Schizophrenia in contemporary mathematics

Errett A. Bishop
University of California, San Diego
La Jolla, California

Distributed in conjunction with
the Colloquium Lectures given at the
University of Montana, Missoula, Montana
August 21—24, 1973,
seventy-eighth summer meeting of the
American Mathematical Society.

SCHIZOPHRENIA IN CONTEMPORARY MATHEMATICS

by

Errett Bishop

During the past ten years I have given a number of lectures on the subject of constructive mathematics. My general impression is that I have failed to communicate a real feeling for the philosophical issues involved. Since I am here today, I still have hopes of being able to do so. Part of the difficulty is the fear of seeming to be too negativistic and generating too much hostility. Constructivism is a reaction to certain alleged abuses of classical mathematics. Unpalatable as it may be to have those abuses examined, there is no other way to understand the motivations of the constructivists.

Brouwer's criticisms of classical mathematics were concerned with what I shall refer to as "the debasement of meaning". His incisive criticisms were one of his two main contributions to constructivism. (His other was to establish a new terminology, involving a re-interpretation of the usual connectives and quantifiers, which permits the expression of certain important distinctions of meaning which the classical terminology does not.)

The debasement of meaning is just one of the trouble spots in contemporary mathematics. Taken all together, these trouble spots indicate that something is lacking, that there is a philosophical deficit of major proportions. What it is that is lacking is perhaps not clear, but the lack in all of its aspects constitutes a syndrome I shall tentatively describe as "schizophrenia".

One could probably make a long list of schizophrenic attributes of contemporary mathematics, but I think the following short list covers most of the ground: rejection of common sense in favor of formalism; debasement of meaning by wilful refusal to accomodate certain aspects of reality; inappropriateness of means to ends; the esoteric quality of the communication; and fragmentation.

Common sense is a quality that is constantly under attack. It tends to be supplanted by methodology, shading into dogma. The codification of insight is commendable only to the extent that the resulting methodology is not elevated to dogma and thereby allowed to impede the formation of new insight. Contemporary mathematics has witnessed the triumph of formalist dogma, which had its inception in the important insight that most arguments of modern mathematics can be broken down and presented as successive applications of a few basic schemes. The experts now routinely equate the panorama of mathematics with the productions of this or that formal system. Proofs are thought of as manipulations of strings of symbols. Mathematical philosophy consists of the creation, comparison, and investigation of formal systems. Consistency is the goal. In consequence meaning is debased, and even ceases to exist at a primary level.

The debasement of meaning has yet another source, the wilful refusal of the contemporary mathematician to examine the content of certain of his terms, such as the phrase "there exists". He refuses to distinguish among the different meanings that might be ascribed to this phrase. Moreover he is vague about what meaning it has for him. When pressed he is apt to take refuge in formalistics, declaring that the meaning of the phrase and the statements of which it forms a part can only be understood in the context of the entire set of assumptions and techniques at his command. Thus he inverts the natural order, which would be to develop meaning first, and then to base his assumptions and techniques on the rock of meaning. Concern about this debasement of meaning is a principal force behind constructivism.

Since meaning is debased and common sense is rejected, it is not surprising to find that the means are inappropriate to the ends. Applied mathematics makes much of the concept of a model, as a tool for dealing with reality by mathematical means. When the model is not an adequate representation of reality, as happens only too often, the means are inappropriate. One gets the impression that some of the model-builders are no longer interested in reality. Their models have become autonomous. This has clearly happened in mathematical philosophy: the models (formal systems) are accepted as the preferred tools for investigating the nature of mathematics, and even as the font of meaning.

Everyone who has taught undergraduate mathematics must have been impressed by the esoteric quality of the communication. It is not natural for "A implies B" to mean "not A or B", and students will tell you so if you give them the chance. Of course, this is not a fatal objection. The question is, whether the standard definition of implication is useful, not whether it is natural. The constructivist, following Brouwer, contends that a more natural definition of implication would be more useful. This point will be developed later. One of the hardest concepts to communicate to the undergraduate is the concept of a proof. With good reason--the concept is esoteric. Most mathematicians, when pressed to say what they mean by a proof, will have recourse to formal criteria. The constructive notion of proof by contrast is very simple, as we shall see in due course. Equally esoteric, and perhaps more troublesome, is the concept of existence. Some of the problems associated with this concept have already been mentioned, and we shall return to the subject again. Finally, I wish to point out the esoteric nature of the classical concept of truth. As we shall see later, truth is not a source of trouble to the constructivist, because of his emphasis on meaning.

The fragmentation of mathematics is due in part to the vastness of the subject but it is aggravated by our educational system. A graduate student in pure mathematics may or may not be required to broaden himself by passing examinations in various branches of pure mathematics, but he will almost certainly not be required or even encouraged to acquaint himself with the philosophy of mathematics, its history, or its applications. We have geared ourselves to producing research mathematicians who will begin to write papers as soon as possible. This anti-social and anit-intellectual process defeats even its own narrow ends. The situation is not likely to change until we modify our conception of what mathematics is. Before important changes will come about in our methods of education and our professional values, we shall have to discover the significance of theorem and proof. If we continue to focus attention on the process of producing theorems, and continue to devalue their content, fragmentation is inevitable.

By devaluation of content I mean the following. To some pure mathematicians the only reason for attaching any interpretation whatever to theorem and proof is that the process of producing theorems and proofs is thereby facilitated. For them content is a means rather than the end. Others feel that it is important to have some content, but don't especially care to find out what it is. Still others, for whom Gödel (see for example [16]) seems to be a leading spokesman, do their best to develop content within the accepted framework of platonic idealism. One suspects that the majority of pure mathematicians, who belong to the union of the first two groups, ignore as much content as they possibly can. If this suspicion seems unjust, pause to consider the modern theory of probability. The probability of an event is commonly taken to be a real number between 0 and 1. One might naively expect that the probabilists would concern themselves with the computation of such real numbers. If so, a quick look at any one of a number of modern texts, for instance the excellent book of Doob [14], should suffice to disabuse him of that expectation. Fragmentation ensues, because much if not most of the theory is useless to someone who is concerned with actually finding probabilities. He will either develop his own semi-independent theories, or else work with ad hoc techniques and rules of thumb. I do not claim that reinvolvement of the probabilists with the basic questions of meaning would of itself reverse the process of fragmentation of their discipline, only that it is a necessary first step. In recent years a small number of constructivists (see [3], [9], [10], [11], [12], [23], and [24]) have been trying to help the probabilists take that step. Whether their efforts will ultimately be appreciated remains to be seen.

When I attempt to express in positive terms that quality in which contemporary mathematics is deficient, the absence of which I have characterized as "schizophrenia", I keep coming back to the term "integrity". Not the integrity of an isolated formalism that prides itself on the maintainance of its own standards of excellence, but an integrity that seeks common ground in the researches of pure mathematics, applied mathematics, and such mathematically oriented disciplines as physics; that seeks to extract the maximum meaning from each new development; that is guided primarily by considerations of content rather than elegance and formal attractiveness; that sees to it that the mathematical representation of reality does not degenerate into a game; that seeks to understand the place of mathematics in contemporary society. This integrity may not be possible of realization, but that is not important. I like to think of constructivism as one attempt to realize at least certain aspects of this idealized integrity. This presumption at least has the possible merit of preventing constructivism from becoming another game, as some constructivisms have tended to do in the past.

In discussing the principles of constructivism, I shall try to separate those aspects of constructivism that are basic to the philosophy from those that are merely convenient (or inconvenient, as the case may be). Four principles stand out as basic:

(A) Mathematics is common sense.

(B) Do not ask whether a statement is true until you know what it means.

(C) A proof is any completely convincing argument.

(D) Meaningful distinctions deserve to be maintained.

