Induction and Recursion

David Pierce

November 12, 2012

Contents

1 Introduction **1**

2 Number theory **2**

3 Algebra **11**

4 Propositional logic **13**

5 First-order logic **20**

6 Set theory **23**

References **29**

1 Introduction

In mathematics we use repeated activity in several ways:

1) to define sets;
2) to prove that all elements of those sets have certain properties;
3) to define functions on those sets.

These three techniques are often confused, but they should not be. Clarity here can prevent mathematical mistakes; it can also highlight important concepts and results such as Fermat's (Little) Theorem, freeness in a category, and Gödel's Incompleteness Theorem. The main purpose of the present article is to show this.

In the 'Preface for the Teacher' of his *Foundations of Analysis* of 1929 [18], Landau discusses to the confusion just mentioned, but without full attention to the logic of the situation. The present article may be considered as a sketch of how Landau's book might be updated. It is indeed a sketch, and that only. I make a number of historical references, when I have been able to consult the cited articles; but the article is not a thorough-going review of the history of the mathematical ideas discussed.

2 Number theory

For an example of the three techniques, suppose we are given the field $\mathbb{R}$ of real numbers.

1. We may define the subset $\mathbb{N}$ of natural numbers by requiring that it contain 1 and that it contain $x + 1$ for each of its elements x. Thus if a real number cannot be shown to be in $\mathbb{N}$ by application of these rules, repeated as needed, then that number is *not* in $\mathbb{N}$, by definition.

2. We may show that $\mathbb{N}$ is closed under addition, since for each element x of $\mathbb{N}$, we have by definition $x + 1 \in \mathbb{N}$, and moreover if $y \in \mathbb{N}$ and $x + y \in \mathbb{N}$, then

$$x + (y + 1) = (x + y) + 1, \qquad (*)$$

which again is in $\mathbb{N}$ by definition. Thus if A is the set of elements y of $\mathbb{N}$ such that, for all x in $\mathbb{N}$, the sum $x + y$ is in $\mathbb{N}$, then A contains 1 and is closed under adding 1. Therefore $A = \mathbb{N}$, by definition of the latter. Similarly, $\mathbb{N}$ is closed under multiplication, since for all x in $\mathbb{N}$ we have $x \cdot 1 = x$, and if $x \cdot y \in \mathbb{N}$, then

$$x \cdot (y + 1) = x \cdot y + x, \qquad (\dagger)$$

which we now know to be in $\mathbb{N}$.

3. Similarly, if we have the Gamma function

$$x \mapsto \int_0^\infty \frac{t^{x-1}}{\mathrm{e}^t}\ \mathrm{d}\,t$$

on $(0, \infty)$, then we can show that $\mathbb{N}$ is closed under it, using integration by parts. Alternatively, we can just define this operation on $\mathbb{N}$ by requiring $\Gamma(1) = 1$ and $\Gamma(x+1) = \Gamma(x) \cdot x$, so that in general $\Gamma(x)$ is what is usually denoted by $(x-1)!$.

The definition of $\mathbb{N}$ here is **inductive**; the proof that $\mathbb{N}$ is closed under addition and multiplication is **inductive**; the second definition of Γ on $\mathbb{N}$ is **recursive.** Alternatively, the definition of $\mathbb{N}$ can also be called recursive; or the second definition of Γ can be called inductive. But I shall argue that either of these alternatives is misleading.

Our inductive definition of $\mathbb{N}$ may be considered as informal. Formally, we may define $\mathbb{N}$ as the *intersection* of the set of all subsets of $\mathbb{R}$ that contain 1 and are closed under adding 1. Alternatively, $\mathbb{N}$ is the *union* of the set of subsets A of $[1, \infty)$ such that $x - 1 \in A$ for every x in $A \setminus \{1\}$. In any case $\mathbb{N}$ is the *smallest* subset of $\mathbb{R}$ that contains 1 and is closed under adding 1. That is to say, $\mathbb{N}$ admits **proof by induction:** every subset B of $\mathbb{N}$ is equal to $\mathbb{N}$, provided we can show that B contains 1 and contains $x + 1$ whenever it contains x. The whole point of the inductive definition of $\mathbb{N}$ is to ensure that $\mathbb{N}$ admits inductive proofs.

To avoid taking $\mathbb{R}$ for granted, we may try a direct axiomatic approach to $\mathbb{N}$. (This is the point of Landau's book.) We can just declare that $\mathbb{N}$ is a set that

1) has an element 1,
2) is closed under an operation $x \mapsto x + 1$, and
3) admits inductive proofs.

Then we obtain the operations of addition and multiplication that we obtained before. Indeed, we already know how to add 1 to an element x of $\mathbb{N}$: the result is simply $x + 1$. If we know what $x + y$ is, then we can use $(*)$ as a definition of $x + (y + 1)$. So we have addition on $\mathbb{N}$. We go on to define $x \cdot 1$ as x, and if we know what $x \cdot y$ is, we define $x \cdot (y + 1)$ as in $(\dagger)$.

This is all true; and yet, in saying it this way, we have cheated. For, it would be *false* to say by analogy that we can make the definition

$$x^1 = x, \tag{‡}$$

and if we know what x^y is, then we can make the definition

$$x^{y+1} = x^y \cdot x. \tag{§}$$

How can this be? Though the reader may not yet be fully in the know, s/he may have observed that our axiomatic treatment of $\mathbb{N}$ has omitted two of Peano's axioms of 1889 [22]:

4) the operation $x \mapsto x + 1$ is injective, but
5) its range does not contain 1.

These axioms turn out not to be needed for the definitions of addition and multiplication on $\mathbb{N}$; but they or at least *something* more is needed for exponentiation on $\mathbb{N}$.

Again, how can this be? Let us first observe that it *is* so, by noting that the Induction Axiom is available in modular arithmetic, although exponentiation as a binary operation is *not* generally definable there. Indeed, in the *Disquisitiones Arithmeticae* of 1801 [10, ¶50], which is apparently the origin of our notion of modular arithmetic, Gauss reports that Euler's first proof of Fermat's Theorem was as follows. Let p be a prime modulus. Trivially $1^p \equiv 1$ (with respect to p or indeed any modulus). If $a^p \equiv a$ (*modulo* p) for some a, then, since $(a+1)^p \equiv a^p + 1$, we conclude $(a+1)^p \equiv a+1$. This can be understood as a perfectly valid proof by induction in the ring with p elements that we denote by $\mathbb{Z}/p\mathbb{Z}$: we have then proved $a^p = a$ for all a in this ring.

However, Dyer-Bennet showed in 1940 [7] that, with respect to a modulus n, all congruences $a \equiv b$ and $c \equiv d$ entail the congruence $a^c \equiv b^d$ if and only if n is one[1] of 1, 2, 6, 42, and 1806. We conclude:

Theorem 1. *For all n in $\mathbb{N}$, The finite ring $\mathbb{Z}/n\mathbb{Z}$ has a binary operation $(x, y) \mapsto x^y$ satisfying the identities (‡) and (§) if and only if $n \in \{1, 2, 6, 42, 1806\}$.*

[1] Dyer-Bennet names G. Birkhoff as having suggested the problem of finding these n and as having found them independently. I found Dyer-Bennet's article through *The on-line encyclopedia of integer sequences.*

Let us observe in passing that the sequence of moduli here arises from what Mazur [21] calls the *self-proving* formulation of Euclid's Proposition IX.20 in the *Elements* [9]: give me some primes, and I'll give you another one by multiplying yours together, adding 1, and finding a prime divisor of the result. Indeed, the product of *no* primes should be considered as 1, and then:

$$\begin{aligned} 1+1 &= 2, \text{ prime;} \\ 2+1 &= 3, \text{ prime;} \\ 2\cdot 3+1 &= 6+1 = 7, \text{ prime;} \\ 2\cdot 3\cdot 7+1 &= 42+1 = 43, \text{ prime;} \\ 2\cdot 3\cdot 7\cdot 43+1 &= 1806+1 = 1807. \end{aligned}$$

It turns out that since $1807 = 13 \cdot 139$, the sequence of moduli in the theorem stops, although of course the set of primes continues to grow.

