Errett Bishop. "Mathematics as a numerical language." In: Intuitionism and Proof Theory, Proc. Conf., Buffalo, N.Y., 1968, North-Holland, Amsterdam, 1970, pp. 53–71.

MATHEMATICS AS A NUMERICAL LANGUAGE

ERRETT BISHOP

The one point on which constructivists agree is their criticism of classical mathematics. Brouwer's great contribution was to analyze intensively the inadequacies of that system. After Brouwer and Kronecker, we all know a reformation is needed, but disagree about what course it should take. Intuitionism, as developed by Brouwer, stresses as basic our intuition of the integers and our intuition of the real numbers; all of mathematics is to be reduced to these two primitive constructs. In my book [1] I proposed, in the spirit of Kronecker rather than Brouwer, that the integers are the only irreducible mathematical construct. This is not an arbitrary restriction, but follows from the basic constructivist goal – that mathematics concern itself with the precise description of finitely performable abstract operations. It is an empirical fact that all such operations reduce to operations with the integers. There is no reason mathematics should not concern itself with finitely performable abstract operations of other kinds, in the event that such are ever discovered; our insistence on the primacy of the integers is not absolute.

Thus by 'constructive' I shall mean a mathematics that describes or predicts the results of certain finitely performable, albeit hypothetical, computations within the set of integers. If a word is needed to delimit this special variety of constructivism, I propose the term 'predictive'. From the predictive point of view, Brouwer's intuitionism at first glance contains elements that are extremely dubious; free choice sequences and allied concepts admit no ready numerical interpretation. Moreover, the numerical content of intuitionistic mathematics is diluted by over-reliance on negativistic techniques. The role of negation in predictive mathematics is philosophically secure, if only because there exist negative statements that do have numerical content. Nevertheless, it is remarkable that a systematic effort to avoid negation leads uniformly to better results.

Constructive mathematics is in its infancy. According to some, it is doomed to the role of scavenger. These people conceive of classical mathematics as establishing the grand design and the imaginative insight, leaving the constructivists to add whatever embellishments their credos demand. Although totally wrong, this viewpoint hints at a truth: The most urgent task of the constructivist is to give predictive embodiment to the ideas and techniques of classical mathematics. Classical mathematics is not totally divorced from reality. On the contrary, most of it has a strongly constructive cast. Much of the constructivization of classical mathematics is therefore routine; constructive versions of many standard results are readily at hand. This makes it easy to miss the point, which is *not* to find a constructive version of this or that, or even of every, classical result. The point is not even to find elegant substitutes for whole classical theories. The point rather is to use classical mathematics, at least initially, as a guide. Much will be of little value to the constructivist, much will be constructive per se, and much will raise fundamental questions which classically are trivial or perhaps do not even make sense. The emphasis will be on the discovery of useful and incisive numerical information. It is the incisiveness and scope of the information, not the elegance of the format, that is relevant.

A given classical result may have no, one, or many constructive versions, none necessarily superior to the rest, because different constructive theorems can represent different numerical aspects of the same classical result, all giving different estimates, useful for different ends. In many instances a classical definition, which makes good constructive sense, no longer represents the correct point of view, and so must be replaced. Finding the correct replacement is often a non-trivial challenge, involving considerations which from the classical standpoint would be absurd. I believe that eventually the influence of constructive on classical mathematics will be greater than the influence the other way. Very possibly classical mathematics will cease to exist as an independent discipline. In the meantime, it behooves constructivists to attach their mathematical and philosophical investigations to mathematics as it exists. Contrary practice has led to numerous irrelevancies and misplaced emphasis. Even a quick look shows that much of the constructivist literature lacks the serious intent the subject demands.

In this short paper I want to indicate what to me are some of the important questions of constructive mathematics today. First, I wish to discuss certain mathematical problems, not because of any special interest or difficulty of these particular problems, but as an indication of the sort of work that needs to be done.

The first example is taken from probability theory. (Modern probabilists seem to have little or no interest in the computation or even the computability of the probabilities with which they deal. It is not surprising the subject is blatantly idealistic.) The Birkhoff ergodic theorem asserts that if T is a measure-preserving transformation on a finite measure space and f is an integrable function, then the averages

$$f^n(x) \equiv \frac{1}{n+1}(f(x)+f(Tx)+ \ldots +f(T^n x))$$

converge as $n \to \infty$ for almost all x. Constructively this result fails. For a counterexample in the style of Brouwer, take T to be rotation of the unit circle through an angle $2\pi\alpha$, where α is a real number which, for all we know, could be equal to 0, but could, on the other hand, be irrational. On the one hand $\{f^n\}$ would converge a.e. to f, and on the other to the constant function $(2\pi)^{-1}\int f(\theta)d\theta$. Thus a constructive proof of Birkhoff's theorem is out of the question. It is shown in [**1**] that the sequence $\{f^n\}$ in general satisfies certain inequalities (of a type first introduced by Doob for the study of martingales, and called *upcrossing inequalities*), which classically imply the sequence converges a.e. These inequalities, which from the classical point of view constitute a considerable strengthening of Birkhoff's theorem and its principal modern generalizations, would seem to afford a satisfactory constructive version of the ergodic theorem, but this is not so. In the case of a completely general measure-preserving transformation, the upcrossing inequalities are probably satisfactory. In other words, they afford a good *equal-hypothesis* substitute for Birkhoff's theorem. However, we would also like a good *equal-conclusion* substitute for Birkhoff's theorem – that is, usable conditions on T that imply the constructive convergence of the sequence $\{f^n\}$ almost everywhere. This is an important open problem.

The next example is taken from algebra. Recently I was asked whether the Hilbert basis theorem – that a polynomial ring over a Noetherian domain is Noetherian – is constructively valid. The answer is easily seen to be 'yes'. Unfortunately, not even the ring of integers is Noetherian from the constructive point of view (and therefore the Hilbert basis theorem is vacuous). For a counterexample in the style of Brouwer, let $\{n_k\}$ be a sequence of integers, for which we are in doubt as to whether they are all equal to 0. The ideal generated by the integers n_k has no finite basis in the constructive sense. The problem is to find a constructively usable reformulation of the definition of a Noetherian ring, which would include the integers and give constructive substance to the Hilbert basis theorem.

