&\geqq x_n-|x_m-x_n|\geqq x_n-m^{-1}-n^{-1}\\ &\geqq x_n-n^{-1}-N^{-1}>N^{-1} \end{aligned}$$

whenever $m\geqq N$. Therefore (2.8.1) is valid.

Conversely, if (2.8.1) is valid, then (2.7 1) holds with $n=N+1$ Therefore $x\in\mathbb{R}^+$.

Assume next that $x \in \mathbb{R}^{0+}$. Then for each positive integer n,

$$x_m \geqq -m^{-1} \geqq -n^{-1} \qquad (m \geqq n).$$

Therefore (2.8.2) is valid with $N_n \equiv n$.

Assume finally that (2.8.2) holds. Then if k, m, and n are positive integers with $m \geqq N_n$, we have

$$x_k \geqq x_m - |x_m - x_k| \geqq -n^{-1} - k^{-1} - m^{-1}.$$

Since m and n are arbitrary, this gives $x_k \geqq -k^{-1}$. Therefore $x \in \mathbb{R}^{0+}$. □

As a corollary of Lemma (2.8), we see that if x and y are equal real numbers, then x is positive if and only if y is positive, and x is nonnegative if and only if y is nonnegative.

It is not strictly correct to say that a real number (x_n) *is* an element of $\mathbb{R}^+$. An element of $\mathbb{R}^+$ consists of a real number (x_n) *and* a positive integer n such that $x_n > n^{-1}$, because an element of $\mathbb{R}^+$ is not presented until both (x_n) and n are given. One and the same real number (x_n) can be associated with two distinct (but equal) elements of $\mathbb{R}^+$ Nevertheless we shall continue to refer loosely to a positive real number (x_n). On those occasions when we need to refer to an n for which $x_n > n^{-1}$, we shall take the position that it was there implicitly all along.

The proof of the following proposition is now easy, and will be left to the reader. For convenience, $\mathbb{R}^*$ represents either $\mathbb{R}^+$ or $\mathbb{R}^{0+}$.

(2.9) **Proposition.** *Let x and y be real numbers. Then*

(a) $x+y \in \mathbb{R}^*$ *and* $xy \in \mathbb{R}^*$ *whenever* $x \in \mathbb{R}^*$ *and* $y \in \mathbb{R}^*$

(b) $x+y \in \mathbb{R}^+$ *whenever* $x \in \mathbb{R}^+$ *and* $y \in \mathbb{R}^{0+}$

(c) $|x| \in \mathbb{R}^{0+}$

(d) $\max\{x, y\} \in \mathbb{R}^*$ *whenever* $x \in \mathbb{R}^*$

(e) $\min\{x, y\} \in \mathbb{R}^*$ *whenever* $x \in \mathbb{R}^*$ *and* $y \in \mathbb{R}^*$.

We now define the order relations on $\mathbb{R}$.

(2.10) **Definition.** Let x and y be real numbers. We define

$$x > y \text{ (or } y < x) \quad \text{if } x - y \in \mathbb{R}^+$$

and

$$x \geqq y \text{ (or } y \leqq x) \quad \text{if } x - y \in \mathbb{R}^{0+}.$$

A real number x is *negative* if $x < 0^*$ – that is, if $-x$ is positive.

Consider real numbers x, x', y, and y' such that (i) $x=x'$, $y=y'$, and $x>y$ We have

$$x'-y'=x-y\in\mathbb{R}^+$$

and therefore (ii) $x'>y'$. We express the fact that (ii) holds whenever (i) is valid by saying that $>$ is a *relation* on $\mathbb{R}$. More formally, a *relation* on a set X is a subset S of $X\times X$ such that if x, x', y, y' are elements of X with $x=x'$, $y=y'$, and $(x,y)\in S$, then $(x',y')\in S$.

We express the fact that $x>y$ if and only if $y<x$ by saying that $>$ and $<$ are *transposed relations.* Similarly, $\geqq$ and $\leqq$ are transposed relations.