Little discoveries like Theorem 1 are a reason to begin the natural numbers with 1 rather than 0. When Henkin in 1960 [15] made some of the observations of the present article, he started the natural numbers with 0 and noted in effect that on $\mathbb{Z}/2\mathbb{Z}$, if (§) (namely $x^{y+1} = x^y \cdot x$) holds, then $0^y = 0$ for all y, since $y = z+1$ for some z; in particular $0^0 \neq 1$, so the equation $x^0 = 1$ fails.

Of course Henkin's argument works in $\mathbb{Z}/n\mathbb{Z}$ for every n that exceeds 1. Still, $\mathbb{Z}/n\mathbb{Z}$ always has an addition given by the identity $(*)$ above (namely $x+(y+1) = (x+y)+1$). At the beginning of the *Disquisitiones,* Gauss notes that addition of integers respects congruence; but apparently he does not feel the need to prove it. However, we may establish the identity $(*)$ on $\mathbb{Z}/n\mathbb{Z}$ as follows.

We assume that we are given the operation $x \mapsto x+1$. As an inductive hypothesis, we assume too that we 'know' $x+b$ for some b; that is, we assume there is an operation $x \mapsto x+b$. But this is not just any operation; it satisfies the identities

$$1+y = y+1, \qquad (x+1)+y = (x+y)+1 \qquad (\P)$$

when $y = b$. Note that these equations are vacuously true when $y = 1$. If we now use $(*)$ when $y = b$ to define $x \mapsto x+(b+1)$, then as a special

case we have

$$1 + (b + 1) = (1 + b) + 1,$$

so by the inductive hypothesis (¶) we have $1 + (b + 1) = (b + 1) + 1$. Similarly

$$\begin{aligned}(x + 1) + (b + 1) &= ((x + 1) + b) + 1 && \text{[by } (*) \text{ when } y = b\text{]}\\ &= ((x + b) + 1) + 1 && \text{[by } (\P) \text{ when } y = b\text{]}\\ &= (x + (b + 1)) + 1. && \text{[by } (*) \text{ when } y = b\text{]}\end{aligned}$$

Thus (¶) holds when $y = b+1$. Therefore on $\mathbb{Z}/n\mathbb{Z}$, as on $\mathbb{N}$, for every y, there is at least one operation $x \mapsto x+y$ satisfying (¶). All we have used to establish this is induction (along with the element 1 and the operation $x \mapsto x + 1$; but the Induction Axiom assumes that these exist).

By induction also, each of the operations $x \mapsto x + y$ satisfying (¶) is unique. Indeed, suppose when $y = b$ there is one such function, but f is another, that is, $f(1) = b + 1$ and $f(x + 1) = f(x) + 1$. Then by (¶) when $y = b$ we have $1 + b = f(1)$, and if $f(a) = a + b$, then $(a + 1) + b = (a + b) + 1 = f(a) + 1 = f(a + 1)$. Thus $x + b = f(x)$ for all x, that is, f is the function $x \mapsto x + y$.

Now we have a unique operation $(x, y) \mapsto x+y$ satisfying (¶). By looking back at the proof, we conclude that $(*)$ is an identity. Indeed, we used this equation to define an operation $x \mapsto x + (b+1)$ from $x \mapsto x+b$, and since these operations are now known to exist uniquely, $(*)$ must hold. However, Peano [22, p. 95] uses this equation by itself as a definition of addition, writing:[2]

> This definition has to be read as follows: if a and b are numbers, and if $(a+b)+1$ has a meaning (that is, if $a+b$ is a number) but $a+(b+1)$ has not yet been defined, then $a + (b + 1)$ means the number that follows $a + b$.

Is Peano correct? Can we take $(*)$ as a definition in his sense? If so, then we should be able to take the equations (‡) and (§) as a definition of exponentiation on, say, $\mathbb{Z}/3\mathbb{Z}$. When Peano makes his remark, he has stated all of his axioms, and $\mathbb{Z}/3\mathbb{Z}$ does not satisfy all of them; still, it satisfies the Induction Axiom, and Peano does not appeal to any other

[2] Peano has $(*)$ in the form $a + (b + 1) = (a + b) + 1$.

axioms, or a lemma derived from them, to justify his remark. Following Peano's procedure in $\mathbb{Z}/3\mathbb{Z}$ then, we get the successive rows of the following table:

x	1	2	3
x^1	1	2	3
x^2	1	1	3
x^3	1	2	3

We make no new row for x^4, since $4 = 1$ in $\mathbb{Z}/3\mathbb{Z}$, so x^4 has already been defined. If we did try to make a row for x^4, using (§), then it would not agree with the row for x^1. Thus, although we can use equations (‡) and (§) to give a definition, in Peano's sense, of exponentiation in $\mathbb{Z}/3\mathbb{Z}$, those equations are not identities under the definition.

Logically then, although we can use the rule $(*)$ by itself to build up an addition table for $\mathbb{N}$ or $\mathbb{Z}/n\mathbb{Z}$, it does not follow that $(*)$ is an identity. This needs an additional argument.

Somewhat modernized, Peano's thinking seems to be this. Let A be the set of all y such that an operation $x \mapsto x + y$ is defined. Then $1 \in A$. Moreover if $b \in A$ then, since we can define $x + (b+1)$ by $(*)$ when $y = b$, it follows that $b+1 \in A$. By induction, A contains all y. But this gives us no unique operation of addition. Indeed, assuming $b \in A$, we can show $b + 1 \in A$ by defining $x + (b + 1)$ as $x + b$ or even 1. What we must do is something like what we did: let A be the set of all y such that an operation $x \mapsto x + y$ is defined *so as to satisfy* (¶).

Now I claim to have shown what I said at the beginning, that the definition of addition by means of $(*)$ should not be called inductive, because such definitions are not generally justified by induction alone. The underlying observation here is not original; again, Henkin makes it, and before him, Landau. (Landau in turn credits Kalmár with the special proof that addition can indeed be established by induction alone. Landau does not mention that only induction is used; nor does he give an example like exponentiation, where induction is definitely not enough.) Using y' for $y + 1$, Landau writes in his 'Preface for the Teacher':

> On the basis of his five axioms, Peano defines $x + y$ for fixed x and all

y as follows:

$$x + 1 = x'$$
$$x + y' = (x + y)',$$

and he and his successors then think that $x + y$ is defined generally; for, the set of y's for which it is defined contains 1, and contains y' if it contains y.

But $x + y$ has *not* been defined.

Landau once shared the confusion of Peano and his successors; the fact of this earlier confusion is a reason for publishing his book. Nevertheless, despite the warnings of Landau, Henkin, and others,[3] confusion about these basic matters persists.

I suggest that Landau himself is a bit confused about what an axiom is; at least, he fails to make a distinction that we find it worthwhile to make today. Peano himself gives *nine* axioms for $\mathbb{N}$, but three of them are the reflexive, symmetric, and transitive properties of equality of numbers, and another is that something equal to a number is a number. Landau rightly sets these aside as being purely logical properties,[4] not specific to elements of $\mathbb{N}$. Peano's remaining five axioms are those mentioned by Landau and also given earlier in the present article. However, two of those, namely that $\mathbb{N}$ contains 1 and is closed under adding 1, are simply features of the structure[5] of $\mathbb{N}$, features whose properties are fixed by the remaining three axioms.