Our third example, from topology, came as a surprise. Elementary algebraic topology should be constructive, but the definition of the singular cohomology groups gives trouble. A singular 1-simplex of the unit circle S^1 is a continuous function ω from the closed unit interval $[0, 1]$ into S^1. Let Ω denote the set of all such ω. A singular 1-cochain c (over the integers $\boldsymbol{Z}$) can be thought of as a function from Ω to $\boldsymbol{Z}$. We wish to define constructively a singular 1-cochain that generates the one-dimensional cohomology group of S^1. Such a cochain will in particular be a non-constant integer-valued function c on Ω. Now the set Ω is arcwise connected, in the sense that if ω_1 and ω_2 are any points of Ω, there exists $\lambda\colon [0, 1] \to \Omega$ with $\lambda(0) = \omega_1$ and $\lambda(1) = \omega_2$. A result of Brouwer says that every integer-valued function on $[0, 1]$ is constant. A corollary is that every integer-valued function on an arcwise connected set is constant. Thus a non-constant integer-valued function c on Ω would counterexample Brouwer's result. Brouwer's result has not been counterexampled. These considerations indicate the difficulties involved in finding a satisfactory constructive version of singular cohomology theory.

Each of the three problems just discussed requires the development of new concepts appropriate to the constructive point of view. None of them is likely to be given an acceptable solution by the application of a general technique of constructivization. Incisive estimates and apt definitions are not to be expected as consequences of general schemes that translate from classical to constructive mathematics, although translation techniques may have value in special instances, as we shall see later.

The most urgent foundational problem of constructive mathematics concerns the numerical meaning of implication. Constructivists have customarily accepted Brouwer's definitions of the mathematical connectives and quantifiers, implication in particular. According to Brouwer, $P \to Q$ means that the existence of a proof of P necessarily entails the existence of a proof of Q, in other words, there is a method that converts a proof of P into a proof of Q. In **[1]** I gave a variant of this definition, fitted to the predictive point of view: $P \to Q$ means '... the validity of the computational facts implicit in the statement P must insure the validity of the computational facts implicit in the statement Q ...'. There is a discrepancy between even this reformulated definition and the predictive goal, since, as defined, $P \to Q$ is not a priori predictive of the results of certain finitely performable computations within the set of integers. Rather than prematurely attempt to resolve this discrepancy, in [**1**] I decided to let the mathematics be the test, and found that in actual practice there was little difficulty in giving numerical interpretations to statements with implications or even nested implications. Although the numerical meaning of implication is a priori unclear, in each particular instance the meaning is clear. We are at liberty to continue to treat the numerical meaning of implication as being provided by the context, but hopefully there exists a philosophical explanation of the empirical fact that intuitionistic implication in each instance admits a numerical interpretation. Such an explanation requires a deeper analysis of the content of a theorem of constructive mathematics. As a point of departure for such an analysis, I examined a number of theorems and proofs of [**1**], and came to the following conclusions.

A *complete* mathematical statement – that is, a theorem conjoined with its proof and with all theorems, proofs, and definitions on which it depends, either directly or indirectly – asserts that a given constructively defined function f, from a given constructively defined set S to the integers, vanishes identically. In other words, it asserts $\forall x A(x)$, where A is the decidable predicate $f(x) = 0$ and x ranges over S.

Most theorems, standing by themselves, are incomplete mathematical statements. An *incomplete mathematical statement* concerns certain entities whose constructions are not described in the statement itself. For instance, the prime number theorem

$$\lim_{n \to \infty} \frac{\pi(n) \log n}{n} = 1$$

implicitly refers to a sequence $\{n_k\}$ of positive integers such that

$$\left| \frac{\pi(n) \log n}{n} - 1 \right| \leqslant k^{-1},$$

whenever $n \geqslant n_k$. The integers n_k can be extracted from the proof of the prime number theorem, presuming it is constructive (as it is). In general, an incomplete mathematical statement asserts the existence of an element y of a set T, such that when the rule for constructing y is given the statement becomes complete. In other words an incomplete statement has the form $\exists y P(y)$, where $P(y)$ is a complete statement with parameter y. According to the above, the incomplete statement therefore has the form

$$\exists y \forall x A(x, y) \tag{1}$$

where x runs over a set S, y runs over a set T, and $A(x, y)$ is a decidable predicate for each x in S and y in T.

An incomplete mathematical statement is a consequence of its completion. Nevertheless, it contains additional information – the intent. If the prime number theorem were written in the form

$$\forall n \forall k \left(n \geqslant n_k \rightarrow \left| \frac{\pi(n) \log n}{n} - 1 \right| \leqslant k^{-1} \right),$$

and n_k replaced by its explicit definition, we might fail to realize that the primary intent was to construct any sequence $\{n_k\}$ making the statement true. The use of incomplete statements aids the intuition and saves time. Moreover, incomplete statements are structural components of implications. An implication $P \rightarrow Q$, where P and Q are incomplete statements, means something quite different from $P' \rightarrow Q'$, where P' is a completion of P and Q' of Q.