Surprisingly many brilliant people refuse to apply common sense to mathematics. A frequent attitude is that the formalization of mathematics has been of great value, because the formalism constitutes a court of last resort to settle any disputes that might arise concerning the correctness of a proof. Common sense tells us, on the contrary, that if a proof is so involved that we are unable to determine its correctness by informal methods, then we shall not be able to test it by formal means either. Moreover the formalism can not be used to settle philosophical disputes, because the formalism merely reflects the basic philosophy, and consequently philosophical disagreements are bound to result in disagreements about the validity of the formalism.

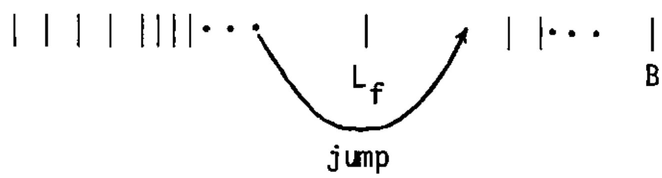
Principle (B) resolves the problem of constructive truth. For that matter, it would resolve the problem of classical truth if the classical mathematicians would accept it. We might say that truth is a matter of convention. This simply means that all arguments concerning the truth or falsity of any given statement about which both parties possess the same relevant facts occur because they have not reached a clear agreement as to what the statement means. For instance in response to the inquiry "Is it true the constructivists believe that not every bounded monotone sequence of real numbers converges?", if I am tired I answer "yes". Otherwise I tell the questioner that my answer will depend on what meaning he wishes to assign to the statement (*), that every bounded monotone sequence or real numbers converges. Moreover I tell him that once he has assigned a precise meaning to statement (*), then my answer to his question will probably be clear to him before I give it. The two meanings commonly assigned to (*) are the classical and the constructive. It seems to me that the classical mathematician is not as precise as he might be about the meaning he assigns to such a statement. I shall show you later one simple and attractive approach to the problem of meaning in classical mathematics. However in the case before us the intuition at least is clear. We represent the terms of the sequence by vertical marks marching to the right, but remaining to the left of the bound B.

$$\begin{array}{c} |\quad|\quad|\quad|\;\;|\;\;|\;\;|\;\;|\;|\,|\,|\,|\;\cdots \qquad\qquad | \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad B \end{array}$$

The classical intuition is that the sequence gets cramped, because there are infinitely many terms, but only a finite amount of space available to the left of B. Thus it has to pile up somewhere. That somewhere is its limit L.

$$\begin{array}{ccc} |\;|\;|\;|\;|\;|\cdots & \qquad | \qquad & | \\ & L & B \end{array}$$

The constructivist grants that some sequences behave in precisely this way. I call those sequences stupid. Let me tell you what a smart sequence will do. It will pretend to be stupid, piling up at a limit (in reality a false limit) L_f. Then when you have been convinced that it really is piling up at L_f, it will take a jump and land somewhere to the right!



Let us postpone a serious discussion of this example until we have discussed the constructive real number system. The point I wish to make now is that under neith interpretation will there be any disagreement as to the truth of (*), once that interpretation has been fixed and made precise.

Whenever a student asks me whether a proof he has given is correct, before answering his question I try to discover his concept of what constitutes a proof. Then I tell him my own concept, (C) above, and ask him whether he finds his argument completely convincing, and whether he thinks he has expressed himself clearly enough so that other informed and intellegent people will also be completely convinced.

Clearly it is impossible to accept (C) without accepting (B), because it doesn't make sense to be convinced that something is true unless you know what it means.

The question often arises, whether a constructivist would accept a non-constructive proof of a numerical result involving no existential quantifiers, such as Goldbach's conjecture or Fermat's last theorem. My answer is supplied by (C): I would want to examine the proof to see whether I found it completely convincing. Perhaps one should keep an open mind, but I find it hard to believe that I would find any proof that relied on the principle of the excluded middle for instance completely convincing. Fortunately the problem is hypothetical, because such proofs do not seem to arise. It does raise the interesting point that a classically acceptable proof of Goldbach's conjecture might not be constructively acceptable, and therefore the classical and the constructive interpretations of Goldbach's conjecture must differ in some fundamental respect. We shall see later that this is indeed the case.

Classical mathematics fails to observe meaningful distinctions having to do with integers. This basic failure reflects itself at all levels of the classical development of mathematics. Consider the number n_0, defined to be 0 if the Riemann hypothesis is true and 1 if it is false. The constructivist does not wish to prevent the classicist from working with such numbers (although he may personally believe that their interest is limited). He does want the calssicist to distinguish such numbers from numbers which can be "computed", such as the number n_1 of primes less than $10^{10^{10}}$. Classical mathematicians do concern themselves sporadically with whether numbers can be "computed", but only on an ad hoc basis. The distinction is not observed in the systematic development of classical mathematics, nor would the tools available to the classicist permit him to observe the distinction systematically even if he were so inclined.

The constructivists are frequently accused of displaying the same insensitivity to shades of meaning of which they accuse the classicist, because they do not distinguish between numbers that can be computed in principle, such as the number n_1 defined above, and numbers that can be computed in fact. Thus they violate their own principle (D). This is a serious accusation, and one that is not easy to refute. Rather than attempting to refute it, I shall give you my personal point of view. First, it may be demanding too much of the constructivist to ask them to lead the way in the development of usable and systematic methods for distinguishing computability in principle from computability in fact. If and when such methods are found, the constructivists will gratefully incorporate them into their mathematics. Second, it is by no means clear that such methods are going to be found. There is no fast distinction between computability in principle and in fact, because of the constant progress of the state of the art among other reasons. The most we can hope for is some good systematic measure of the efficiency of a computation. Until such is found, the problem will continue to be treated on an ad hoc basis.

I was careful not to call the number n_0 defined above an integer. Whether we do call it an integer is of no real importance, as long as we distinguish it in some way from numbers such as n_1. For instance we might call n_0 an integer and call n_1 a constructive integer. The constructivists have not accepted this terminology, in part because of Brouwer's influence, but also because it does not accord with their estimate of the relative importance of the two concepts. I shall reserve the term "integer" for what a classicist might call a constructive integer and put aside, at least for now, the problem of what would be an appropriate term for what is classically called an integer (assuming that the classical notion of an integer is indeed viable).

Thus we come to the crucial question, "What is an integer?" As we have already seen, the question is badly phrased. We are really looking for a definition of an integer that will be an efficient tool for developing the full content of mathematics. Since it is clear that we always work with representations of integers, rather than integers themselves (whatever those may be), we are really trying to define what we mean by a representation of an integer. Again, an integer is represented only when some intelligent agent constructs the representation, or establishes the convention that some artifact constitutes a representation. Thus in its final version the question is, "How does one represent an integer?" In practice we shall not be so meticulous as all this in our use of language. We shall simply speak of integers, with the understanding that we are really speaking of their representations. This causes no harm, because the original concept of an integer, as something invariant standing behind all of its representations, has just been seen to be superfluous. Moreover we shall not constantly trouble to point out that (representations of) integers exist only by virtue of conventions established by groups of intelligent beings. After this preliminary chatter, which may seem to have been unnecessary, we present our definition of an integer, dignified by the title of the

Fundamental Constructivist Thesis

Every integer can be converted in principle to decimal form by a finite, purely routine, process.

Note the phrase "in principle". It means that although we should be able to program a computer to produce the decimal form of any given integer, there are cases in which it would be naive to run the program and wait around for the result.

Everything else about integers follows from the above thesis plus the rules of decimal arithmetic that we learned in elementary school. Two integers are equal if their decimal representations are equal in the usual sense. The order relations and the arithmetic of integers are defined in terms of their decimal representations.

With the constructive definition of the integers, we have begun our study of the technical implementation of the constructivist philosophy. Our point of view is to describe the mathematical operations that can be carried out by finite beings, man's mathematics for short. In contrast, classical mathematics concerns itself with operations that can be carried out by God. For instance, the above number n_0 is classically a well-defined integer because God can perform the infinite search that will determine whether the Riemann hypothesis is true. As another example, the smart sequences previously discussed may be able to outwit you and me (or any other finite being), but they will not be able to outwit God. That is why statement (*) is true classically but not constructively.