Burris gives these three axioms at the head of 'The Dedekind–Peano Number System', an appendix to his *Logic for Mathematics and Computer Science* [3], an excellent book from which I have learned a lot. After stating the axioms, Burris defines addition as Peano does. As we have seen, the definition *is* justified; but it is not *obviously* justified. The student may come away from that appendix with the wrong impression.

[3] Dedekind was perhaps the first to give such a warning; see below.

[4] Perhaps Peano himself, or one of the followers mentioned by Landau, had already done this setting aside.

[5] That is, they are formally entailed by considering $\mathbb{N}$ as a structure in a signature with a constant for 1 and a singulary operation-symbol for $x \mapsto x + 1$. See §3 below.

A similarly wrong impression may be got from Mac Lane and Birkhoff's *Algebra* [20, p. 15], where right after the Peano axioms are given, it is said that the natural numbers can serve to index iterates of singulary ('unary') operations. If this is so, then one might expect elements of $\mathbb{Z}/3\mathbb{Z}$ to serve as indices of iterated products—that is, as exponents of powers—in $\mathbb{Z}/3\mathbb{Z}$ (or in any $\mathbb{Z}/n\mathbb{Z}$, or $\mathbb{Z}$ itself); but as we have seen, this is not possible.

Also in his 'Preface for the Teacher', Landau warns,

> My book is written, as befits such easy material, in merciless telegram style (**'Axiom', 'Definition', 'Theorem', 'Proof',** occasionally **'Preliminary Remark',** rarely words which do not belong to one of these five categories).

But the material is *not* easy. Perhaps it is the assumption that the material *is* easy that led Landau and others to be confused about it in the first place. Such confusion could have real mathematical consequences: it might lead one to replace Fermat's Theorem with a false belief that $a^{p+1} \equiv a$ is an identity *modulo* p.

Landau is not concerned with noting what can be proved by induction alone; the point of his book is just to construct the fields of real and complex numbers, so the analyst can use them in good conscience. Nonetheless, it seems worthwhile to note that, after defining addition on $\mathbb{N}$ as we have done, we can go on to establish, again by induction alone, that addition is commutative and associative. Using also that the operation $x \mapsto x + 1$ is injective, we obtain that addition is cancellative (in the sense that $x + z = y + z$ implies $x = y$). Again by induction alone, there is a unique operation of multiplication, which is commutative and associative, and which distributes over addition, although it need not be cancellative. Thus we have $\mathbb{N}$ as a commutative semi-ring; but then we also have the set $\{1, \ldots, n\}$ as a commutative semi-ring when we define $x \mapsto x + 1$ on this set as in the following table, so that this operation is

injective, and proof by induction is available.[6]

x	1	...	$n-1$	n
$x+1$	2	...	n	1

Here $n + 1 = 1$, and then by induction $n + x = x$, so n is neutral for addition; also then, every element has an additive inverse, so the set $\{1, \ldots, n\}$ becomes the ring $\mathbb{Z}/n\mathbb{Z}$. Of course, once one has the ring-structure of $\mathbb{Z}$, derived perhaps from the semi-ring structure of $\mathbb{N}$, then it is easy to show that congruence *modulo* n respects this structure, so that $\mathbb{Z}/n\mathbb{Z}$ is ring. Still, it seems worthwhile to note that most of this *has already been shown* in establishing the semi-ring structure of $\mathbb{N}$, because the very same arguments work for $\mathbb{Z}/n\mathbb{Z}$.

Again, the attempt to define exponentiation by induction alone breaks down almost completely. For every n in $\mathbb{N}$, for every element x of $\mathbb{Z}/n\mathbb{Z}$, there is of course a function $y \mapsto x^y$ from $\mathbb{N}$ to $\mathbb{Z}/n\mathbb{Z}$ satisfying the identities (‡) and (§); but this needs more than induction. The full result is the following.

Theorem 2. *For every set A that has an element 1 and is closed under an operation $x \mapsto x + 1$, the following are equivalent conditions.*

1. *The operation $x \mapsto x+1$ is injective, but its range does not contain 1, and no proper subset of A contains 1 and is closed under the operation.*
2. *For every set B that has an element c and is closed under an operation $x \mapsto f(x)$, there is a unique function g from A to B satisfying the identities*

$$g(1) = c, \qquad\qquad g(x+1) = f(g(x)).$$

Dedekind's work on the natural numbers predates Peano's, and his mathematical understanding seems to be more profound. He gives the forward direction of Theorem 2 in his *Nature and Meaning of Numbers* of 1887

[6] It may be said that we do not know what $\{1, \ldots, n\}$ means unless we have defined the ordering of $\mathbb{N}$, so that $\{1, \ldots, n\} = \{x \in \mathbb{N} \colon 1 \leqslant x \leqslant n\}$. The theorem that $\mathbb{N}$ can indeed be linearly ordered in the usual way does require all of the Peano axioms. Without proving this theorem though, we can still define $\{1, 2\}$, then $\{1, 2, 3\}$, and so on, as far as we like.

[6, II(126), p. 85]. It is an accident of history that the Peano axioms are usually so called. Peano does give us some notation, which has perhaps helped solidify the ideas behind it. Russell and Whitehead may have helped spread the notation through the *Principia Mathematica*: their sign $\supset$ for implication is derived from Peano's reversed C, and they use Peano's sign $\in$ for membership of an individual in a class (originally the sign is an epsilon, for the Greek ἐστί 'is' [26, pp. 25–26]). Dedekind himself does not distinguish between this membership relation and the subset relation: he used the same sign for both, looking something like a 3 or perhaps a *reversed* epsilon ([6, II(3), p. 46] or [5, p. 3]).

Henkin [15, p. 337] proves the reverse direction of Theorem 2, but does not explicitly mention any earlier proof. If the theorem is difficult, the difficulty lies in recognizing that such a theorem *might* be true; once one can recognize this possibility, proving the theorem is not that hard.[7] Similarly, it is not so hard to prove the independence of Euclid's Parallel Postulate from his others; but it seems to have taken more than two thousand years to recognize the possibility of such a proof.

3 Algebra

Theorem 2 can be understood as being about *algebras* in a *signature* with one constant and one singulary operation-symbol. In the most general sense, an **algebra** is a set with some distinguished operations and elements; those operations and elements are given symbols, which compose the **signature** of the algebra. Theorem 2 is about an algebra $(A, 1, x \mapsto x + 1)$, or more simply $\mathfrak{A}$. In a word, the second condition in the theorem is that $\mathfrak{A}$ admits **recursion**; more elaborately, the algebra admits **recursive definition of homomorphisms.** Another way to say this is that the algebra is **absolutely free:** that is, it is a free object in the category of *all* algebras of its signature. In the first condition, the

[7] One can obtain the function g as the union of the set of all sets $\{(1, c), (2, f(c)), (3, f^2(c)), \ldots, (n, f^{n-1}(c))\}$; this is Dedekind's approach. Alternatively, g is the intersection of the set of all relations R from A to B such that $1\ R\ c$ and $(x + 1)\ R\ (f(y))$ whenever $x\ R\ y$. In the words of Enderton [8, p. 35], these are the bottom-up and top-down approaches, respectively.

part about not having certain proper subsets is that $\mathfrak{A}$ has no proper *subalgebras*; in a word, $\mathfrak{A}$ is **minimal.**

Again, such minimality is not enough to establish recursion. Dedekind [6, ¶130, p. 89] notes this, observing in effect that there is no homomorphism from $\mathbb{Z}/2\mathbb{Z}$ to $\mathbb{Z}/3\mathbb{Z}$, even though the former admits induction. (Nonetheless, Dedekind does refer to definition by recursion as definition by induction.)