Let us grant that mathematics is concerned with statements of the form (1), and inquire how an implication $P \rightarrow Q$, where P and Q are statements of form (1), is to be interpreted as a statement of form (1). The interpretation we shall develop is due to Gödel **[2]**. Let P be $\exists y \forall x A(x, y)$ and Q be $\exists v \forall u B(u, v)$, so that $P \rightarrow Q$ is

$$\exists y \forall x A(x, y) \rightarrow \exists v \forall u B(u, v). \tag{2}$$

Pure thought translates this into

$$\forall y (\forall x A(x, y) \rightarrow \exists v \forall u B(u, v)). \tag{3}$$

For a given value of y, which we now fix, (3) asserts that in case the statement $A(x, y)$ is true for all x then v can be constructed to have certain properties. The rule for constructing v will consist of a certain finitely describable procedure, some stages of which perhaps assume the truth of $\forall x A(x, y)$. For example, at a certain stage we may need to know that a certain integer d is non-zero, in order to be able to perform a division, and it may be necessary to make use of the hypothesis $\forall x A(x, y)$ to derive the needed inequality $d \neq 0$. In fact, such an occurrence is no obstacle to giving a universally valid definition of v. We simply take v to be some convenient constant in case $d = 0$ (*which* constant does not matter). Such considerations lead us to expect we can extend out original construction, and give a universal construction for an element v of the set in question, with the property that in case $\forall x A(x, y)$ does hold, then v will have the value originally prescribed. In other words, we conjecture that if we can prove the statement

$$\forall x\, A(x, y) \to \exists v \forall u\, B(u, v), \tag{4}$$

we can actually prove the seemingly stronger statement

$$\exists v(\forall x\, A(x, y) \to \forall u\, B(u, v)), \tag{5}$$

which in turn is equivalent to

$$\exists v \forall u(\forall x\, A(x, y) \to B(u, v)). \tag{6}$$

Thus we conjecture that $P \to Q$ is actually equivalent to the seemingly stronger statement

$$\forall y \exists v \forall u(\forall x\, A(x, y) \to B(u, v)). \tag{7}$$

To go deeper, fix values of y, v and u, and consider the statement

$$\forall x\, A(x, y) \to B(u, v) \tag{8}$$

which asserts that the particular decidable statement $B(u, v)$ is a necessary consequence of the totality of decidable statements $A(x, y)$ for all x. Experience indicates that a proof of a statement such as (8) actually deduces the truth of $B(u, v)$ from the truth of finitely many of the statements $A(x, y)$. Moreover, reflection indicates the difficulty of exhibiting an instance of a proof of a statement such as (8) that actually uses infinitely many of the statements $A(x, y)$ to prove the statement $B(u, v)$. Let us therefore conjecture that if we can prove (8) we can actually construct elements $x_1, \ldots, x_n$ such that

$$A(x_1, y) \wedge \ldots \wedge A(x_n, y) \to B(u, v). \tag{9}$$

Now (9) is a finitary statement involving decidable propositions. Therefore the rules of classical logic hold, so that there exists k with $1 \leqslant k \leqslant n$ such that

$$A(x_k, y) \to B(u, v). \tag{10}$$

Thus we conjecture that if we can prove (8) we can prove

$$\exists x(A(x, y) \to B(u, v)). \tag{11}$$

Thus we conjecture that (7) is actually equivalent to the seemingly stronger statement

$$\forall y \exists v \forall u \exists x(A(x, y) \to B(u, v)). \tag{12}$$

Using the axiom of choice, we transform (12) into

$$\exists \bar{v} \exists \bar{x} \forall y \forall u (A(\bar{x}(y, u), y) \to B(u, \bar{v}(y))). \tag{13}$$

This is our candidate for a numerical version of $P \to Q$, and these considerations probably explain why intuitionistic implication in actual practice admits of numerical interpretation. Gödel in **[2]** proves for a particular formal system, designed to accommodate large portions of constructive mathematics, that any proof of (2) can be transformed into a proof of (13). Whether or not we wish to commit ourselves to this or any other formal system, Gödel's result strengthens our conjecture that (13) is the proper numerical version of $P \to Q$, and leads us to define a new type of implication, which I shall call *numerical implication* (or *Gödel implication*). With P and Q as above, $P \to Q$ is defined to be the statement (13). Since it appears that intuitionistic implication in practice amounts to numerical implication, it would seem that for the philosophical unity of predictive mathematics we should abandon intuitionistic implication and work with numerical implication exclusively. Before we definitely accept such a change, we should check in more detail that numerical implication is actually being used, and, just as important, that our intuitions can adjust to the change. I believe experience will prove that numerical implication is at least as natural and easy to use as intuitionistic implication.

Another important foundational problem is to find a formal system that will efficiently express existing predictive mathematics. I think we should keep the formalism as primitive as possible, starting with a minimal system and enlarging it only if the enlargement serves a genuine mathematical need. In this way the formalism and the mathematics will hopefully interact to the advantage of both. As a point of departure, we take a formal system such as used by Gödel [**2**]. Another version is given in Spector [**6**], where the relevant system is called Σ_2. It is closely related to the system of Kleene and Vesley [**3**], with the free-choice type axioms left out.

To give a quick sketch, our system Σ formalizes the theory of functions of certain types. The types are defined inductively as follows. The primitive type is [**0**], and a function of type [**0**] is an integer. If $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n$ are types, $\boldsymbol{t} \equiv (\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n)$ is a type, and a function of type $\boldsymbol{t}$ is an n-tuple $(f_1, \ldots, f_n)$, where f_i is of type $\boldsymbol{t}_i$. If $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ are types, $\boldsymbol{t}_1 \to \boldsymbol{t}_2$ is a type, and a function of type $\boldsymbol{t}_1 \to \boldsymbol{t}_2$ is a function from the set of all functions of type $\boldsymbol{t}_1$ into the set of all functions of type $\boldsymbol{t}_2$. The types $\boldsymbol{t}_1 \to (\boldsymbol{t}_2 \to \boldsymbol{t}_3)$ and $(\boldsymbol{t}_1, \boldsymbol{t}_2) \to \boldsymbol{t}_3$ are considered to be the same, and the types $\boldsymbol{t} \to (\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n)$ and