You may think that I am making a joke, or attempting to put down classical mathematics, by bringing God into the discussion. This is not true. I am doing my best to develop a secure philosophical foundation, based on meaning rather than formalistics, for current classical practice. The most solid foundation available at present seems to me to involve the consideration of a being with non-finite powers--call him God or whatever you will--in addition to the powers possessed by finite beings.

What powers should we ascribe to God? At the very least, we should credit him with limited omniscience, as described in the following limited principle of omniscience (LPO for short): If $\{n_k\}$ is any sequence of integers, then either $n_k = 0$ for all k or there exists a k with $n_k \neq 0$. By accepting LPO as valid, we are saying that the being whose capabilities our mathematics describes is able to search through a sequence of integers to determine whether they all vanish or not.

Let us return to the technical developemnt of constructive mathematics, since it is simpler, and postpone the further consideration of classical mathematics until later. Our first task is to develop an appropriate language to describe the mathematics of finite beings. For this we are indebted to Brouwer. (See references [1], [6], [15], [20], and [21] for a more complete exposition than we are able to give here.) Brouwer remarked that the meanings customarily assigned to the terms "and", "or", "not", "implies", "there exists", and "for all" are not entirely appropriate to the constructive point of view, and he introduced more appropriate meanings as necessary.

The connective "and" causes no trouble. To prove "A and B", we must prove A and also prove B, as in classical mathematics. To prove "A or B" we must give a finite, purely routine method which after a finite number of steps either leads to a proof of A or to a proof of B. This is very different from the classical use of "or"; for example LPO is true classically, but we are not entitled to assert it constructively because of the constructive meaning of "or".

The connective "implies" is defined classically by taking "A implies B" to mean "not A or B". This definition would not be of much value constructively. Brouwer therefore defined "A implies B" to mean that there exists an argument which shows how to convert an arbitrary proof of A into a proof of B. To take an example, it is clear that "{(A implies B) and (B implies C)} implies (A implies C)" is always true constructively; the argument that converts arbitrary proofs of "A implies B" and "B implies C" into a proof of "A implies C" is the following: given any proof of A, convert it into a proof of C by first converting it into a proof of B and then converting that proof into a proof of C.

We define "not A" to mean that A is contradictory. By this we mean that it is inconceivable that a proof of A will ever be given. For example, "not 0 = 1" is a true statement. The statement "0 = 1" means that when the numbers "0" and "1" are expressed in decimal form, a mechanical comparison of the usual sort checks that they are the same. Since they are already in decimal form, and the comparison in question shows they are not the same, it is impossible by correct methods to prove that they are the same. Any such proof would be defective, either technically or conceptually. As another example, "not (A and not A)" is always a true statement, because if we prove not A it is impossible to prove A--therefore, it is impossible to prove both.

Having changed the meaning of the connectives, we should not be surprised to find that certain classically accepted modes of inference are no longer correct. The most important of these is the principle of the excluded middle--"A or not A". Constructively, this principle would mean that we had a method which in finitely many, purely routine, steps would lead to a proof of disproof of an arbitrary mathematical assertion A. Of course we have no such method, and nobody has the least hope that we ever shall. It is the principle of the excluded middle that accounts for almost all of the important unconstructivities of classical mathematics. Another incorrect principle is "(not not A) implies A". In other words, a demonstration of the impossibility of the impossibility of a certain construction, for instance, does not constitute a method for carrying out that construction.

I could proceed to list a more or less complete set of constructively valid rules of inference involving the connectives just introduced. This would be superfluous. Now that their meanings have been established, the rest is common sense. As an exercise, show that the statment

$$\text{"}(A \to 0 = 1) \leftrightarrow \text{not } A\text{"}$$

is constructively valid.

The classical concept of a set as a collection of objects from some pre-existent universe is clearly inappropriate constructively. Constructive mathematics does not postulate a pre-existent universe, with objects lying around waiting to be collected and grouped into sets, like shells on a beach. The entities of constructive mathematics are called into being by the constructing intelligence. From this point of view, the very question "What is a set?" is suspect. Rather we should ask the question, "What must one do to construct a set?". When the question is posed this way, the answer is not hard to find.

Definition. To construct a set, one must specify what must be done to construct an arbitrary element of the set, and what must be done to prove two arbitrary elements of the set are equal. Equality so defined must be shown to be an equivalence relation.

As an example, let us construct the set of rational numbers. To construct a rational number, define integers p and q and prove that $q \neq 0$. To prove that the rational numbers p/q and p_1/q_1 are equal, prove $pq_1 = p_1q$.

While we are on the subject, we might as well define a function $f : A \to B$. It is a rule that to each element x of A associates an element $f(x)$ of B, equal elements of B being associated to equal elements of A.

The notion of a subset A_0 of a set A is also of interest. To construct an element of A_0, one must first construct an element of A, and then prove that the element so constructed satisfies certain additional conditions, characteristic of the particular subset A_0. Two elements of A_0 are equal if they are equal as elements of A.

Contrary to classical usage, the scope of the equality relation never extends beyond a particular set. Thus it does not make sense to speak of elements of different sets as being equal, unless possibly those different sets are both subsets of the same set. This is because for the constructivist equality is a convention, whose scope is always a given set; all this is conceptually quite distinct from the classical concept of equality as identity. You see now why the constructivist is not forced to resort to the artifice of equivalence classes!

After this long digression, consider again the quantifiers. Let $A(x)$ be a mathematical assertion depending on a parameter x ranging over a set S. To prove "$\forall x A(x)$", we must give a method which to each element x of S associates a proof of $A(x)$. Thus the meaning of the universal quantifier "$\forall$" is essentially the same as it is classically.

We expect the existential quantifier "$\exists$", on the other hand, to have a new meaning. It is not clear to the constructivist what the classicist means when he says "there exists". Moreover, the existential quantifier is just a glorified version of "or", and we know that a reinterpretation of this connective was necessary. Let the variable x range over the set S. Then to prove "$\exists x A(x)$" we must construct an element x_0 of S, according to the principles laid down in the definition of S, and then prove the statement "$A(x_0)$".

Again, certain classical uses of the quantifiers fail constructively. For example, it is not correct to say that "not $\forall x A(x)$ implies $\exists x$ not $A(x)$." On the other hand, the implication "not $\exists x A(x)$ implies $\forall x$ not $A(x)$" is constructively valid. I hope all this accords with your common sense, as it does with mine.

Perhaps you see an objection to these developments--that they appear to violate constructivist principle (D) above. By accomodating our terminology to the mathematics of finite beings, have we not replaced the classical system, that does not permit the systematic development of constructive meaning, by a system that does not permit the systematic development of classical meaning? In my opinion the exact opposite is true--the constructive terminology just introduced affords as good a framework as is presently available for expressing the content of classical mathematics.

If you wish to do classical mathematics, first decide what non-finite attributes you are willing to grant to God. You may wish to grant him LPO and no others. Or you may wish to be more generous and grant him EM, the principle of the excluded middle, possibly augmented by some version of the axiom of choice. When you have made your decision, avail yourself of all the apparatus of the constructivist, and augment it by those additional powers (LPO or EM or whatever) that you have granted to God. Although you will be able to prove more theorems than the constructivist will, because your being is more powerful than his, his theorems will be more meaningful than yours. Moreover to each of your theorems he will be able to associate one of his, having exactly the same meaning. For example, if LPO is the only non-finite attribute of your God, then each of your theorems "A" he will restate and prove as "LPO implies A". Clearly the meaning will be preserved. On the other hand, if he proves a theorem "B", you will also be able to prove "B", but your "B" will be less meaningful than his. The classical interpretation of even such simple results as Goldbach's conjecture is weaker than the constructive interpretation. In both cases the same phenomena--the results of certain finitely performable computations--are predicted, but the degree of conviction that the predicted phenomena will actually be observed is greater in the constructive case, because to trust the classical predictions one must believe in the theoretical validity of the concept of a God having the specified attributes, whereas to trust the constructive predictions one must only believe in the theoretical validity of the concept of a being who is able to perform arbitrarily involved finite operatio

It would thus appear that even a constructive proof of such a result as "the number of zeros in the first n digits of the decimal expansion of Π does not exceed twice the number of ones" would leave us in some doubt as to whether the prediction is correct for any particular value of n, say for $n = 1000$. We have brought mathematics down to the gut level. My gut tells me to trust the construct prediction and be wary of the classical prediction. I see no reason that yours should not tell you to trust both, or to trust neither.