If we understand an element of a set as a nullary operation on the set, then Theorem 2 can be understood as a special case of the following.

Theorem 3. *An arbitrary algebra is absolutely free if and only if:*

1) the algebra is minimal,
2) each of its operations is injective,
3) the ranges of any two of those operations are disjoint.

To establish the sufficiency of the three conditions,[8] one proceeds as one would for the corresponding implication in Theorem 2: given an algebra $\mathfrak{A}$ meeting those conditions and an arbitrary algebra $\mathfrak{B}$ of the same signature, one obtains a unique homomorphism from $\mathfrak{A}$ to $\mathfrak{B}$, either as the intersection of the set of all *relations* from A to B with the appropriate properties, or as the union of the set of certain *partial* functions from A to B. For the necessity of the three conditions, one observes that all absolutely free algebras of a given signature are isomorphic to one another; then it is enough to observe that one example meets the conditions. In the situation of Theorem 2, one already has such an example, or rather, one *assumes* that one has an example, namely $(\mathbb{N}, 1, x \mapsto x + 1)$. In the case of a signature with no constants, not only are all absolutely free algebras isomorphic to one another, but they are identical with one another: they are empty.

In case there *are* constants, the natural example of a free algebra would seem to be the *term algebra,* as described for example by Hodges [16, §1.3, p. 14]. I want to work out one case, by way of arguing that it is indeed

[8] See note 7 above. Enderton [8, p. 39] establishes sufficiency in case the signature has one binary, one singulary, and arbitrarily many nullary operation-symbols. We shall look at a similar case in §4 below.

mathematically worthwhile to be aware of Theorem 3 in its generality, and not only Theorem 2.

4 Propositional logic

In his 'automathography' [13, p. 206], Halmos writes:

> An exposition of what logicians call the propositional calculus can annoy and mystify mathematicians. It looks like a lot of fuss about the use of parentheses, it puts much emphasis on the alphabet, and it gives detailed consideration to 'variables' (which do not in any sense vary). Despite (or because of?) all the pedantic machinery, it is hard to see what genuine content the subject has. Insofar as it talks about implications between propositions, everything it says looks obvious. Does it really have any mathematical value?
>
> Yes it does... Question: what is the propositional calculus? Answer: the theory of free Boolean algebras with a countably infinite set of generators.

It is good that Halmos found a way to understand logic that satisfied him; but he seems to have missed the point. For one thing, propositional logic is not just about free *Boolean* algebras: it is about certain *absolutely free* algebras, in the sense above, from which Boolean algebras are obtained as *quotients.* This is how the fuss and pedantry arises that Halmos decries; but I think it is inevitable, and I hope to make it a little more interesting below.

Meanwhile, logic should be understood as *foundational* for mathematics. One *can,* generally *does,* and probably *must* learn mathematics before symbolic logic; but if one wants to formalize one's work, at least by way of checking for mistakes, then one will reduce one's mathematics to logic, and not the other way around as Halmos does.

I shall describe here a version of the prositional calculus that Church [4, ch. I] develops from that of Łukasiewicz. We first choose a set V of propositional 'variables'. In the algebraic sense above, these variables will indeed be constants. One does generally want V to be countably infinite, but this will make little difference for us. Actually, it is perhaps

philosophically best to consider V as finite, as long as it can be made as large as necessary to cover any particular situation.

We define a set of propositional formulas by three rules:

1) every variable is a formula,
2) 0 is a formula, and
3) if F and G are formulas, then so is the *implication* $(F \to G)$.

Thus we obtain a *term algebra* in the signature $V \cup \{0, \to\}$. We may understand this definition to pick out the formulas from the set of all *strings* of our symbols, just as our original definition of $\mathbb{N}$ picked out its elements from $\mathbb{R}$. This set of strings might be formalized as the set of functions from sets $\{1, \ldots, n\}$ or $\{0, \ldots, n-1\}$ into our set of symbols. Again though, this way of thinking is backwards or anachronistic, if we think of symbolic logic as being developed for the sake of formalizing the notions of sets and functions and numbers in the first place. One should understand a string of symbols as something that can be written down, on paper, left to right in the usual way. From the given definition of formula, it is easy enough to show that any particular written string is a formula; and that is all we need.

We want to know that the algebra of formulas is *absolutely free.* That is, we want formulas to have **unique readability.** If we accept Theorem 3, then since by definition the algebra of formulas is minimal, we need only show that (1) the operation $(F, G) \mapsto (F \to G)$ is injective and (2) its range contains neither 0 nor a variable. These facts correspond to the two Peano axioms other than Induction; but in the present case they are theorems. One of the theorems is trivial, since all implications have more than one symbol. The other theorem is that if $(F \to G)$ and $(H \to K)$ are the same formula, then F and H are the same formula (and hence also G and K are the same), assuming F, G, H, and K are formulas in the first place. This follows from the lemma that no proper initial segment of a formula is a formula. One can prove this by induction on the *lengths* of formulas: that is, one can prove it as a theorem about the algebra $\mathbb{N}$. Again this might be anachronistic, if one is developing logic in order to formalize $\mathbb{N}$. Alternatively, one can prove simultaneously by induction in the algebra of formulas that every formula neither (1) has a formula as a proper initial segment, nor (2) is itself a proper initial segment of another formula.

It may be legitimate to consider unique readability of formulas as being obvious. From school arithmetic, we have the sense that we can always put in enough brackets to make a given expression unambiguous. In our present system, *every* implication is surrounded by brackets; is it not obvious that this is enough to ensure unique readability? Church seems to think so. He first notes in passing [4, p. 38, n. 91] that brackets can be dispensed with entirely by using the prefix notation of Łukasiewicz, but does not dwell on the issue. Later [4, §10, pp. 70–71] he gives an algorithm for determining whether a given string is a formula. Given a string beginning and ending with brackets, the algorithm aims to detect (by counting brackets) an arrow of implication in the string such that, if the string *is* a formula, then the two strings between the arrow and the outer brackets must be formulas. Then the algorithm is applied in turn to those two strings, and so on. Church calls the arrow found by the algorithm the *principal implication sign* of the formula. Now, by definition an implication must have *a* principal implication sign; but Church does not exactly *prove* that his algorithm finds such a sign in every well-formed implication. Even if this is granted, there remains the question of whether the sign is unique.[9]

Whether one proves it or not, unique readability should be stated clearly, in order to be able to emphasize, as we shall do below, what might be called the 'non-unique readability' of *theorems.*

Meanwhile, we need unique readability of formulas to define **interpretations,** namely functions h from the algebra of formulas into $\{0, 1\}$ such that $h(0) = 0$ and

$$h((F \to G)) = 1 \text{ if and only if } h(F) = 0 \text{ or } h(G) = 1.$$

[9]Burris proves unique readability for the Łukasiewicz notation only, but seems to take unique reability of bracketed infix notation as obvious [3, pp. 39, 142, 197]. Enderton notes in effect [8, p. 30] that unique readability is immediate if the formula $(F \to G)$ is understood to be the ordered triple $(F, \to, G)$. Again this is anachronistic: it begs the question of whether we *can* write down formulas in such a way that an arbitrary implication can be unambiguously analyzed as an ordered triple $(F, \to, G)$. If writing the formula as a *string* $(F, \to, G)$ is enough, then it is enough to write it without the commas, as $(F \to G)$. But the brackets *are* needed; why is that? Here one may start to feel the need to *prove* that the brackets are enough; and Enderton does in fact supply a proof.