If $x<y$ or $x=y$, then $x\leqq y$. The converse is not valid: as we shall see later, it is possible that we have $x\leqq y$ without being able to prove that $x<y$ or $x=y$. For this reason it was necessary to define the relations $<$ and $\leqq$ independently of each other.

The following rules for manipulating inequalities are easily proved from Proposition (2.9). We omit the proofs.

(2.11) **Proposition.** *For all real numbers x, y, z, and t,*

(a) $x<z$ *whenever either* $x<y$ *and* $y\leqq z$ *or* $x\leqq y$ *and* $y<z$

(b) $x\leqq z$ *whenever* $x\leqq y$ *and* $y\leqq z$

(c) $x+y\leqq z+t$ *whenever* $x\leqq z$ *and* $y\leqq t$

(d) $x+y<z+t$ *whenever* $x\leqq z$ *and* $y<t$

(e) $xy\leqq zy$ *whenever* $x\leqq z$ *and* $y\geqq 0$*

(f) $xy<zy$ *whenever* $x<z$ *and* $y>0$*

(g) *if* $x<y$, *then* $-x>-y$

(h) *if* $x\leqq y$, *then* $-x\geqq -y$

(i) $\max\{x,y\}\geqq x$

(j) $\min\{x,y\}\leqq x$

(k) *if* $x\leqq y$ *and* $y\leqq x$, *then* $x=y$

(l) $|x|\geqq 0$*

(m) $|x+y|\leqq|x|+|y|$.

An important property of the relation $<$, of which we shall make no use, is the *antisymmetry* property, which states that at most one of the relations $x<y$ and $y<x$ is valid for given real numbers x and y. This negative statement has no place in the affirmative mathematics we are trying to develop, except as motivation. Its place is taken by the affirmative statement (k) of Proposition (2.11). As a general principle, negative statements are only for counterexamples and motivation; they are not to be used in subsequent work.

(2.12) **Definition.** For real numbers x and y we write $x \neq y$ if and only if $x < y$ or $x > y$.

Inequality $\neq$ is a relation because both $<$ and $>$ are relations. As motivation we have the negative statement that at most one of the relations $x \neq y$, $x = y$ can hold for given real numbers x and y. In other words, at most one of the relations $x < y$, $x > y$, $x = y$ can hold. This is clear from the definitions.

The following proposition defines the *inverse* x^{-1} of a real number $x \neq 0^*$, and derives the basic properties of the operation $x \mapsto x^{-1}$.

(2.13) **Proposition.** *Let x be a nonzero real number (so that $|x| \in \mathbb{R}^+$). There exists a positive integer N with $|x_m| \geqq N^{-1}$ for $m \geqq N$. Define*

$$y_n \equiv (x_{N^3})^{-1} \qquad (n < N)$$

and

$$y_n \equiv (x_{nN^2})^{-1} \qquad (n \geqq N).$$

Then

$$x^{-1} \equiv (y_n)_{n=1}^{\infty}$$

is a real number which is positive if x is positive, and negative if x is negative; also $xx^{-1} = 1^*$.

If t is any real number for which $xt = 1^*$, then $t = x^{-1}$. The operation $x \mapsto x^{-1}$ is a function. If $x \neq 0$ and $y \neq 0$, then $(xy)^{-1} = x^{-1}y^{-1}$. If $\alpha \neq 0$ is rational, then $(\alpha^*)^{-1} = (\alpha^{-1})^*$. If $x \neq 0$, then $(x^{-1})^{-1} = x$.

Proof: Our definitions guarantee that $|y_n| \leqq N$ for all n.

Consider positive integers m and n. Write

$$j \equiv \max\{m, N\}, \qquad k \equiv \max\{n, N\}.$$

Then

$$\begin{aligned} |y_m - y_n| &= |y_m||y_n||x_{jN^2} - x_{kN^2}| \\ &\leqq N^2((jN^2)^{-1} + (kN^2)^{-1}) = j^{-1} + k^{-1} \leqq m^{-1} + n^{-1}. \end{aligned}$$

Therefore x^{-1} is a real number.