In common with other constructivists, I also have gut feelings about the relative merits of the classical and constructive versions of those results which, unlike Goldbach's conjecture, assert the existence of certain quantities. If we let "A" be any such result, with the constructive interpretation, then the constructive version of the corresponding classical result will be (for instance) "LPO $\rightarrow$ A", as we have seen. My feeling is that it is likely to be worth whatever extra effort it takes to prove "A" rather than "LPO $\rightarrow$ A".

The linguistic developments I have outlined could be taken as the basis for a formalization of constructive (and therefore of classical) mathematics. So as not to create the wrong impression, I wish to emphasise again certain points that have already been made.

Formalism

The devil is very neat. It is his pride
To keep his house in order. Every bit
Of trivia has its place. He takes great pains
To see that nothing ever does not fit.

And yet his guests are queasy. All their food,
Served with a flair and pleasant to the eye,
Goes through like sawdust. Pity the perfect host!
The devil thinks and thinks and he cannot cry.

Constructivism

Computation is the heart
Of everything we prove.
Not for us the velvet wisdom
Of a softer love.

If Aphrodite spends the night,
Let Pallas spend the day.
When the sun dispels the stars
Put your dreams away.

There are at least two reasons for developing formal systems for constructive mathematics. First, it is good to state as concisely and systematically as we are able some of the objects, constructions, terminology, and methods of proof. The development of formal systems that catch these aspects of constructive practice should help to sharpen our understanding of how best to organize and communicate the subject. Second and more important, informal mathematics is the appropriate language for communicating with people, but formal mathematics is more appropriate for communicating with machines. Modern computer languages (see the report [30], for example), while rich in facilities, seem to be lacking in philosophical scope. It might be worthwhile to investigate the possibility that constructive mathematics would afford a solid philosophical basis for the theory of computation, and constructive formalism a point of departure for the development of a better computer language. Certainly recursive function theory, which has played a central role in the philosophy of computation, is inadequate to the task.

The development of a constructive formalism at any given level would seem to be no more difficult than the development of a classical formalism at the same level. See [17], [18], [20], [21], [22], and [27] for examples. For a discussion of constructive formalism as a computer language see [2].

Let us return to the technical development of constructive mathematics, and ask what is meant constructively by a function $f : \mathbb{Z} \to \mathbb{Z}$ (where $\mathbb{Z}$ is the set of integers). We improve the classical treatment right away-instead of talking about ordered pairs, we talk about rules. Our definition takes a function $F : \mathbb{Z} \to \mathbb{Z}$ to be a rule that associates to each (constructively defined) integer n a (constructively defined) integer $f(n)$, equal values being associated to equal arguments. For a given argument n, the requirement that $f(n)$ be constructively defined means that its decimal form can be computed by a finite, purely routine process. That's all there is to it. Functions $f : \mathbb{Z} \to Q$, $f : Q \to Q$, $f : \mathbb{Z}^+ \to Q$ are defined similarly. (Here Q is the set of rational numbers and $\mathbb{Z}^+$ the set of positive integers.) A function with domain $\mathbb{Z}^+$ is called a sequence, as usual.

Now that we know what a sequence of rational numbers is, it is easy to define a real number. A real number is a Cauchy sequence of rational numbers! Again, I have improved on the classical treatment, by not mentioning equivalence classes. I shall never mention equivalence classes. To be sure we completely understood this definition, let us expand it a bit. Real numbers are not pre-existent entities, waiting to be discovered. They must be constructed. Thus it is better to describe how to construct a real number, than to say what it is. To construct a real number, one must

(a) construct a sequence $\{x_n\}$ of rational numbers

(b) construct a sequence $\{N_n\}$ of integers

(c) prove that for each positive integer n we have

$$|x_i - x_j| \le \frac{1}{n} \text{ whenever } i \ge N_n \text{ and } j \ge N_n.$$

Of course, the proof (c) must be constructive, as well as the constructions (a) and (b).

Two real numbers $\{a_n\}$ and $\{b_n\}$ (the corresponding convergence parameters (b) and proofs (c) are assumed without explicit mention) are said to be equal if for each positive integer k there exists a positive integer N_k such that $|a_n - b_n| \leq \frac{1}{k}$ whenever $n \geq N_k$. It can be shown that this notion of equality is an equivalence relation. Addition and multiplication of real numbers are also defined in the same way as they are defined classically. The order relation, on the other hand, is more interesting. If $a = \{a_n\}$ and $b = \{b_n\}$ are real numbers, we define $a < b$ to mean that there exist positive integers M and N such that $a_n \leq b_n - \frac{1}{M}$ whenever $n \geq N$. Then it is easily shown that $a < b$ and $b < c$ imply $a < c$, that $a < b$ implies $a - c < b - c$, and so forth. Some care must be exercised in defining the relation $\leq$. We could define $a \leq b$ to mean that either $a < b$ or $a = b$. An alternate definition would be to define it to mean that $b < a$ is contradictory. We shall not use either of these, although our definition turns out to be equivalent to the latter.

Definition. $a \leq b$ means that for each positive integer M there exists a positive integer N such that $b_n \geq a_n - \frac{1}{M}$ whenever $n \geq N$.

To make the choice of this definition plausible, I shall construct a certain real number H.

$$H = \sum_{n=1}^{\infty} \alpha_n 2^{-n}$$

where $\alpha_n = 0$ in case every even integer between 4 and n is the sum of two primes, and $\alpha_n = 1$ otherwise. (More precisely, H is given by the Cauchy sequence $\{a_n\}$, with $a_n = \sum_{k=1}^{n} \alpha_k 2^{-k}$, and the sequence $\{N_n\}$ of convergence parameters, where $N_n = n$.) Clearly we wish to have $H \geq 0$. It certainly is according to the definition we have chosen. (The real number 0 of course is the Cauchy sequence of rational numbers all of whose terms are 0.) On the other hand, we would not be entitled to assert that $H \geq 0$ if we had defined $H \geq 0$ to mean that either $H > 0$ or $H = 0$, because the assertion "$H > 0$ or $H = 0$" means that we have a finite, purely routine method for deciding which; in this case, a finite, purely routine method for proving or disproving Goldbach's conjecture!

Most of the usual theorems about $\leq$ and $<$ remain true constructively, with the exception of trichotomy. Not only does the usual form "$a < b$ or $a = b$ or $a > b$" fail, but such weaker forms as "$a < b$ or $a \geq b$", or even "$a \leq b$ or $a \geq b$" fail as well. For example, we are not entitled to assert "$0 < H$ or $0 = H$ or $0 > H$". If we consider the closely related number $H' = \sum_{n=1}^{\infty} a_{2n}(-2)^{-n}$, we are not even entitled to assert that "$H' \geq 0$ or $H' \leq 0$".

Since trichotomy is so fundamental, we might expect constructive mathematics to be hopelessly enfeebled because of its failure. The situation is saved, because trichtomy does have a constructive version, which of course is considerably weaker than the classical.

<u>Theorem</u>. For arbitrary real numbers a, b, and c, with $a < b$, either $c > a$ or $c < b$.