(So one thinks of 0 as *falsehood,* and 1 as *truth.*) An interpretation then is determined by its restriction to the set V of variables. In the spirit of Halmos, we may note that an interpretation is a homomorphism into the two-element field, if we there understand $x \to y$ to be $x \cdot y + x + 1$. Moreover, if two formulas are understood to be **equivalent** if their images are equal under every interpretation, then the set of equivalence-classes of formulas can indeed be understood as a Boolean algebra, at least when we make the usual definitions: the *negation* $\neg F$ is $(F \to 0)$, the *disjunction* $(F \vee G)$ is $(\neg F \to G)$, and the *conjunction* $(F \wedge G)$ is $\neg(\neg F \vee \neg G)$. But again, the point of logic is to show how such a Boolean algebra can be obtained in the first place. The Boolean algebra of equivalence-classes of propositional formulas in two variables can be depicted in the Hasse diagram in Figure 1; but a question that logic must take up is how a

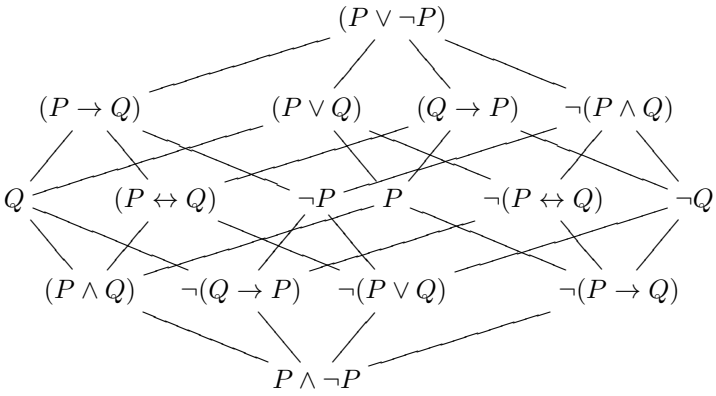

Figure 1: Hasse diagram for an algebra of equivalence-classes of formulas

node in the diagram can be named so that its position in the diagram can be inferred, even if the name is seen in isolation from other names and the diagram itself. A naming scheme as in Figure 2 would not be very useful.

In the foreword of his *Algebra,* Lang writes [19, p. v]:

> Unfortunately, a book must be projected in a totally ordered way on the page axis, but that's not the way mathematics 'is', so readers have

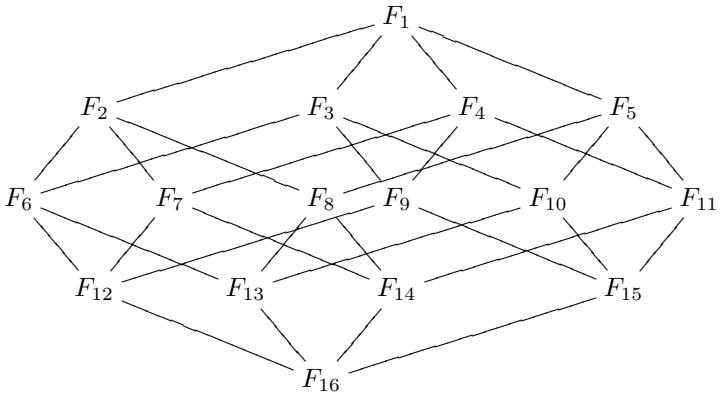

Figure 2: Bad names for equivalence-classes of formulas

> to make choices how to reset certain topics in parallel for themselves, rather than in succession.

Logic might be said to investigate how this projecting can be done so that readers are indeed able to recover what they want.

Again, to define an interpretation of formulas, we use the unique readability of formulas. We use this also to define a certain binary operation on the set of formulas. Let us use the symbol $*$ for this operation, so that we can define it by

$$F * G = \begin{cases} H, & \text{if } G \text{ has the form } (F \to H); \\ G, & \text{otherwise.} \end{cases}$$

Such an operation is a **rule of inference** (this one being called *Modus Ponens* or Detachment). We do not actually need unique readability in order to define, as **logical axioms,** all formulas of one of the forms

$$(F \to (G \to F)),$$
$$((F \to (G \to H)) \to ((F \to G) \to (F \to H))),$$
$$(((F \to 0) \to 0) \to F).$$

We obtain the set of **theorems** by taking the set of logical axioms and closing under $*$ within the set of formulas. The set of theorems can then be understood as a minimal algebra in a signature with $*$ and with a constant for each logical axiom. However, unlike the algebra of formulas, the algebra of theorems is not free.

Because of the freeness of the algebra of formulas, each formula has a unique *parsing tree*; for example, one of the logical axioms has the parsing tree in Figure 3, where P, Q, and R are propositional variables. Strictly the tree is an *ordered* tree, in the sense that left branches must

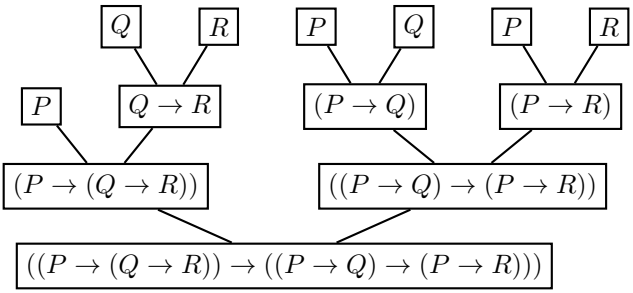

Figure 3: a parsing tree

be distinguished from right branches.

The minimality of the algebra of theorems means that each theorem has a *proof,* which also can be understood as a tree; but this proof is not unique. For example, the theorem $(P \to P)$ has the proof shown in Figure 4, where F stands for a formula $(G \to P)$, where G can be *any* formula. We may debate whether a proof is 'really' a tree, as opposed to a linearly ordered refinement of the tree, such as

$$\begin{gathered}((P \to ((G \to P) \to P)) \to ((P \to (G \to P)) \to (P \to P))),\\ (P \to ((G \to P) \to P)),\\ ((P \to (G \to P)) \to (P \to P)),\\ (P \to (G \to P)),\\ (P \to P).\end{gathered}$$

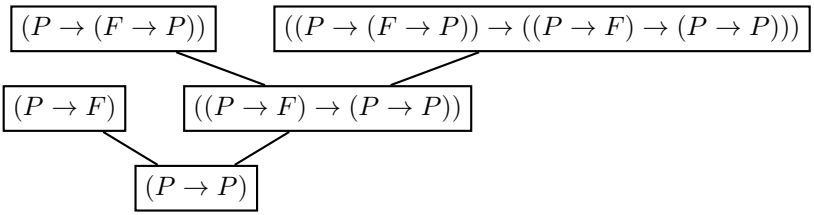

Figure 4: a proof, where F is $(G \to P)$

Normally a proof is something that can be *read,* and reading is done linearly; in the language of Lang, a proof is a totally ordered projection on the page axis. But to *understand* a proof is to understand the relation of each of its steps to other steps; and this relation places those steps as nodes of a tree.

As there is an algorithm to determine whether an arbitrary string of symbols is a formula, so there is an algorithm to determine whether an arbitrary formula is a theorem in the present sense. Indeed, by induction in the algebra of theorems, if F is a theorem, then 1 is its only interpretation, that is, the formula is **logically true.** The converse of this implication is also true, although the proof is not obvious.[10] And there is an algorithm to determine whether a formula is logically true: just check its interpretations under all of the (finitely numerous) interpretations of its variables: that is, check its *truth table.* If the formula is indeed logically true, then we can find a proof of it.