$(\boldsymbol{t} \to \boldsymbol{t}_1, \ldots, \boldsymbol{t} \to \boldsymbol{t}_n)$ are considered to be the same. The system Σ contains variables of the various types, for functions of the various types. Variables of the various types are combined in meaningful ways to form terms. A variable of type $\boldsymbol{t}$ is also a term of type $\boldsymbol{t}$. If $\mathfrak{u}$ is a term of type $\boldsymbol{t}_1 \to \boldsymbol{t}_2$ and $\mathfrak{v}$ a term of type $\boldsymbol{t}_1$, then $\mathfrak{u}(\mathfrak{v})$ is a term of type $\boldsymbol{t}_2$. The constant 0 is a term of type $[\mathbf{0}]$. For each term $\mathfrak{u}$ of type $[\mathbf{0}]$, the *successor* $\mathfrak{u}'$ of $\mathfrak{u}$ is a term of type $[\mathbf{0}]$. If $\mathfrak{u}$ is a term of type $\boldsymbol{t}$ and $x_1, \ldots, x_n$ are distinct variables of arbitrary types $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n$, then $\lambda(x_1, \ldots, x_n)\mathfrak{u}$ is a term of type $(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n) \to \boldsymbol{t}$, to be interpreted as the function whose value at $(f_1, \ldots, f_n)$ is the result of replacing x_i by f_i in $\mathfrak{u}$ $(1 \leqslant i \leqslant n)$. The primitive formulas of Σ are of the form $\mathfrak{u} = \mathfrak{v}$, where $\mathfrak{u}$ and $\mathfrak{v}$ are terms of type $[\mathbf{0}]$. By means of the connectives $\wedge$, $\vee$ and $\to$, and the quantifiers $\exists$ and $\forall$, applied to primitive formulas, arbitrary formulas are obtained. The statement 'not A' is defined to mean '$A \to 0 = 1$.' The axioms and rules of inference include the axioms and rules of the intuitionistic predicate calculus (rules and axioms Al through A10 and Bl through B4 of **[6]**), axioms for equality (axioms Cl through C4 of **[6]**), the induction rule (rule D of **[6]**), and the axiom of choice for all types (an extension of axiom E of **[6]**). Also included is an axiom expressing the meaning of the λ-operator. For convenience we might also include axioms for certain functions, as is done in **[3]**.

Our first problem is to interpret sets in the system Σ. To each set A we associate a formula $A'(x, y)$ containing no quantifiers. (Then $A'(x, y)$ is decidable for given values of the free variables x and y.) The set A is defined by taking $x \in A$ to mean $\forall y A'(x, y)$. (Of course, x may stand for a finite sequence $(x_1, \ldots, x_n)$ of variables of various types, and the same is true of y. Note that it would be incorrect to define $x \in A$ by a formula of the type $\exists z\ \forall y A'(x, y, z)$; in such a situation, the value of z is necessary to completely determine an element of A; hence we should write $(x, z) \in A$ rather than $x \in A$, and define it by the formula $\forall y A'(x, y, z)$.) Each set A has a relation of equality, which means we must define $x_1 =_A x_2$ by a formula

$$\exists z \forall y A''(x_1, x_2, y, z).$$

Of course $=_A$ must be shown to be an equivalence relation. The special equality $x_1 = x_2$ defined as $\forall y(x_1(y) = x_2(y))$ is called *functional equality*. Whenever the equality relation on a set A is not defined, functional equality will be meant. In general we require the equality relation on any set to be weaker than functional equality. If A' and A'' contain a variable u, in addition to those described above, we have a family of sets indexed by the parameter u. This construction will only be used in the special case of a family of subsets, as described below. If A and B are any sets, defined respectively by formulas A', A'' and B', B'', we define the set $C \equiv F(A, B)$ of all functions from A to B as follows. Write the formula

$$\forall x(x \in A \to f(x) \in B) \wedge \forall x_1 \forall x_2(x_1 \in A \wedge x_2 \in A \wedge x_1 =_A x_2 \to f(x_1) =_B f(x_2))$$

in the form $\exists u \forall v C'(f, u, v)$. Take $(f, u) \in C$ to mean $\forall v C'(f, u, v)$. Similarly take $(f_1, u_1) =_C (f_2, u_2)$ to mean

$$\forall x(x \in A \to f_1(x) =_B f_2(x)).$$

To define a subset B of a set A, according to **[1]**, we must define an element x of A, perform the construction of an additional function u, and check that certain additional conditions are satisfied. Thus B will be determined by a certain formula B', and $(x, u) \in B$ will mean

$$x \in A \wedge \forall v B'(x, u, v).$$

We take $(x_1, u_1) =_B (x_2, u_2)$ to mean $x_1 =_A x_2$. In case B' contains a variable w in addition to x, u and v, we obtain a family of subsets of A. The union and the intersection of a family of subsets of A are defined in obvious ways. In case B_1 and B_2 are subsets of A, the formula $B_1 \subset B_2$ is defined in an obvious way, and $B_1 = B_2$ is by definition the formula $B_1 \subset B_2 \wedge B_2 \subset B_1$. The statement $B_1 \subset B_2 \wedge B_2 \subset B_3 \to B_1 \subset B_3$, for instance, is a formula in Σ containing as subformulas the formulas A', A'', B_1', B_2' and B_3'. (This particular statement is of course provable in Σ.)

There is no difficulty in extending the above ideas to complemented sets. (A complemented set, relative to a family $\mathscr{F}$ of real-valued functions on a set A, is an ordered pair (B_1, B_2) of subsets of A such that for all $x \in B_1$ and $y \in B_2$ there exists f in $\mathscr{F}$ with $f(x) \neq f(y)$.) We first run into difficulty with Borel sets. In **[1]** we consider a set A, a family $\mathscr{F}$ of real-valued functions on A, and a family $\mathscr{M}$ of complemented subsets of A relative to $\mathscr{F}$; we define a Borel set generated by $\mathscr{M}$ to be a complemented subset obtainable inductively from the two following techniques of construction:

(1). The elements of $\mathscr{M}$ are Borel sets.

(2). A countable union (or countable intersection) of Borel sets already constructed is a Borel set.

Properties valid for an arbitrary Borel set are often proved by induction, corresponding to the inductive character of the definition just given. The type of induction in question also occurs in Brouwer's definition of the constructive ordinals. These definitions seem not to be formalizable in Σ. To extend the system Σ to subsume the theory of Borel sets, it would be necessary to include the type of induction in question. W. A. Howard tells me he has constructed such an extension. We shall consider this matter no further, since there is another approach, which will be explained later.