<u>Proof</u>. Choose integers M and N_0 such that $a_n \leq b_n - \frac{1}{M}$ whenever $n \geq N_0$. Choose integers N_a, N_b, and N_c such that $|a_n - a_m| \leq (6M)^{-1}$ whenever $n, m \geq N_a$, $|b_n - b_m| \leq (6M)^{-1}$ whenever $n, m \geq N_b$, $|c_n - c_m| \leq (6M)^{-1}$ whenever $n, m \geq N_c$. Set $N = \max\{N_0, N_a, N_b, N_c\}$. Since a_N, b_N, and c_N are all

rational numbers, either $c_N < \frac{1}{2}(a_N + b_N)$ or $c_N \geq \frac{1}{2}(a_N + b_N)$. Consider first the case $c_N \geq \frac{1}{2}(a_N + b_N)$. Since $a_N \leq b_N - M^{-1}$, it follows that $a_N \leq c_N - (2M)^{-1}$. For each $n \geq N$ we therefore have

$$\begin{aligned} a_n &\leq a_N + (6M)^{-1} \leq c_N - (2M)^{-1} + (6M)^{-1} \\ &\leq c_n + (6M)^{-1} - (2M)^{-1} + (6M)^{-1} = c_n - (6M)^{-1}. \end{aligned}$$

Therefore $a < c$. In the other case, $c_N < \frac{1}{2}(a_N + b_N)$, it follows similarly that $c < b$. This completes the proof of the theorem.

Do not be deceived by the use of the word "choose" in the above proof, which is simply a carry-over from classical usage. No choice is involved, because M and N_0, for instance, are fixed positive integers, defined explicitly by the proof of the inequality $a < b$. Of course we could decide to substitute other values for the original values of M and N_0, if we desired, so some choice is possible should we wish to exercise it. If we do not explicitly state what choice we wish to make, it will be assumed that the values of M and N_0 given by the proof of $a < b$ are chosen.

The number H, which is constructively a well-defined real number, is classically rational, because if the Goldbach conjecture is true then $H = 0$, and if the conjecture is false then $H = 2^{-n+1}$, where n is the first even integer for which it fails. We are not entitled to assert constructively that H is rational: if it is rational, then either $H = 0$ or $H \neq 0$, meaning that either Goldbach's conjecture is true or else it is false; and we are not entitled to assert this constructively, until we have a method for deciding which. We are not entitled to assert H is irrational either, because if H is irrational, then $H \neq 0$, therefore Goldbach's conjecture is false, therefore H is the *rational number* 2^{-n+1}, *a contradiction! Thus* H *cannot be asserted to be* rational, although its irrationality is contradictory. (I am indebted to Halsey Royden for this amusing observation.)

It is easy to prove the existence of many irrational numbers, by proving the uncountability of the real numbers, as a corollary of the Baire category theorem. For the present, let us merely remark that $\sqrt{2}$ is irrational. Of course, $\sqrt{2}$ can be defined by constructing successive decimal approximations. It is therefore constructively well-defined. The classical proof of the irrationality of $\sqrt{2}$ shows that if $\frac{p}{q}$ is any rational number then $\frac{p^2}{q^2} \neq 2$. Since both $\frac{p^2}{q^2}$ and 2 can be written with denominator q^2, it follows that

$$\left|\frac{p}{q} - \sqrt{2}\right| \cdot \left|\frac{p}{q} + \sqrt{2}\right| = \left|\frac{p^2}{q^2} - 2\right| \geq \frac{1}{q^2} .$$

Since clearly $\frac{p}{q} \neq \sqrt{2}$ in case $\frac{p}{q} < 0$ or $\frac{p}{q} > 2$, to show that $\frac{p}{q} \neq \sqrt{2}$ we may assume $0 \leq \frac{p}{q} \leq 2$. Then

$$\left|\frac{p}{q} - \sqrt{2}\right| \geq \left|\frac{p}{q} + \sqrt{2}\right|^{-1} \cdot \frac{1}{q^2} \geq |2 + 2|^{-1} \cdot \frac{1}{q^2} = \frac{1}{4q^2} .$$

Therefore $\sqrt{2} \neq \frac{p}{q}$. Thus $\sqrt{2}$ is (constructively) irrational.

The failure of the usual form of trichotomy means that we must be careful in defining absolute values and maxima and minima of real numbers. For example, if $x = \{x_n\}$ is a real number, with sequence $\{N_n\}$ of convergence parameters, then $|x|$ is defined to be the Cauchy sequence $\{|x_n|\}$ of rational numbers (with sequence $\{N_n\}$ of convergence parameters). Similarly, min $\{x, y\}$ is defined to be the Cauchy sequence $\{\min\{x_n, y_n\}\}_{n=1}^{\infty}$, and max $\{x, y\}$ to be $\{\max\{x_n, y_n\}\}_{n=1}^{\infty}$.

This definition of min, in particular, has an amusing consequence. Consider the equation

$$x^2 - xH' = 0.$$

Clearly 0 and the number H' are solutions. Are they the only solutions? It depends on what we mean by "only". Clearly min $\{0, H'\}$ is a solution, and we are unable to identify it with either 0 or H'. Thus it is a third solution! The reader might like to amuse himself looking for others. This discussion incidentally makes the point that if the product of two real numbers is 0 we are not entitled to conclude that one of them is 0. (For example, $x(x - H') = 0$ does not imply that $x = 0$ or $x - H' = 0$: set $x = \min\{0, H'\}$.)

The constructive real number system as I have described it is not accepted by all constructivists. The intuitionists and the recursive function theorists have other versions.

For Brouwer, and his followers (the intuitionists), the constructive real numbers described above do not constitute all of the real number system. In addition there are incompletely determined real numbers, corresponding to sequences of rational numbers whose terms are not specified by a master algorithm. Such sequences are called "free-choice sequences", because the creating subject, who defines the sequence, does not completely commit himself in advance but allows himself some freedom of choice along the way in defining the individual terms of the sequence.

There seem to be at least two motivations for the introduction of free-choice sequences into the real number system. First, since each constructive real number can presumably be described by a phrase in the English language, on superficial consideration the set of constructive real numbers would appear to be countable. On closer consideration this is seen not to be the case: Cantor's uncountability theorem holds, in the following version. If $\{x_n\}$ is any sequence of real numbers, there exists a real number x with $x \neq x_n$ for all n. Nevertheless it appears that Brouwer was troubled by a certain aura of the discrete clinging to the constructive real number system $\mathbb{R}$. Second, every function anyone has ever been able to construct from $\mathbb{R}$ to $\mathbb{R}$ has turned out to be continuous, in fact uniformly continuous on bounded subsets. (The function f that is 1 for $x \geq 0$ and 0 for $x < 0$ does not count, because for those real numbers x for which we have no proof of the statement "$x \geq 0$, or $x < 0$" we are unable to compute $f(x)$.) Brouwer had hopes of proving that every function from $\mathbb{R}$ to $\mathbb{R}$ is continuous, using arguments involving free choice sequences. He even presented such a proof [7]. It is fair to say that almost nobody finds his proof intelligible. It can be made intelligible by replacing Brouwer's arguments at two critical points by axioms, that Kleene and Vesley [21] call "Brouwer's principle" and "the bar theorem". My objection to this is, that by introducing such a theorem as "all $f : \mathbb{R} \to \mathbb{R}$ are continuous" in the guise of axioms, we have lost contact with numerical meaning. Paradoxically this terrible price buys little or nothing of real mathematical value. The entire theory of free-choice sequences seems to me to be made of very tenuous mathematical substance.

If it is fair to say that the intuitionists find the constructive concept of a sequence generated by an algorithm too precise to adequately describe the real number system, the recursive function theorists on the other hand find it too vague. They would like to specify more precisely what is meant by an algorithm, and they have a candidate in the notion of a recursive function. They admit only sequence of integers or rational numbers that are recursive (a concept we shall not define here: see [20] for détails). Their reasons are, that the concept is more precise than the naive concept of an algorithm, that every naively defined algorithm has turned out to be recursive, and it seems unlikely we shall ever discover an algorithm that is not recursive. This requirement that every sequence of integers must be recursive is wrong on three fundamental grounds. First and most important, there is no doubt that the naive concept is basic, and the recursive concept derives whatever importance it has from some presumption that every algorithm will turn out to be recursive. Second, the mathematics is complicated rather than simplified by the restriction to recursive sequences. If there is any doubt as to this, it can be resolved by comparing some of the recursivist developments of elementary analysis with the constructivist treatment of the same material. Even if one is oriented to running material on a computer, the recursivist formulation would constitute an obstacle, because very likely the recursive presentation would be translated into computer language by first translating into common constructive terminology (at least mentally) and then translating that into the language of whatever computer was being used. Third, no gain in precision is actually achieved. One of the procedures for defining the value of a recursive function is to search a sequence of integers one by one, and choose the first that is non-zero, having first proved that one of them is non-zero. Thus the notion of a recursive function is at least as imprecise as the notion of a correct proof. The latter notion is certainly no more precise than the naive notion of a (constructive) sequence of integers.