But this is a special feature of propositional logic; it fails in *first-order logic.*[11] I am going to try to describe this logic as tersely as possible for present purposes; but the logic is less neat than propositional logic. Still,

[10] One approach is the following. For every formula F, if h is an interpretation, define F^h as F itself if $h(F) = 1$, and otherwise let F^h be $(F \to 0)$. If the variables of F are among $P_0, \ldots, P_{n-1}$, then one shows, by induction in the algebra of formulas, that the formula $({P_0}^h \to \cdots \to ({P_{n-1}}^h \to F^h)\cdots)$ is a theorem. In particular, if F is logically true, then $({P_0}^h \to \cdots \to ({P_{n-1}}^h \to F)\cdots)$ is always a theorem. Then one can in turn eliminate each ${P_i}^h$, using the theorem $((P \to G) \to ((\neg P \to G) \to G))$.

[11] More precisely first-order *predicate* logic, or perhaps first-order *quantifier* logic.

we should appreciate that it is one solution found, after decades if not millenia of labor, to the problem of how mathematics can be expressed.

5 First-order logic

In first-order logic, the role of propositional variables is taken by **atomic formulas,** which express equality or other relations between *individuals*; these individuals are denoted by **terms,** that is, members of the appropriate *term algebra.*[12] Again this algebra has a signature, comprising n-ary operation-symbols for various n, including 0; but there are also **individual variables,** which—as far as constructing terms is concerned—play the role of constants, that is, nullary operation-symbols. Terms can be understood as *polynomials.*

First order logic also introduces new operations on formulas: $F \mapsto \exists x\ F$ and $F \mapsto \forall x\ F$, where x is an individual variable. (One can understand $\forall x\ \varphi$ as an abbreviation of $\neg\exists x\ \neg\varphi$.) Every occurrence of x in $\exists x\ F$ or $\forall x\ F$ is *bound.* A formula in which all occurrences of variables are bound is a **sentence.** The set of sentences is not defined inductively; rather, the function that assigns to each formula its set of *free* variables is defined recursively, and then the set of sentences is the inverse image of the empty set of variables under this function. Thus the definition of sentences does require unique readability of formulas.

An **interpretation** is still a function on the set of formulas. Its codomain is no longer $\{0, 1\}$, but is instead the family of subsets of finite Cartesian powers of some set M. The interpretation may then be denoted by $\mathfrak{M}$, or more precisely $\varphi \mapsto \varphi^{\mathfrak{M}}$. Here $\varphi^{\mathfrak{M}}$ can be understood as the *solution set* of the formula φ in $\mathfrak{M}$, so that if φ has n free variables, then $\varphi^{\mathfrak{M}}$ is a subset of M^n. The interpretation is *not* determined by its restriction to the set of variables; indeed, variables are no longer formulas. The interpretation is determined by its restriction to the set of atomic formulas. But now further analysis is possible. Each atomic formula is a combination of a relation-symbol (possibly the equals sign) and the appropriate number of terms. Each term t then has an interpretation $t^{\mathfrak{M}}$, which is an operation on M. In particular, if t has n variables, then $t^{\mathfrak{M}}$ is a function from M^n

[12] See pages 12 and 14.

to M. The map $t \mapsto t^{\mathfrak{M}}$ is defined recursively from certain operations $S^{\mathfrak{M}}$ on M that are assigned to the operation-symbols S in the logic. Then the interpretation of an atomic formula is determined by the relations $S^{\mathfrak{M}}$ on M assigned to the relation-symbols S in the logic. The whole interpretation $\mathfrak{M}$ is now determined by the set M and the operations and relations $S^{\mathfrak{M}}$ on M for the various symbols S. The interpretation is a **structure** on M.

If σ is a sentence, then by the foregoing account $\sigma^{\mathfrak{M}}$ should be a subset of M^0. But M^n can be understood as the set of functions from $\{0, \ldots, n-1\}$ to M; in particular, M^0 is the set of functions from the empty set to M. There is only one such function, the empty set, which can be denoted by 0; and then $M^0 = \{0\}$, which can be denoted by 1. Thus $\sigma^{\mathfrak{M}}$ is either 0 or 1: it is 0, if σ is **false** in $\mathfrak{M}$, and 1 if σ is **true.** More generally a formula φ with n free variables can be considered as true in $\mathfrak{M}$ if $\varphi^{\mathfrak{M}} = M^n$; but then a formula that is not true is not necessarily false.

A structure in which every formula in a given set is true is a **model** of that set. That set then **entails** every formula that is true in every model of the set. A formula entailed by the empty set is **logically true.** A set of formulas that is closed under entailment is a **theory.** By this definition, a theory has no obvious algebraic structure whereby the theory is *minimal* in the sense discussed earlier. Nonetheless, a theory can be given such an algebraic structure: this is the import of **Gödel's Completeness Theorem** of 1930 [11].

An algebraic structure on theories is a **proof system.** A proof system then is a set of *logical axioms*—certain formulas—and *rules of inference*—certain operations on formulas; the logical axioms can be considered as nullary operations. We require the logical axioms to be logically true, and the rules of inference to yield only formulas that are entailed by their arguments. Gödel's theorem is that there is a proof system[13] that

[13] As logical axioms of this proof system, we can take those of propositional logic, along with the formulas $((\varphi \to \psi) \to (\varphi \to \forall x\ \psi))$ where x does not occur freely in φ. The remaining axioms are $(\forall x\ \vartheta \to \vartheta^x_t)$, where t is a constant or variable, and ϑ^x_t is the result of replacing all free occurrences of x in ϑ with t; if t is a variable, it must not occur freely in a subformula of ϑ where x is bound. As rules of inference, we can take $*$ as before, along with the operations $\varphi \mapsto \forall x\ \varphi$.

is **complete** in the sense that every algebra of formulas with respect to this system is already a theory.[14]

The set T of formulas entailed by a set Γ of formulas is also the smallest algebra of formulas (with respect to a complete proof system) that includes Γ. Then Γ is a set of **axioms** for T. This set T is a theory, but it need not be **complete** as a theory; that is, there may be a sentence such that neither itself nor its negation is entailed by Γ. (For a formula with free variables, neither it nor its negation need belong to a given complete theory.) Presburger showed in 1930 that we can write down a set Γ of axioms such that (1) the semigroup $(\mathbb{N}, 1, +)$ is a model of Γ, and (2) the theory entailed by Γ is complete. More precisely, although Γ is infinite, we can write down as much of it as we need. Indeed, Γ contains, for each n in $\mathbb{N}$, the axiom

$$\bigvee_{k=1}^{n} (x = k \vee \exists y\ x = ny + k),$$

where k as a term stands for $1 + \cdots + 1$ (with k occurrences of 1) and ny stands for $y + \cdots + y$ (with n occurrences of y); and Γ contains finitely many other axioms.[15] However, by *Gödel's Incompleteness Theorem* of 1931 [12], the same cannot be done for the semi-ring $(\mathbb{N}, 1, +, \cdot)$: while the set of formulas that are true in this structure is complete in the present sense, it is not entailed by a set of formulas determined by a rule.[16]

[14] Gödel's proof requires introduction of new relation-symbols; Henkin's improvement [14] uses new constants. Church [4] presents both proofs.

[15] The other axioms are $x + y = y + x$, $x + (y + z) = (x + y) + z$, $x \neq y \to \exists z\ (x + z = y \vee y + z = x)$, $x + y \neq x$, and $x + y \neq 1$. See Hodges [16, pp. 85 & 128]; I have not consulted Presburger's original paper. One way to prove the theorem is to introduce a relation-symbol $<$ so that the atomic formula $x < y$ means $\exists z\ x+z = y$; and also, for each n in $\mathbb{N}$, to allow $x \equiv y \pmod{n}$ as an atomic formula. In the enlarged signature, every formula is equivalent to a formula with the same free variables, but without quantifiers; also, for every quantifier-free sentence, itself or its negation is entailed by the axioms, since itself or its negation is true in $(\mathbb{N}, 1, +)$, and this structure embeds in every model of the axioms.