It appears that the theory of the standard abstract structures – groups, metric spaces, differentiable manifolds, and so forth, can be developed within Σ. For example, a metric space (X, ρ) can be realized in Σ by the formulas X', X'' determining the set X, by the element (ρ, v) of the set $F(X\times X, \boldsymbol{R}^{0+})$, and by an element w that arises when the formula

$$x \in X \wedge y \in X \wedge \rho(x, y) = 0 \rightarrow x =_X y$$

is translated to the form $\exists w \forall t P(w, t, \rho) = 0$. We shall abbreviate $M \equiv (X', X'', \rho, v, w)$, and introduce the notation $M \in \mathscr{MET}$ to represent the fact that X' and X'' are formulas defining a set X, that ρ is a metric on X, that v and w are the moduli introduced above, and that a certain formula $\forall y \mathscr{M}'(\rho, v, w, y)$, which contains X' and X'' as subformulas and expresses the fact that ρ is a metric on X and v and w are the moduli described, is valid. As a formula of Σ, $M \in \mathscr{MET}$ stands for $\forall y \mathscr{M}'(\rho, v, w, y)$. We might call $\mathscr{MET}$ a *large set* (or a class), as distinguished from the sets already defined (which we sometimes call *small sets*).

Although the metric spaces form a large set, rather than a small set, the compact (metric) spaces can be regarded as a small set, as follows. Let T be the set of totally bounded pseudo-metrics on the integers $\boldsymbol{Z}$. For each ρ in T, let $\boldsymbol{Z}_\rho$ be the completion of $\boldsymbol{Z}$ with respect to ρ. The set of all such completions is the set of all compact spaces.

In the same way, there is a small set of locally compact spaces.

A (small) category C is an analog of a small set. The set X of objects of C is specified by a formula X', and the equality relation on X is taken to be functional equality. Another formula Y' defines the sets $\mathrm{Hom}(x_1, x_2)$ of morphisms, so that $z \in \mathrm{Hom}(x_1, x_2)$ means $\forall y Y'(x_1, x_2, z, y)$. A third formula gives the equality relations on the sets $\mathrm{Hom}(x_1, x_2)$. There is a function e which to each x in X assigns an element (the *identity element*) e_x of $\mathrm{Hom}(x, x)$. There is a function which to each $z_1 \in \mathrm{Hom}(x_1, x_2)$ and $z_2 \in \mathrm{Hom}(x_2, x_3)$ assigns an element (their product) of $\mathrm{Hom}(x_1, x_3)$. The usual axioms for a category must be satisfied.

Of course, all metric spaces, or all groups, or all vector spaces over $\boldsymbol{R}$ cannot be regarded as a small category. However, if a suitable cardinality restriction – such as separability or countability – is imposed on the individual objects, many classical categories can be considered as small categories in our sense. Examples are the compact (metric) spaces, the locally compact (metric) spaces, the countable metric spaces, the (separable) Banach spaces, the countable groups, the locally compact (metric) groups, the (metric) differentiable manifolds, and the countably generated vector spaces over $\boldsymbol{R}$. Functors are easily treated, and a suitable framework for homological algebra is thereby provided. For example, one can formalize the definition of the singular homology functor from the small category of locally compact spaces to a certain small category of abelian groups, and presumably derive in Σ the standard properties of singular homology.

Following the procedure we used for defining a large set, we can define a large category. For example, to make $\mathscr{M}\mathscr{E}\mathscr{T}$ into a large category, we must first define the set Hom (M_1, M_2) of morphisms connecting given objects $M_1 \in \mathscr{M}\mathscr{E}\mathscr{T}$ and $M_2 \in \mathscr{M}\mathscr{E}\mathscr{T}$ of $\mathscr{M}\mathscr{E}\mathscr{T}$. Again the membership relation $z \in$ Hom (M_1, M_2) is defined by a formula $\forall y \mathscr{M}''(M_1, M_2, y, z)$, where our notation indicates that $\mathscr{M}''$ is a formula with variables ρ_1, v_1, w_1, ρ_2, v_2, w_2, y, z containing X_1', X_1'', X_2' and X_2'' as subformulas. Another formula will give the equality relation on Hom (M_1, M_2). Again the function e which assigns to each $M \in \mathscr{M}\mathscr{E}\mathscr{T}$ an identity element of Hom (M, M) and the function describing the product of morphisms must be given. The structure of these functions needs further elucidation. They will not be fixed functions of the system Σ, because their types will depend on the types of their arguments, which are not fixed. Presumably their definitions will have the same form independently of type.

Continuing along the above lines, we should have no difficulty in defining functors between large categories. Whether the theory would be useful in constructive mathematics is not clear. On the one hand, it is possible that small categories are adequate for the applications, but it is also possible that something more general than a large category might be needed, to define which we would need to enlarge the system Σ.

Consider a positive measure μ on a compact space X, that is, a nonnegative linear functional $f \to \int f \mathrm{d}\mu$ on the space $C(X)$ of all continuous real-valued functions on X. We wish to define the measures of certain subsets of X. It would be extremely awkward to attempt to formalize the theory as given in [**1**] in the system Σ, because there is no set of Borel sets and therefore no set of measurable sets in Σ. Thus we must either extend Σ or take another approach. It turns out that a modification of the approach of **[1]** gives a theory that is not only formalizable in Σ but improves the original version. The idea is to define a *partial set* S to be a triple $(f, \{f_j\}, M)$, where $f \in C(X)$, $0 \leqslant f \leqslant 1$, $f_j \in C(X)$ and $0 \leqslant f_j$ for each positive integer j, and M is a positive integer, such that

(a). $\omega(S) \equiv \lim\limits_{j\to\infty} \int f_j \mathrm{d}\mu$ exists,

(b). $f_M > \min\{f, 1-f\}$,

(c). $f_1 \leqslant f_2 \leqslant \ldots$.