The desire to achieve complete precision, whatever that is, is doomed to frustration. What is really being sought is a way to guarantee that no disagreements will arise. Mathematics is such a complicated activity that disagreements are bound to arise. Moreover, mathematicians will always be tempted to try out new ideas that are so complicated or innovative that their meaning is questionable. What is important is not to develop some framework, such as recursive function theory, in the vain hope of forestalling questionable innovations, but rather to subject every development to intense scrutiny (in terms of the meaning, not on formal grounds).

Recursive functions come into their own as the source of certain counter-examples in constructive mathematics, the most famous being the word-problem in the theory of groups. Since the concept of a (constructively) recursive sequence is narrower than the concept of a (constructive) sequence, it is easier to demonstrate that there exist no recursive sequences satisfying a given condition G. Such a demonstration makes it extremely unlikely that a (constructive) sequence satisfying G will be found without some radically new method for defining sequences being discovered, a discovery that many view as almost out of the question.

Although some every beautiful counter-examples have been given by means of recursive functions, they have also been used as a source of counter-examples in many situations in which a prior technique due to Brouwer [20] would have been both simpler and more convincing. Brouwer's idea is to counterexample a theorem A by proving $A \to \mathrm{LPO}$. Since nobody seriously thinks LPO will ever be proved, such a counter-example affords a good indication that A will never be proved. As an instance, Brouwer has shown that the statement that every bounded monotone sequence of real numbers converges implies LPO.

Another source of Brouwerian counter-examples is the statement LLPO (for the "lesser limited principle of omniscience"), that if $\{n_k\}$ is any sequence of integers, then either the first non-zero term, if one exists, is even or else the first non-zero term, if one exists, is odd. Clearly LPO $\rightarrow$ LLPO, but there seems to be no way to prove that LLPO $\rightarrow$ LPO. Nevertheless, we are just as sceptical that LLPO will ever be proved. Thus A $\rightarrow$ LLPO is another type of *Brouwerian counter-example for* A. As an instance, the statement that "either $x \geq 0$ or $x \leq 0$ for each real number x" implies LLPO, in fact is equivalent to it.

Thus we are so sceptical that the statements LPO, LLPO, and their ilk will ever be proved that we use them for building counter-examples. The strongest counter-example to A would be to show that a proof of A is inconceivable, in other words to prove "not A", but proving "A $\rightarrow$ LPO" or "A $\rightarrow$ LLPO" is almost as good. In fact, I personally find it inconceivable that LPO (or LLPO for that matter) will ever be proved. Nevertheless I would be reluctant to accept "not LPO" as a theorem, because my belief in the impossibility of proving LPO is more of a gut reaction prompted by experience than something I could communicate by arguments I feel would be sure to convince any objective, well-informed, and intelligent person. The acceptance of "not LPO" as a theorem would have one amusing consequence, that the theorems of constructive mathematics would not necessarily be classically valid (on a formal level) any longer. It seems we are doomed to live with "LPO" and "there exists a function from $[0, 1]$ to $\mathbb{R}$ that is not uniformly continuous" and similar statements, of whose impossibilities we are not quite sure enough to assert their negations as theorems.

The classical paradoxes are equally viable constructively, the simplest perhaps being "this statement is false." The concept of the set of all sets seems to be paradoxical (i.e., lead to a contradiction) constructively as well as classically. Informed common sense seems to be the best way of avoiding these paradoxes of self reference. Their spectre will always be lurking over both classical and constructive mathematics. Hermann Weyl made the meticulous avoidance of self reference the basis of a whole new development of the real number system (see Weyl [32]) that has since become known as predicative mathematics. Weyl later abandoned his system in favor of intuitionism. I see no better course at present than to recognise that certain concepts are inherently inconsistent and to familiarize oneself with the dangers of self-reference.

Not only is there insufficient time, but I would not be competent to review all of the recent advances of constructive mathematics, including those of ad hoc constructivism as well as those taking place under constructivist philosophies at variance with those that I have presented here, for example the recursivist constructivism of Markov and his school in Russia. (I have been told that some of the recent advances in differential equations have tended to present that subject in a more constructive light. Perhaps Felix Browder will give us some information about those developments.) I shall restrict myself in what remains to selected developments with which I am familiar, that seem to me to be of special interest.

Brouwer [6] was the first to develop a constructive theory of measure and integration, and the intuitionist tradition (see [19] and [31] for instance) in Holland carried the development further, working with Lebesgue measure on $\mathbb{R}^n$. In [1] I worked with arbitrary measures (both positive and negative) on locally compact spaces, recovering much of the classical theory. The Daniell integral was developed in full generality in [5]. The concept of an integration space postulates a set X, a linear subset L of the set of all partially-defined functions from X to $\mathbb{R}$, and a linear functional I from L to $\mathbb{R}$ *having the properties*

(1) if $f \in L$, then $|f| \in L$ and $\min\{f, 1\} \in L$

(2) if $f \in L$ and $f_n \in L$ for each n, such that $f_n \geq 0$ and

$\sum_{n=1}^{\infty} I(f_n)$ converges to a sum that is less than $I(f)$, then

$\sum_{n=1}^{\infty} f_n(x)$ converges and is less than $f(x)$, for some x in

the common domain of f and the functions f_n

(3) $I(p) \neq 0$ for some $p \in L$

(4) $\lim_{n\to\infty} I(\min\{f, n\}) = I(f)$ and $\lim_{n\to\infty} I(\min\{|f|, n^{-1}\}) = 0$ for all f in L.

We define L_1 to consist of all partially defined functions f from X to $\mathbb{R}$ such that there exists a sequence $\{f_n\}$ of elements of L such that (a) $\sum_{n=1}^{\infty} I(|f_n|)$ converges and (b) $\sum_{n=1}^{\infty} f_n(x) = f(x)$ whenever $\sum_{n=1}^{\infty} |f_n(x)|$ converges.

It turns out to be possible to extend I to L_1, in such a way that (X, L_1, I) also satisfy the axioms, and in addition L_1 is complete under the metric $\rho(f, g) = I(|f - g|)$.

The only real problem in recovering the classical Daniell theory is posed by the classical result that if $f \in L_1$ then the set $A_t = \{x \in X : f(x) \geq t\}$

is integrable for all $t > 0$ (in the sense that its characteristic function χ_t, defined by $\chi_t(x) = 1$ if $f(x) \geq t$ and $\chi_t(x) = 0$ if $f(x) < t$, is in L_1). The constructive version is that A_t is integrable for all except countably many $t > 0$. The proof of this requires a rather complex theory, called the theory of profiles. Y. K. Chan informs me that he has been able to simplify the theory of profiles considerably. He has also effected a considerable simplification in another trouble-spot of [5], the proof that a non-negative linear functional I on the set $L = C(X)$ of continuous functions on a compact space X satisfies the axioms for an integration space presented above. (Axiom (2) is the troublemaker.)

Constructive integration theory affords the point of departure for some *recent constructivizations of parts of probability theory. There is no* (constructive) way to prove even the simplest cases of the ergodic theorem, such that if T denotes rotation of a circle X through an angle α, then for each integrable function $f : X \to \mathbb{R}$ and almost all x in X, the averages

$$f_N(x) = \frac{1}{N} \sum_{n=1}^{N} f(T^n x)$$

converge. (The difficulty comes about because we are unable to decide for instance whether $\alpha = 0$.) One way to recover the essence of the ergodic theorem constructively, and in fact deepen it considerably, is to show that the sequence $\{f_N\}$ satisfies certain integral inequalities, analogous to the upcrossing inequalities (see [14]) of martingale theory. This was done in the context of the Chacon-Ornstein ergodic theorem in [1], and even more generally in [3].