[16] Gödel assigns to each formula φ a *code,* which is in $\mathbb{N}$. We can treat operations on $\mathbb{N}$ as composing a kind of algebra: the constants of this algebra are $x \mapsto x + 1$, the projections $(x_0, \ldots, x_{n-1}) \mapsto x_i$, and the constant functions $(x_0, \ldots, x_{n-1}) \mapsto c$; and there is one binary operation, converting an n-ary operation f and an $n + 2$-ary operation g into the $n + 1$-ary operation h given recursively by $h(\vec{x}, 1) = f(\vec{x})$ and $h(\vec{x}, y + 1) = g(\vec{x}, y, h(\vec{x}, y))$. The operations in the *minimal* algebra are

It is tedious to work through the number-theoretic details of Gödel's original argument, but a corresponding incompleteness result is more readily established for set theory. Moreover, if one agrees with Landau that the analyst ought not just to *assume* the real numbers, but should *construct* them from the natural numbers, then perhaps one ought to construct the natural numbers too, and not just assume them (as Landau does); and this construction can be done in set theory. Finally, set theory is a context for considering the *inductive* definitions and the minimal algebras that we have mentioned.

6 Set theory

From ordinary language, we have the notion of a *collective noun.* Singular in form, a collective noun refers to many things as one thing. The Russell Paradox of 1902 [23] is that there is no most general collective noun; for if there were, it could be the word *set,* but then there would be a set of all sets that did not contain themselves, and this set would both contain itself and not.

Nonetheless, there can be a most general collective noun for our purposes; as this noun, let us use the word *collection.* For mathematical study, we distinguish certain collections as **sets.** We determine the properties of sets axiomatically. Our language for doing this is first-order logic with no constants, no operation-symbols, and only one relation-symbol: Peano's $\in$ as mentioned above. For us this is a binary symbol that takes, as its right argument, a set, and as its left argument, a possible element of the set. It is then simplest to take *both* arguments of $\in$ as sets. Our variables then will range over sets alone.

We need not have an official symbol for equality, but can approach the matter as follows. A given singulary formula $\varphi(x)$ will define a collection, namely the collection of all sets a such that $\varphi(a)$ is true. Such a collection

called *recursive,* and a subset of $\mathbb{N}$ is called *recursive* if it is $f^{-1}(1)$ for some singulary recursive operation f. Then Gödel's theorem is that there is no complete theory entailed by formulas that are true in $(\mathbb{N}, 1, +, \cdot)$ and whose codes compose a recursive set. It follows that such a set of codes cannot even be *recursively enumerable,* that is, be the range of a recursive operation.

will be called a **class.** We consider two classes to be **equal** if they have the same members.

In particular, if a is a set, then we have the formula $x \in a$ (with the parameter a). We consider the class defined by this formula to be a itself. That is, every set is a class, namely the class of its members. In particular, two *sets* are equal if they have the same members. We can now use the expression $x = y$ as an abbreviation of the formula

$$\forall z\ (z \in x \leftrightarrow z \in y).$$

Now we can state the following.

Axiom (Equality)**.** *Equal sets are members of the same sets:*

$$\forall x\ \forall y\ (x = y \rightarrow \forall z\ (x \in z \leftrightarrow y \in z)).$$

Alternatively, if the sign of equality were an official relation-symbol, then the sentence

$$\forall x\ \forall y\ (x = y \leftrightarrow \forall z\ (z \in x \leftrightarrow z \in y))$$

would probably be taken as an axiom, the Axiom of Extensionality, as in Zermelo's original treatment of 1908 [27, p. 201]. Then the above Equality Axiom would be taken as logically true. Indeed, more would be logically true, namely

$$\forall x\ \forall y\ (x = y \rightarrow (\varphi(x) \leftrightarrow \varphi(y)))$$

for all singulary formulas φ (possibly with parameters). However, for us the truth of all of these sentences is a *theorem,* which can be established by induction in the algebra of formulas.

For our purposes, we introduce two more axioms:

Axiom (Empty Set)**.** *The empty class is a set.*

Axiom (Adjunction)**.** *For all sets a and b, there is a set whose members are b and the members of a.*

The empty set is denoted by 0; the set whose members are b and the members of a is $a \cup \{b\}$. The Empty-Set and Adjunction axioms together

can be understood as saying simply that every *finite* collection of sets is a set. In particular,

$$0, \qquad \{0\}, \qquad \{0, \{0\}\}, \qquad \{0, \{0\}, \{0, \{0\}\}\},$$

and so forth are sets, each being the collection of all previously known sets. This sequence is 0, $0'$, $0''$, $0'''$, and so forth, where in general $x' = x \cup \{x\}$. We may call x' the **successor** of x. We now define 1 as $0'$, and we define $\mathbb{N}$ as the smallest collection of sets that contains 1 and is closed under succession. However, we should be able to *prove* that the three Peano Axioms are satisfied. The Axiom of Induction is satisfied by definition of $\mathbb{N}$. Then we have the following lemma, perhaps the simplest application of this axiom.

Lemma. *Every element of* $\mathbb{N}$ *is the successor of some set.*

Proof. 1 is the successor of 0, and every successor of a successor is in particular a successor of some set. □

Theorem 4. 1 *is not the successor of any element of* $\mathbb{N}$.

Proof. If $1 = a'$, this means $\{0\} = a \cup \{a\}$, so in particular $a = 0$. But 0 is not in $\mathbb{N}$, by the lemma. □

It is not so easy to prove that succession is injective on $\mathbb{N}$. We want to show that every element of $\mathbb{N}$ is a' for some *unique* set a that is either 0 or an element of $\mathbb{N}$. By the theorem, this is true when $a = 1$. Suppose it is true when $a = b$, but now $b' = c'$, that is, $b \cup \{b\} = c \cup \{c\}$. If $b \neq c$, then we must have $b \in c$ and $c \in b$. This peculiar possibility can be ruled out by another axiom; but there is a better way.

A class $\boldsymbol{C}$ defined by a formula φ is a **subclass** of a class $\boldsymbol{D}$ defined by a formula ψ if $\forall x\, (\varphi(x) \to \psi(x))$; we then use $\boldsymbol{C} \subseteq \boldsymbol{D}$ as an abbreviation of this sentence. Of course a **subset** of a class is a subclass that is a set. We noted (on page 11) that Dedekind confused the membership and subset relations; but now we must carefully distinguish. A class *contains* its elements, but *includes* its subclasses. In particular, the successor of a

set a both contains a and includes a, and it is the smallest set to do so. A class $\boldsymbol{C}$ is **transitive** if it includes all of its elements, that is,

$$\forall x\ (x \in \boldsymbol{C} \to x \subseteq \boldsymbol{C})$$

or $\forall x\ \forall y\ (y \in x \wedge x \in \boldsymbol{C} \to y \in \boldsymbol{C})$. A subclass $\boldsymbol{C}$ of a class $\boldsymbol{D}$ is **proper** if it is not the whole class $\boldsymbol{D}$; in that case we write $\boldsymbol{C} \subset \boldsymbol{D}$. We can now state and prove the following.

Theorem 5. *The class of transitive sets contains* 0 *and is closed under succession, and succession is injective on this class.*

Proof. Trivially 0 is transitive. Suppose a is transitive and $b \in a \cup \{a\}$. Then either $b \in a$ or $b = a$, and in each case $b \subseteq a$, so $b \subseteq a'$. Thus a' is transitive.