Assume now that $x > 0^*$. Then by (2.8), $x_n > 0$ for all sufficiently large n. Hence $y_n > K_x^{-1}$ (where K_x is the canonical bound for x) for all sufficiently large n. It follows from (2.8) that $x^{-1} > 0^*$. A similar proof shows that $x^{-1} < 0^*$ whenever $x < 0^*$.

Let k be the maximum of the canonical bounds for x and x^{-1}. Write $xx^{-1} \equiv (z_n)$. Then

$$z_n \equiv x_{2nk}y_{2nk} \equiv x_{2nk}(x_{2nN^2k})^{-1} \qquad (n \geqq N)$$

Therefore

$$\begin{aligned} |z_n - 1^*| &= |x_{2nN^2k}|^{-1}|x_{2nk} - x_{2nN^2k}| \\ &\leqq |y_{2nk}|((2nk)^{-1} + (2nN^2k)^{-1}) \leqq n^{-1} \end{aligned}$$

for $n \geqq N$. It follows that $xx^{-1} = 1^*$.

If t is any real number with $xt=1^*$, then

$$x^{-1}=x^{-1}(xt)=(x^{-1}x)t=(xx^{-1})t=t.$$

If $x=x'$, then

$$x'x^{-1}=xx^{-1}=1^*.$$

Therefore $x^{-1}=(x')^{-1}$. It follows that $x\mapsto x^{-1}$ is a function.

If $x\neq 0$ and $y\neq 0$, then

$$(xy)x^{-1}y^{-1}=xx^{-1}yy^{-1}=1^*.$$

Therefore $x^{-1}y^{-1}=(xy)^{-1}$.

If $\alpha\neq 0$ is rational, then $\alpha^*\equiv(\alpha,\alpha,\ldots)$. Therefore

$$(\alpha^*)^{-1}=(\alpha^{-1},\alpha^{-1},\ldots)=(\alpha^{-1})^*.$$

For each x in $\mathbb{R}^+$, x^{-1} is in $\mathbb{R}^+$, and thus $(x^{-1})^{-1}$ exists. Since $x^{-1}x=xx^{-1}=1^*$, it follows that $(x^{-1})^{-1}=x$. Similarly $(x^{-1})^{-1}=x$ if x is negative Therefore $(x^{-1})^{-1}=x$ whenever $x\neq 0$. □

Of course, we often write x/y instead of xy^{-1} when x and y are real numbers with $y\neq 0$.

As the previous propositions show, $(\alpha\beta)^*=\alpha^*\beta^*$, $(\alpha+\beta)^*=\alpha^*+\beta^*$, $(-\alpha)^*=-\alpha^*$, $(|\alpha|)^*=|\alpha^*|$, and $(\alpha^{-1})^*=(\alpha^*)^{-1}$ for all rational numbers α and β. Also $\alpha\triangle\beta$ if and only if $\alpha^*\triangle\beta^*$, where $\triangle$ stands for any of the relations $=$, $<$, $>$, and $\neq$. This situation is expressed by saying that the map $\alpha\mapsto\alpha^*$ is an *order isomorphism* from $\mathbb{Q}$ into $\mathbb{R}$. This justifies identifying $\mathbb{Q}$ with a subset of $\mathbb{R}$, as we previously identified $\mathbb{Z}$ with a subset of $\mathbb{Q}$. Henceforth we make no distinction between a rational number α and the corresponding real number α^*

The next lemma shows that the n^{th} rational approximation x_n to a real number $x\equiv(x_n)$ actually approximates x to within n^{-1}.

(2.14) **Lemma.** *For each real number $x\equiv(x_n)$, we have*

$$|x-x_n|\leqq n^{-1}\qquad(n\in\mathbb{Z}^+).$$

Proof: By (2.4) and the definition of $|\ |$, the m^{th} rational approximation to $n^{-1}-|x-x_n|$ is

$$n^{-1}-|x_{4m}-x_n|\geqq n^{-1}-((4m)^{-1}+n^{-1})=-(4m)^{-1}>-m^{-1}.$$

By (2.7), we have $n^{-1}-|x-x_n|\in\mathbb{R}^{0+}$. Therefore $|x-x_n|\leqq n^{-1}$. □