John Nuber [23] takes another route. He presents sufficient conditions, close to being necessary, that the sequence $\{f_N\}$ actually converges a.e., in the context of the classical Birkhoff ergodic theorem. More recently, in an unpublished manuscript, he has generalized his conditions to the context of the classical Chacon-Ornstein theorem.

Y. K. Chan has done much to constructivize the theory of stochastic processes. His paper [10] unifies the two classically distinct cases of the renewal theorem into one constructive result. His paper [12] contains the following theorem:

Theorem. Let μ_1 and μ_2 be probability measures on $\mathbb{R}$, and f_1 and f_2 their characteristic functions (Fourier transforms). Let g be a continuous function on $\mathbb{R}$, with $|g| \leq 1$. Then for every $\varepsilon > 0$ there exist $\delta > 0$ and $\theta > 0$, depending only on ε and the moduli of continuity of f_1, f_2, and g, such that

$$\left|\int g d\mu_1 - \int g d\mu_2\right| < \varepsilon$$

whenever $|f_1 - f_2| < \delta$ on $[-\theta, \theta]$.

A simple corollary is Levy's theorem, that if $\{\mu_n\}$ is a sequence of probability measures on $\mathbb{R}$, whose characteristic functions $\{f_n\}$ converge uniformly on compact sets to some function f, then μ_n converges weakly to a probability measure μ whose characteristic function is f.

Levy's theorem is classically an important tool for proving convergence of measures. Chan shows that this is also true constructively, by using it to get constructive proofs of the central limit theorem and of the Levy-Khintchine formula for infinitely divisible distributions.

Chan's papers [9] and [11] are primarily concerned with the problem of constructing a stochastic process. In [9] he gives a constructive version of Kolmogorov's extension theorem. In [11], he constructivizes the classical derivation of a time homogeneous Markov process from a strongly continuous semi-group of transition operators. In addition he proves Ionescu Tulcea's theorem and a supermartingale convergence theorem.

H. Cheng [13] has given a very pretty version of the Riemann mapping theorem and Caratheodory's results on the convergence of mapping functions. He defines a simply connected proper open subset U of the complex plane $\mathbb{C}$ to be mappable relative to some distinguished point z_0 of U if for each $\varepsilon > 0$ there exist finitely many points $z_1, \ldots, z_n$ in the complement of U such that any continuous path beginning at z_0 and having distance $\geq \varepsilon$ from each of the points $z_1, \ldots, z_n$ lies entirely in U. He shows that mappability is necessary and sufficient for the existence of a mapping function. He goes on to study the dependence of the mapping function on the domain, by defining natural metrics on the space D of domains (with distinguished points z_0) and the space M of mapping functions, and proving that the function $\lambda : D \to M$ that associates to each domain its mapping function is a homeomorphism. He thus extends and constructivizes the classical Caratheodory results. Many of his estimates are similar to those developed by Warschawski in his studies of the mapping function.

The problem of constructivizing the classical theory of uniformization is still open. (Even reasonable conjectures seem difficult to come by.) So is the problem of (constructively) constructing canonical maps for multiply connected domains, as far as I know.

It is natural to define two sets to have the same cardinality if they are in one-one correspondence. The constructive theory of cardinality seems hopelessly involved, due to the constructive failure of the Cantor-Bernstein lemma, and for other reasons as well.

Progress has been made however in constructivizing the theory of ordinal numbers. Brouwer [8] defines ordinals to be ordered sets that are built up from non-empty finite sets by finite and countable addition. F. Richman [26] gives a more general definition. Simple in appearance, his definition constructivizes the property of induction in just the right way. An ordinal number (or well-ordered set) is a set S with a binary relation $<$ such that

(1) if $a < b$ and $b < c$, then $a < c$

(2) one and only one of the relations $a < b$, $b < a$, $a = b$ holds for given elements a and b of S

(3) let T be any subset of S with the property that every element b of S, such that $a \in T$ for each a in S with $a < b$, belongs to T; then $T = S$.

Richman shows that each Brouwerian ordinal satisfies (1), (2), and (3). He gives examples of ordinals (in his sense) that are not Brouwerian. He shows that every subset of an ordinal is an ordinal (under the induced order). He uses his theory to constructivize the classical theorems of Zippin and Ulm concerning existence and uniqueness of p-groups with prescribed invariants.

The above examples might give the impression that the constructivization of classical mathematics always proceeds smoothly. I shall now give some other examples, to show that in fact it does not.

In [1] the Gelfand theory of commutative Banach algebras was constructivized to a certain extent. The theory has to be considered unsatisfactory, not because the classical content is not recovered (it is), but because it is so ugly. It is almost certain that a prettier constructivization will someday be found.

Stolzenberg [28] gives a meticulous analysis of some of the considerations involved in constructivizing a particular classical theory, the open mapping theorem and related material. Again, an incisive constructivization is not obtained.

J. Tennenbaum [29] gives a deep and intricate constructive version of Hilbert's basis theorem. Consider a commutative ring A with unit. It would be tempting to call A (constructively) Noetherian if for each sequence $\{a_n\}$ of elements of A there exists an integer N such that for $n \geq N$ the element a_n is a linear combination of $a_1, \ldots, a_{n-1}$ with coefficients in A. This notion would be worthless--not even the ring of integers is Noetherian in this sense. In case A is discrete (meaning that the equality relation for A is decidable), the appropriate constructive version of Noetherian seems to be the following (as given in [29]).

<u>Definition</u>. A sequence $\{a_n\}$ of elements of A is almost eventually zero if for each sequence $\{n_k\}$ of positive integers there exists a positive integer k such that $a_n = 0$ for $k \leq n \leq k + n_k$.

<u>Definition</u>. A basis operation r for A is a rule that to each finite sequence $a_1, \ldots, a_n$ of elements of A assigns an element $r(a_1, \ldots, a_n)$ of A of the form $a_n + \lambda_1 a_1 + \ldots + \lambda_{n-1} a_{n-1}$, where each λ_i belongs to A.

<u>Definition</u>. A is Notherian if it has a basis operation r such that for each sequence $\{a_n\}$ of elements of A the associated sequence $\{r(a_1, \ldots, a_n)\}$ is almost eventually zero.

Tennenbaum proved the appropriateness of his version of Noetherian by checking out the standard cases and proving the Hilbert basis theorem. He also extended his definition and results to the case of a not-necessarily discrete ring A. The theory in that case is so complex that it cannot be considered satisfactory.

In spite of the pioneering efforts of Kronecker, and continued work by many algebraists, resulting in many deep theorems, the systematic constructivization of algebra would seem hardly to have begun. The problems are formidable. A very tentative suggestion is that we should restrict our attentions to algebraic structures endowed with some sort of topology, with respect to which all operations and maps are continuous. The work of Tennenbaum quoted above might provide some ideas of how to accomplish this. The task is complicated by the circumstance that no completely suitable constructive framework for general topology has yet been found.

The constructivization of general topology is impeded by two obstacles. First, the classical notion of a topological space is not constructively viable. Second, even for metric spaces the classical notion of a continuous function is not constructively viable; the reason is that there is no constructive proof that a (pointwise) continuous function from a compact (complete and totally bounded) metric space to $\mathbb{R}$ is uniformly continuous. Since uniform continuity for functions on a compact space is the useful concept, pointwise continuity (no longer useful for proving uniform continuity) is left with no useful function to perform. Since uniform continuity can not be formulated in the context of a general topological space, the latter concept also is left with no useful function to perform.

In [1] I was able to get along by working mostly with metric spaces and using various ad hoc definitions of continuity: one for compact spaces, another for locally compact spaces, and another for the duals of Banach spaces. The unpublished manuscript [4] was an attempt to develop constructive general topology systematically. The basic idea is that a topological space should consist of a set X, endowed with both a family of metrics and a family of boundedness notions, where a boundedness notion on X is a family S of subsets of X, (called bounded subsets), whose union is X, closed under finite unions and the formation of subsets.