Then we define $x \in S$ to mean that there exists $\delta > 0$ with $f(x) > f_m(x)+\delta$ for all positive integers m, and $x \in \sim S$ to mean there exists $\delta > 0$ with $1-f(x) > f_m(x)+\delta$ for all m. Consider $x \in S$ and $y \in \sim S$. Now either $f_M(x) \geqslant f(x)$ or $f_M(x) \geqslant 1-f(x)$, by (b). Since the former inequality contradicts $x \in S$, we have in fact $f_M(x) \geqslant 1-f(x)$. Hence $1-f(x) < f(x)$. Similarly $f(y) < 1-f(y)$. Hence $x \neq y$. Write $S' \equiv \{x : x \in \sim S\}$ and $S'' \equiv \{x : x \in \sim S\}$. The ordered pair

$$\alpha(S) \equiv (S', S'')$$

is a complemented set in the sense of **[1]**, which means that if $x \in S'$ and $y \in S''$ then $x \neq y$. The *complement* of S is defined to be the triple $\sim S \equiv (1-f, \{f_j\}, M)$. Clearly $\alpha(\sim S) = ((\sim S)', (\sim S)'') = (S'', S')$, which by definition (see **[1]**) is $-(S', S'') = -\alpha(S)$.

Now if $T \equiv (g, \{g_j\}, N)$ is a second partial set, we define $S < T$ to mean that for each positive integer m there exists a positive integer $\gamma \equiv \gamma(m)$ such that $|f-g| \leqslant f_\gamma - g_m$. Clearly, $S < T$ if and only if $\sim S < \sim T$. Also $S < U$ whenever $S < T$ and $T < U$, and $x \in T$ whenever $x \in S$ and $S < T$. Hence $S < T$ implies that $\alpha(S) < \alpha(T)$, in the sense that $x \in \alpha(S) \to x \in \alpha(T)$ and $x \in -\alpha(S) \to x \in -\alpha(T)$.

For the partial set S considered above, write $\int S \equiv \int f \mathrm{d}\mu$. For the partial sets S and T given above, $S < T$ implies

$$\left|\int S - \int T\right| = \left|\int f \mathrm{d}\mu - \int g \,\mathrm{d}\mu\right| \leqslant \int |f-g| \mathrm{d}\mu \leqslant$$
$$\leqslant \int (f_\gamma - g_m) \mathrm{d}\mu \leqslant \int f_\gamma \mathrm{d}\mu \leqslant \omega(S).$$

The *union* $S \vee T$ of the partial sets S and T is

$$S \vee T \equiv (\max\{f, g\}, \{f_j + g_j\}_{j=1}^{\infty}, M+N).$$

It is easily seen that $S \vee T$ is a partial set, with $\omega(S \vee T) = \omega(S)+\omega(T)$. Similarly, $S \wedge T$ is defined to be

$$(\min\{f, g\}, \{f_j+g_j\}_{j=1}^{\infty}, M+N).$$

It is a partial set. We have

$$\sim(S \vee T) = \sim S \wedge \sim T \quad \text{and} \quad \sim(S \wedge T) = \sim S \vee \sim T.$$

If $S_1 < S_2$ and $T_1 < T_2$, then

$$S_1 \vee S_2 < T_1 \vee T_2 \quad \text{and} \quad S_1 \wedge S_2 < T_1 \wedge T_2 .$$

Also

$$\alpha(S \vee T) < \alpha(S) \cup \alpha(T) \quad \text{and} \quad \alpha(S \wedge T) < \alpha(S) \pm \alpha(T).$$

A *measurable set* $S \equiv \{S(n)\}_{n=1}^{\infty}$ is a sequence $S(1) < S(2) < \ldots$ of partial sets, with

$$S(n) \equiv (f(n, \cdot), \{f_j(n, \cdot)\}_{j=1}^{\infty}, M(n)),$$

such that $\lim_{n\to\infty} \omega(S(n)) = 0$. The limit

$$\mu(S) \equiv \lim_{n\to\infty} \int f(n, \cdot)\mathrm{d}\mu,$$

called the *measure* of S, exists. If T is another measurable set, we define $(S \vee T)(n) \equiv S(n) \vee T(n)$ and $(S \wedge T)(n) \equiv S(n) \wedge T(n)$ for each n. Also $(\sim S)(n) \equiv \sim S(n)$. It follows that $S \vee T$, $S \wedge T$ and $\sim S$ are measurable sets, and the usual algebraic laws are valid.

For each measurable set S, we take $x \in S$ to mean $x \in S(n)$ for some n, and $x \in \sim S$ to mean $x \in \sim S(n)$ for all n. Correspondingly we write $\alpha(S) \equiv \bigcup_{n=1}^{\infty} \alpha(S(n))$. Then $\alpha(\sim S) = -\alpha(S)$.

If the measure $\mu(S)$ of the measurable set S is positive, then $\int f(n, \cdot)\mathrm{d}\mu >$ $> \omega(S_n)$ for some n, and a construction similar to that of **[1]** gives a point $x \in S(n)$. Hence $x \in S$, so that $\alpha(S)$ is non-void.

If $\{S(\cdot, k)\}_{k=1}^{\infty}$ is a sequence of measurable sets, such that

$$C \equiv \lim_{n\to\infty} \mu(S(\cdot, 1) \vee \ldots \vee S(\cdot, k))$$

exists, a somewhat complicated definition basically similar to that given in **[1]**, leads to a measurable set $S \equiv \bigvee_{k=1}^{\infty} S(\cdot, k)$ having the following properties:

(1). $\mu(S) = C$

(2). $\alpha(S) < \bigcup_{k=1}^{\infty} \alpha(S(\cdot, k))$.

An operation $\bigwedge$ is defined similarly. The theory of chapter 6 of [1] can be developed in this framework.