(2.15) **Lemma.** *If $x\equiv(x_n)$ and $y\equiv(y_n)$ are real numbers with $x<y$, then there exists a rational number α with $x<\alpha<y$.*

Proof: By (2.4), we have $y-x\equiv(y_{2n}-x_{2n})_{n=1}^{\infty}$. Since $y-x\in\mathbb{R}^+$, by (2.7) there exists n in $\mathbb{Z}^+$ with $y_{2n}-x_{2n}>n^{-1}$. Write

$$\alpha\equiv\tfrac{1}{2}(x_{2n}+y_{2n}).$$

Then

$$\alpha-x\geqq\alpha-x_{2n}-|x_{2n}-x|\geqq\tfrac{1}{2}(y_{2n}-x_{2n})-(2n)^{-1}>0.$$

Also,

$$y-\alpha\geqq y_{2n}-\alpha-|y_{2n}-y|\geqq\tfrac{1}{2}(y_{2n}-x_{2n})-(2n)^{-1}>0.$$

Therefore $x<\alpha<y$. □

As a corollary, for each x in $\mathbb{R}$ and r in $\mathbb{R}^+$ there exists α in $\mathbb{Q}$ with $|x-\alpha|<r$. Here is another corollary.

(2.16) **Proposition.** *If $x_1,\ldots,x_n$ are real numbers with $x_1+\ldots+x_n>0$, then $x_i>0$ for some i $(1\leqq i\leqq n)$.*

Proof. By (2.15), there exists a rational number α with $0<\alpha<x_1+\ldots+x_n$. For $1\leqq i\leqq n$ let a_i be a rational number with

$$|x_i-a_i|<(2n)^{-1}\alpha.$$

Then

$$\sum_{i=1}^{n}a_i\geqq\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}|x_i-a_i|>\tfrac{1}{2}\alpha.$$

Therefore $a_i>(2n)^{-1}\alpha$ for some i. For this i it follows that

$$x_i\geqq a_i-|x_i-a_i|>0.\quad\square$$

(2.17) **Corollary.** *If x, y, and z are real numbers with $y<z$, then either $x<z$ or $x>y$.*

Proof: Since $z-x+x-y=z-y>0$, either $z-x>0$ or $x-y>0$, by (2.16). □

The next lemma gives an extremely useful method for proving inequalities of the form $x\leqq y$.

(2.18) **Lemma.** *Let x and y be real numbers such that the assumption $x>y$ implies that $0=1$. Then $x\leqq y$.*

Proof: Without loss of generality, we take $y=0$. For each n in $\mathbb{Z}^+$, either $x_n\leqq n^{-1}$ or $x_n>n^{-1}$. The case $x_n>n^{-1}$ is ruled out, since it implies that $x>0$. Therefore $-x_n\geqq-n^{-1}$ for all n, and so $-x\geqq0$. Thus $x\leqq0$. □

(2.19) **Theorem.** *Let (a_n) be a sequence of real numbers, and let x_0 and y_0 be real numbers with $x_0 < y_0$ Then there exists a real number x such that $x_0 \leqq x \leqq y_0$ and $x \neq a_n$ for all n in $\mathbb{Z}^+$.*

Proof: We construct by induction sequences (x_n) and (y_n) of rational numbers such that

(i) $x_0 \leqq x_n \leqq x_m < y_m \leqq y_n \leqq y_0 \quad (m \geqq n \geqq 1)$

(ii) $x_n > a_n$ or $y_n < a_n \quad (n \geqq 1)$

(iii) $y_n - x_n < n^{-1} \quad (n \geqq 1)$.

Assume that $n \geqq 1$ and that $x_0, \ldots, x_{n-1}$, $y_0, \ldots, y_{n-1}$ have been constructed. Either $a_n > x_{n-1}$ or $a_n < y_{n-1}$. In case $a_n > x_{n-1}$, let x_n be any rational number with $x_{n-1} < x_n < \min\{a_n, y_{n-1}\}$, and let y_n be any rational number with $x_n < y_n < \min\{a_n, y_{n-1}, x_n + n^{-1}\}$. Then the relevant inequalities are satisfied In case $a_n < y_{n-1}$, let y_n be any rational number with $\max\{a_n, x_{n-1}\} < y_n < y_{n-1}$, and x_n any rational number with $\max\{a_n, x_{n-1}, y_n - n^{-1}\} < x_n < y_n$. Again, the relevant inequalities are satisfied. This completes the induction.