For example, let C be the set of all real valued functions $f : \mathbb{R} \to \mathbb{R}$, bounded and (uniformly) continuous on finite intervals. Each finite interval of $\mathbb{R}$ induces a metric on C (the uniform metric on that interval). In addition, there is a natural boundedness notion S. A subset E of C belongs to S if there exists $r > 0$ such that $|f| \le r$ for all f in E. A sequence $\{f_n\}$ of elements of C converges to an element f of C if it converges with respect to each of the metrics on C, and if it is bounded.

The notion of a continuous function from one such space to another, as given in [4], is somewhat involved and will not be repeated here. It was possible to develop a theory that seems to accomodate the known examples and to have certain pleasing functorial qualities, but the theory is somehow not convincing--for one thing, it is too involved. For another, there is a certain sort of space--let us call it a ball space--that does not fit well into the theory.

Definition. A ball space is a set X, together with a function that to each $r \ge 0$ and point x of X associate a subset $B(x, r)$ of X (to be thought of as the closed ball of radius r about x) satisfying the following axioms.

(a) $B(x, r) \subset B(x, s)$ if $r \le s$

(b) $B(x, 0) = \{x\}$

(c) $B(x, r) = \cap \{B(x, s) : s > r\}$

(d) if $y \in B(x, r)$, then $x \in B(y, r)$

(e) if $y \in B(x, r)$ and $z \in B(y, s)$, then $z \in B(x, r + s)$

(f) $\cup \{B(x, r) : r \ge 0\} = X$.

Duals of Banach spaces are particular instances of ball spaces, as are various other function spaces.

Algebraic topology, at least at the elementary level, should not be too difficult to constructivize. There is a problem with defining singular cohomology constructively, as pointed out in [2]. Richman [25] points out that the classical Vietoris homology theory is not satisfactory constructively, and he gives a new version that constructively (and also classically) has certain features that are more desirable.

I would like to conclude these lectures by discussing some of the tasks that face constructive mathematics.

Of primary importance is the systematic constructive development of enough of algebra for a pattern to begin to emerge. Of course, it may be that much of the classical theory is inherently unconstructivizable, and that constructive algebra will go its own way. It is too early to tell.

Less critical, but also of interest, is the problem of a convincing constructive foundation for general topology, to replace the ad hoc definitions in current use. It would also be good to see a constructivization of algebraic topology actually carried through, although I suspect this would not pose the critical difficulties that seem to be arising in algebra.

To sum up, the first task is to constructivize as much of existing classical mathematics as is suitable for constructivization. As this is being done, we should increasingly turn our attention to questions of the efficiency of our algorithms, and bridge the gap between constructive mathematics on the one hand and numerical analysis and the theory of computation on the other. Since constructive mathematics is the study of what is theoretically computable, it should afford a sound philosophical basis for the theory of computation.

Our terminology and technical devices need constant re-examination as to whether they are the most appropriate tools for extracting the full meaning from our material. It seems to me that the meaning of implication, in particular, should be thoroughly studied, and other possible candidates investigated. Such statements as "$(A \to B) \to C$" have a rather tenuous meaning, and in many instances of proofs of such statements, something more is actually being proved. Work of Gödel [17] raises some interesting possibilities about possible re-definitions of implication, which seem to be very difficult to implement in usable generality, and which also seem to run counter to natural modes of thought. There seems to be no reason in principle that we should not be able to develop a viable terminology that incorporates more than one meaning for some or all of the quantifiers and connectives.

More important than any of these technical problems is the broader problem of involving ourselves more deeply with the meaning of mathematics at all levels. This is the simplest and most general statement of the constructivist program, and the technical developments are intended as a means to that end.

References

1. Bishop, E., "Foundations of Constructive Analysis", McGraw-Hill, New York, 1967.

2. Bishop, E., Mathematics as a Numerical Language, in "Intuitionism and Proof Theory", North-Holland, Amsterdam, 1970.

3. Bishop, E., A Constructive Ergodic Theorem, Jour. Math. Mech., vol. 17 (1968), pp. 631-640.

4. Bishop, E., Notes on Constructive Topology, U.C.S.D., 1971.

5. Bishop, E. and H. Cheng, "Constructive Measure Theory", A.M.S. memoire 116, Providence, 1972.

6. Brouwer, L.E.J., Begründung der Mengenlehre Unabhängig vom Logischen Satz vom Ausgeschlossenen Dritte II: Theorie der Punktmengen, Ver. Kon. Akad. Wet. (Amsterdam) vol. 12 (1919), pp. 1-33.

7. Brouwer, L.E.J., Beweiss dass Jede Volle Funktion Gleichmässig Stetig ist, Kon. Ned. Akad. Wet. (Amsterdam), vol. 27 (1924), pp. 189-193.

8. Brouwer, L.E.J., Zur Begründung der Intuitionistischen Mathematik III, Math. Ann. vol. 96 (1927), pp. 451-488.

9. Chan, Y. K., A Constructive Approach to the Theory of Stochastic Processes", Trans. A.M.S. vol. 165 (1972), pp. 37-44.

10. Chan, Y.K., A Constructive Version of the Renewal Theorem, to appear.

11. Chan, Y.K., Notes on Constructive Probability Theory, to appear.

12. Chan, Y.K., On Constructive Convergence of Measures on the Real Line, to appear.

13. Cheng, H., Ph.D. Dissertation, University of California at San Diego, 1971.

14. Doob, J. L., "Stochastic Process" Wiley, New York, 1953.

15. Frankel, A. and Y. Bar-Hillel, "Foundations of Set Theory", North-Holland, Amsterdam, 1958.

16. Godel, K., What is Cantor's Continuum Problem?", Amer. Math. Monthly, vol. 54 (1947), pp. 515-525.

17. Gödel, K., Über Eine Bisher Noch Nicht Benützte Erweiterung des Finiten Standpunktes, Dialectica, vol. 12 (1958), pp. 280-287.

18. Goodman, N. D. and J. Myhill, The Formalization of Bishop's Constructive Mathematics, in "Topics in Algebraic Geometry, and Logic", Springer, Berlin, 1972.

19. Heyting, A., Note on the Riesz-Fischer Theorem, Proc. Kon. Akad. Wet. (Amsterdam) vol. 54 (1951) pp. 35-40.

20. Kleene, S. C., "Introduction to Metamathematics", Van Nostrand, Princeton, 1952.

21. Kleen, S. C. and R. Vesley, "The Foundations of Intuitionistic Mathematics", North-Holland, Amsterdam, 1965.

22. Myhill, J., Formal Systems of Intuitionistic Analysis III-Lawlike Analysis, to appear.

23. Nuber, J., A Constructive Ergodic Theorem, Trans. A.M.S., vol. 164 (1972), pp. 115-137.

24. Nuber, J., A Constructive Proof of the Chacon-Ornstein Theorem, to appear.

25. Richman, F., A Constructive Modification of Vietoris Homology, to appear.

26. Richman, F., The Constructive Theory of Countable Abelian P-Groups, to appear in Pac. Jour. Math.

27. Spector, C., Provably Recursive Functionals of Analysis, in Amer. Math. Soc. Symposium on pure mathematics, No. 5, Providence, 1962, pp. 1-27.

28. Stolzenberg, G., A Critical Analysis of Banach's Open Mapping Theorem, to appear.

29. Tennenbaum, J., Ph.D. Dissertation, University of California at San Diego, 1973.

30. van der Meulen, S. G. and C. H. Lindsey, "Informal Introduction to Algol 68", Math. Centrum, Amsterdam, 1969.

31. van Rootselaar, B., Generalization of the Brouwer Integral, Ph.D. Dissertation Amsterdam, 1954.

32. Weyl, H., "Das Kontinuum", Leipzig, 1918. Reprinted by Chelsea, New York, 1932.