To formalize in Σ the notion of an abstract measure space, definition 1 of chapter 7 of [1] must be rewritten as follows. A *measure space* is a family $\mathscr{M} \equiv \{A_t\}_{t \in T}$ of complemented subsets of a set X relative to a certain family $\mathscr{F}$ of real-valued functions on X, a map $\mu: T \to \boldsymbol{R}^{0+}$, and an additional structure as follows: The void set $\emptyset$ is an element A_{t_0} of $\mathscr{M}$, and $\mu(t_0) = 0$. If s and t are in T, there exists an element $s \vee t$ of T such that $A_{s \vee t} < A_s \cup A_t$. Similarly, there exist operations $\wedge$ and $\sim$ on T, corresponding to the set-theoretic operations $\cap$ and $-$. The usual algebraic axioms are assumed, such as $\sim (s \vee t) = \sim s \wedge \sim t$. Certain measure-theoretic axioms, such as $\mu(s \vee t) + \mu(s \wedge t) = \mu(s) + \mu(t)$, are also assumed. Finally, there exist operations $\bigvee$ and $\bigwedge$. If, for example, $\{t_n\}$ is a sequence such that $C \equiv \lim_{k \to \infty} \mu(t_1 \vee \ldots \vee t_k)$ exists, then $\bigvee \{t_n\}$ is an element of T with measure C. Certain axioms for $\bigvee$ and $\bigwedge$ are assumed. If T is the family of measurable sets of a compact space relative to a measure μ, and the set-theoretic function $\mu: T \to \boldsymbol{R}^{0+}$ and the associated operations are defined as indicated above, the result is a measure space in the sense just described.

Considerations such as the above indicate that essentially all of the material of [1], appropriately modified, can be comfortably formalized in Σ.

Much effort has been expanded in developing formal systems to accommodate Brouwer's theory of free choice sequences and related constructs. I am of the opinion that this is not the appropriate approach to Brouwer's ideas. Presuming we are satisfied with Σ as a vehicle for predictive mathematics, I think we should realize the non-predictive portions of intuitionistic mathematics as part of the metatheory of Σ, rather than trying to incorporate them in some modification of Σ. (Aspects of Brouwer's approach to free choice sequences lead me to think he might have been sympathetic to this point of view.) Now the metatheory of Σ is based on the work of Gödel [2], who shows that every function proved in Σ to exist can be constructed (simply by unwinding the existence proof) by means of certain canonical operations. Kreisel [5] has shown that every function f, of the type mapping sequences of integers into integers, constructed by means of Gödel's canonical operations, is continuous, in the sense that if ω_0 is any sequence of positive integers there exists a positive integer N such that if ω and ω' are any sequences of integers with $|\omega| \leqslant \omega_0$ and $\omega(n) = \omega'(n)$ for all $n \leqslant N$ then $f(\omega) = f(\omega')$. This can be regarded as the central result of

Brouwer's theory of spreads. Thus we can develop Brouwer's ideas as a metatheory. In particular, Brouwer's result that every $f: \boldsymbol{R} \to \boldsymbol{R}$ is continuous emerges as a meta-theorem, which states that if we construct within Σ a function $f: \boldsymbol{R} \to \boldsymbol{R}$ having a certain property, then we can construct a continuous f having the same property. In fact more is true, the canonical function which is constructed by unwinding the formalized proof of the existence of f is continuous. After these results have been proved as metatheorems, it becomes constructively meaningful to build a new formal system to incorporate the metatheory. This is not attractive at present, since the theory of spreads has found no significant mathematical applications.

In [**1**] I remarked that the constructive theory of Banach algebras, given in chapter 11, was forced and unnatural, and that some metatheory was indicated, to smooth the transition from the constructive to the classical proofs. Although it is too early to speak with assurance, it appears that for the particular application I had in mind the metatheory in question is provided by the numerical (or Gödel) interpretation of implication. The principle result of chapter 11 of [**1**] involves the notion of a partial ideal P of a Banach algebra $\mathfrak{A}$, determined by elements $x_1, \ldots, x_n$ of $\mathfrak{A}$ and a totally bounded subset A of $\mathfrak{A}$, and defined as

$$\begin{aligned} P &\equiv P(x_1, \ldots, x_n; A) \\ &\equiv \{x_1 y_1 + \ldots + x_n y_n : y_i \in A \text{ for } 1 \leqslant i \leqslant n\}. \end{aligned}$$

If $x_1, \ldots, x_n$ generate an ideal I of $\mathfrak{A}$, then the distance of I to the identity e is 1, so that for each A the distance of P to e is at most 1. Thus the statement $S(x_1, \ldots, x_n)$, that $x_1, \ldots x_n$ generate an ideal, can be written

$$\forall A \forall \alpha \operatorname{dist}(P(x_1, \ldots, x_n; A), e) > \alpha,$$

where A ranges over the totally bounded subsets of $\mathfrak{A}$ and α over the positive constants < 1. Now we want to constructivize the result Γ, that if $x_1, \ldots x_n$ generate an ideal I, then for each x in A there exists a complex number z such that $x_1, \ldots, x_n, x - ze$ generate an ideal; in other words, the statement $S(x_1, \ldots, x_n) \to \exists z S(x_1, \ldots, x_n, x - ze)$. A simple and trivial modification of the classical proof of Γ gives a constructive proof of the weaker statement Γ'

$$S(x_1, \ldots, x_n) \to \forall B \forall \beta \exists z \operatorname{dist}(P(x_1, \ldots, x_n, x - ze; B), e) > \beta,$$

where B ranges over the same set as A and β the same set as α. The numerical meaning of Γ' is

$$\forall B \forall \beta \exists z \exists A \exists \alpha' Q,$$

where Q is the statement

$$\operatorname{dist}(P(x_1, \ldots, x_n; A), e) > \alpha \to \operatorname{dist}(P(x_1, \ldots, x_n, x-ze; B), e) > \beta.$$

This is just the constructive substitute for Γ which was proved in [1] by extremely tedious considerations.