From (i) and (iii) it follows that

$$|x_m - x_n| = x_m - x_n < y_n - x_n < n^{-1} \qquad (m \geqq n).$$

Similarly $|y_m - y_n| < n^{-1}$ for $m \geqq n$. Therefore $x \equiv (x_n)$ and $y \equiv (y_n)$ are real numbers. By (i) and (iii), they are equal. By (i), $x_n \leqq x$ and $y_n \geqq y$ for all n. If $a_n < x_n$, then $a_n < x$ and so $a_n \neq x$. If $a_n > y_n$, then $a_n > y = x$ and so $a_n \neq x$. Thus x has the required properties. □

Theorem (2.19) is the famous theorem of Cantor, that the real numbers are uncountable. The proof is essentially Cantor's "diagonal" proof. Both Cantor's theorem and his method of proof are of great importance.

The time has come to consider some counterexamples. Let (n_k) be a sequence of integers, each of which is either 0 or 1, for which we are unable to prove either that $n_k = 1$ for some k or that $n_k = 0$ for all k. This corresponds to what Brouwer calls "a fugitive property of the natural numbers". For example, such a sequence can be defined as follows. Let n_k be 0 if $u^t + v^t \neq w^t$ for all integers u, v, w, t with $0 < u, v, w \leqq k$ and $3 \leqq t \leqq 2 + k$. Otherwise let n_k be 1. Then we are unable to prove $n_k = 1$ for some k, because this would disprove Fermat's last theorem. We are unable to prove $n_k = 0$ for all k, because this would prove Fermat's last theorem.

Now define $x_k \equiv 0$ if $n_j = 0$ for all $j \leqq k$, and $x_k \equiv 2^{-m}$ otherwise, where m is the least positive integer such that $n_m = 1$. Then $x \equiv (x_k)$ is a

nonnegative real number, but we are unable to prove that $x>0$ or $x=0$. Since nothing is true unless and until it has been proved, it is untrue that $x>0$ or $x=0$.

Of course, if Fermat's last theorem is proved tomorrow, we shall probably still be able to define a fugitive sequence (n_k) of integers. Thus it is unlikely that there will ever exist a constructive proof that for every real number $x \geqq 0$ either $x>0$ or $x=0$. We express this fact by saying that there exists a real number $x \geqq 0$ such that it is *not* true that $x>0$ or $x=0$.

In much the same way we can construct a real number x such that it is *not* true that $x \geqq 0$ or $x \leqq 0$.

3. Sequences and Series of Real Numbers

We develop methods for defining a real number in terms of approximations by other real numbers.

(3.1) **Definition.** A sequence (x_n) of real numbers *converges* to a real number x_0 if for each k in $\mathbb{Z}^+$ there exists N_k in $\mathbb{Z}^+$ with

$$|x_n - x_0| \leqq k^{-1} \qquad (n \geqq N_k). \tag{3.1.1}$$

The real number x_0 is then called a *limit* of the sequence (x_n). To express the fact that (x_n) converges to x_0 we write

$$\lim_{n \to \infty} x_n = x_0$$

or

$$x_n \to x_0 \quad \text{as } n \to \infty$$

or simply $x_n \to x_0$.

A sequence (x_n) of real numbers is said to *converge*, or be *convergent*, if there exists a limit x_0 of (x_n).

It is easily seen that if (x_n) converges to both x_0 and x_0', then $x_0 = x_0'$.

A convergent sequence is *bounded:* there exists r in $\mathbb{R}^+$ such that $|x_n| \leqq r$ for all n

A convergent sequence of real numbers is not determined until the limit x_0 and the sequence (N_k) are given, as well as the sequence (x_n) itself. Even when they are not mentioned explicitly, these quantities are implicitly present. Similar comments apply to many subsequent definitions, including the following.