Certain other results in the theory of Banach algebras, not treated in [1], cannot be constructivized so simply. For example, an unpublished result Ω of the author, first published in [7], states that if γ is a differentiable arc in complex n-space C^n, then every continuous function $f\colon \gamma \to C$ can be uniformly approximated on γ by polynomials in the coordinates $z_1, \ldots, z_n$ of C^n. Stolzenberg informs me he has constructivized the classical proof, in case γ is analytic, by no means a simple task. To try to get a cheap constructivization, note that for each k the distance $d_k(f, \gamma)$ of f to the polynomials of degree k, with coefficients bounded in absolute value by k, is computable. (The proof is left to the reader.) Thus our statement Ω has the form

$$\forall f \forall \gamma \forall m \exists k \exists j (j > m \wedge d_k^j(f, \gamma) < m^{-1}),$$

where f ranges over the set of functions from γ to C, γ over the differentiable arcs in C^n, and $d_k^j(f, \gamma)$ is the jth rational approximation to the real number $d_k(f, \gamma)$. Actually this is an abbreviation of

$$\forall f \forall \gamma \forall m (\forall y\, A(f, \gamma, y) \to \exists k \exists j (j > m \wedge d_k^j(f, \gamma) < m^{-1})),$$

or

$$\forall f \forall \gamma \forall m \exists k \exists j \exists y (A(f, \gamma, y) \to (j > m \wedge d_k^j(f, \gamma) < m^{-1})),$$

where now the variables f, γ, m, k, j and y range over all functions of certain types, and $\forall y\, A(f, \gamma, y)$ is the statement that f and γ actually belong to the above-mentioned sets. Thus our approximation statement Ω is an $\forall\exists$-theorem, which means it can be written in the form $\forall u \exists v P(u, v)$, where P is decidable. Now it is an empirical observation that $\forall\exists$-theorems of classical mathematics tend to be constructively valid, so we suspect there is a metatheorem to that effect. For instance, we might try to prove the metatheorem M that every $\forall\exists$-theorem provable in the system Σ' obtained by adjoining the axiom of the excluded middle to Σ is constructively valid. (Presuming the classical proof of the above theorem Ω can be formalized within Σ', such a metatheorem M would imply that Ω is constructively valid.) Now Spector [6] has proved such a metatheorem M. Unfortunately, his proof involves an inductive procedure whose meaning is unclear, and so the question is still open. At least for certain special theories, such a metatheorem could be extremely useful.

The system Σ, with the predictive interpretation of implication, can certainly be presented as a programming language like fortran and algol. As stated before, each theorem T of Σ has the form $\exists x \forall y A(x, y)$, where x is constructed in the proof of Σ. We should be able to write a compiler for the language Σ, so that whenever a proof of such a theorem T is read into our computer, the computer will compile a program to compute the constructed quantity x. What do we mean by a program to compute a given function x of a given type? Without loss of generality, we consider only functions x whose values are finite sequences of integers of a given length n. Our question can be answered by induction. Presumably we know what it means to program the computer to compute a given integer. If x is an n-tuple $(x_1, \ldots, x_n)$ of functions, to program the computer to compute x we program it to compute each of the functions $x_1, \ldots, x_n$. Finally, if x is of type $t \to [0]$, a program to compute x is a program to compute $x(u)$ for an arbitrary argument u of x. Since $x(u)$ is determined by the values of u, the only information the program will need about u is the values of u at certain of its arguments $\gamma_1, \ldots, \gamma_k$, where of course γ_i may be a function of $u(\gamma_1), \ldots, u(\gamma_{i-1})$, and k itself may be determined in the course of the computations. Thus the program for computing x will contain as subprograms the programs for computing $\gamma_1, \ldots, \gamma_k$. The computation of $u(x)$ will request the values $u(\gamma_1), \ldots, u(\gamma_k)$ at certain junctures, which will be supplied by the program for computing u we supply the machine when we request $x(u)$ in conjunction with the programs for computing the γ_i (which are part of the program for computing x). Thus by induction we see what it means to program the computation of a given function x. Of course, the program may call on other programs, or subroutines, representing proofs of theorems referred to in the proof of the given theorem. Definitions may be called as well. Types of functionals will presumably be established by type declarations, as in algol.

As an example, consider the theorem of Koksma [**4**] that the set A of all $\theta > 1$, for which the powers $\{\theta^n\}_{n=1}^{\infty}$ are equidistributed modulo 1, is a full subset of the set of real numbers > 1, which means its complement has measure 0. Jonathan Tennenbaum has asked whether it is possible to explicitly exhibit an element θ of A, in the sense we can compute an arbitrary term of the decimal expansion of θ. The answer hinges on whether Koksma's proof is constructive. It is, except at one point, and that point is easily constructivized. The upshot is that the answer to Tennenbaum's question is 'yes'. Thus, having realized Σ as a programming language, we could feed the formalized constructivized version of Koksma's theorem into the computer, then feed in a positive integer n, and request the machine to output the nth term of the decimal expansion of the fixed θ constructed in the proof. We might have to wait a long time.

It would be interesting to take Σ as the point of departure for a reasonable programming language, and to write a compiler.

ADDED IN PROOF. The author wishes to thank G. Kreisel, J. Myhill, and G. Stolzenberg for correcting inaccuracies in the original draft. Myhill notes that in view of such definitions as that of $(f, u) \in C$, it would be more comfortable to extend Σ by adding as axioms the formulas which specify the numerical interpretation of implication. On further research, I have not been able to find in [5] or elsewhere in Kreisel's work the quoted result, that every f mapping sequence of integers into integers is continuous. Kreisel proves a somewhat weaker result, a form of pointwise continuity. However, the quoted result is valid. The numerical interpretation of statement (2), as given above, is valid only when each of the variables x, y, u, v ranges over a *basic set* – a set for which no computations are necessary to check that an element belongs to the set. In Σ, all variables range over basic sets; in informal mathematics, all statements can presumably be phrased in terms of variables ranging over basic sets.

REFERENCES

[1] E. BISHOP, *Foundations of constructive analysis*, McGraw-Hill, New York (1967).

[2] K. GÖDEL, Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes, *Dialectica* **12** (1958) 280–287.

[3] S. C. KLEENE and R. VESLEY, *The foundations of intuitionistic mathematics*, North-Holland, Amsterdam (1965).

[4] J. F. KOKSMA, Ein mengentheoretischer Satz über die Gleichverteilung modulo Eins, *Composito Math.* **2** (1935) 250–258.

[5] G. KREISEL, Functionals, ordinals, species, in *Logic, methodology and philosophy of science*, Amsterdam (1968) 145–159.

[6] C. SPECTOR, Provably recursive functionals of analysis, *Proc. Symp. on pure mathematics, Am. Math. Soc., Providence* **5** (1962) 1–27.

[7] STOLZENBERG, G., Uniform approximation on smooth curves, *Acta Math.* **115** (1966) 185–198.