Finally suppose a and b are distinct transitive sets, and $a' \subseteq b'$. Then $a \in b'$, so $a \in b$ (since $a \neq b$), hence $a \subseteq b$, and then $a \subset b$. Therefore b is not a subset of a (since otherwise a would be a proper subset of itself), so $b \notin a$. Thus $b \notin a'$, so $b' \neq a'$. □

Corollary. *Succession is injective on* $\mathbb{N}$.

Proof. Containing 0 and being closed under succession, the class of transitive sets also contains 1; but by definition $\mathbb{N}$ is included in every collection, and in particular every class, that contains 1 and is closed under succession. □

Thus, instead of obtaining $\mathbb{N}$ as a subset of $\mathbb{R}$, or just assuming it exists so as to satisfy the Peano axioms, we can construct $\mathbb{N}$, satisfying the Peano axioms, on the basis of three simple axioms about sets.

This construction does not give us $\mathbb{N}$ as a set. We may promulgate the Axiom of Infinity, whereby $\mathbb{N}$ or rather $\{0\} \cup \mathbb{N}$ is a set by fiat. But what does this mean? We may state as an axiom that *some* set contains 0 and is closed under succession; we may even state that there is a smallest such set; but we cannot even conclude that there is a smallest

such *class* without something like the following axiom or rather scheme of axioms.[17]

Axiom Scheme (Separation)**.** *Every subclass of a set is a set.*

If we assume now there is *one* set that contains 0 and is closed under succession, then the intersection of the class of all such sets is a set, called ω; and no proper subclass of this has the same closure properties. But even without the assumption, we can obtain ω as a class. One way to do this is as follows. A class $\boldsymbol{C}$ is **well-founded** if each of its sub*sets* has an element that is disjoint from that subset, that is,

$$\forall x\,(x \subseteq \boldsymbol{C} \to \exists y\,(y \in x \wedge \forall z\,(z \in y \to z \notin x))).$$

A *set* whose every member is the successor of a member is not well-founded. Now we define ω as the class of all *transitive, well-founded* sets a such that

1) every nonempty member of a is the successor of a member of a, and
2) if a is not empty, it has a member whose successor is *not* in a.

Then $0 \in \bigcup\omega$, and with a bit of work, $\bigcup\omega$ is closed under succession. By the Separation axioms, $\bigcup\omega$ is the *smallest* class that contains 0 and is closed under succession. Also $\bigcup\omega = \omega$.

Note why the members of ω must be required to be well-founded. Having $\mathbb{N}$ informally as above, we may imagine a set $\{a_k\colon k \in \mathbb{N}\}$, where ${a_{k+1}}' = a_k$ in each case, and also the function $k \mapsto a_k$ is injective. This set meets the two numbered criteria for being in ω, but is not well-founded.

However, possibly $\{0\} \cup \mathbb{N} \cup \{a_k\colon k \in \mathbb{N}\}$ is a set, but its every infinite subclass contains an element of $\{0\} \cup \mathbb{N}$. Then the set is well-founded. It might even be transitive, if $a_k = \{0\} \cup \mathbb{N} \cup \{a_n\colon k < n\}$ in each case. We have no formal way prevent such a set from belonging to ω. This is what troubled Skolem [24]: set theory is not categorical, but can have a non-standard model, in which the class defined by the formula for ω is actually larger than intended.

[17] The theory with the Equality, Empty-Set, Adjunction, and Separation axioms is called General Set Theory by Boolos [1, p. 196], but is called STZ by Burgess [2, p. 223], for **S**zmielev and **T**arski with **Z**ermelo's Axiom of Separation.

We may assume that there is a standard model of set theory in which ω really is $\{0\} \cup \mathbb{N}$; but a formalization of this would be a literally infinite statement, and we can quote such a statement only in part:

ω contains 0 and $0'$ and $0''$ and..., and nothing else.

Suppose we nonetheless believe that there is still a standard model of set theory, a model in which ω really is $\mathbb{N} \cup \{0\}$. In that case, the formal sentences that are true in this model compose a collection. Then, using the idea of Gödel [12],[18] we can assign to each such formula a *code,* which is a certain set. Although we could avoid doing so, it will be convenient now to introduce 0 as an official constant, and $'$ as an official function-symbol. We number the symbols used in our formulas, as for example in the following scheme:

$\rightarrow$	$\neg$	$\exists$	(	)	$\in$	0	$'$	x	y	z	...
0	1	2	3	4	5	6	7	8	9	10	...

Then the **code** of a formula φ is a finite sequence $\ulcorner\varphi\urcorner$ of elements of ω. For example, $\ulcorner\exists x \neg 0' \in x\urcorner$ is $(2, 8, 1, 6, 7, 5, 8)$. Such a sequence can be understood as a function from some element of ω to ω; in particular, it is a certain finite *set* of ordered pairs. This is not a formal theorem of set theory; it is just the observation that, by the Empty-Set and Adjunction axioms, for every ordered pair (a, b) of sets, there is a *set* $\{\{a\}, \{a, b\}\}$, which (as Kuratowski showed in 1921 [17]) can be identified with the ordered pair; and then every finite collection of ordered pairs can be built up one by one into a set.

If a formula φ has length n, then $\ulcorner\varphi\urcorner$ is an element of the class denoted by ω^n. Thus the codes of all formulas belong to the class that can be denoted by $\bigcup_{n\in\omega} \omega^n$ or $\omega^{<\omega}$. If the collection of codes of all true sentences of set theory were itself a class, then there would be a singulary formula ϑ defining the class of codes of all singulary formulas φ such that $\varphi(\ulcorner\varphi\urcorner)$ is false. In this case,

$$\vartheta(\ulcorner\vartheta\urcorner) \leftrightarrow \neg\vartheta(\ulcorner\vartheta\urcorner),$$

which is absurd. This is a variant of the Russell Paradox, known as something like Tarski's Theorem on the Undefinability of Truth. In stating his

[18] See note 16 on page 22.

version of the theorem, Tarski [25, p. 247] notes his debt to Gödel. The point now is that, if $\mathbb{N}$ is a set or even just a class, then the collection of codes of true sentences of set theory is not a class. In short, some mathematical collections are definitely not classes.

In the present context, a form of Gödel's Incompleteness Theorem can be established as follows. The collection of codes of formally provable sentences is a class, at least on the assumption that $\mathbb{N}$ is a class. Then there is a class consisting of the codes of all singulary formulas φ such that $\neg\varphi(\ulcorner\varphi\urcorner)$ has a formal proof. This last class is defined by a formula ψ. Then $\psi(\ulcorner\psi\urcorner)$ is true if and only if the code of its negation is in the class of codes of provable sentences. If this class never contains the codes of both a sentence and its negation, our set theory is **consistent**; but in this case neither $\psi(\ulcorner\psi\urcorner)$ nor its negation is provable, although $\neg\psi(\ulcorner\psi\urcorner)$ is true.

If we do not want to assume that $\mathbb{N}$ is a class, we can still argue as follows. There is a smallest class $\boldsymbol{C}$ comprising the codes of provable sentences. Let ψ define the class of codes $\ulcorner\varphi\urcorner$ such that $\neg\varphi(\ulcorner\varphi\urcorner)$ is in $\boldsymbol{C}$. If $\psi(\ulcorner\psi\urcorner)$ had a formal proof, then it would be true, and so both $\psi(\ulcorner\psi\urcorner)$ and $\neg\psi(\ulcorner\psi\urcorner)$ would be in $\boldsymbol{C}$. If $\neg\psi(\ulcorner\psi\urcorner)$ had a formal proof, then both it and $\psi(\ulcorner\psi\urcorner)$ would be true.

References

[1] George Boolos, *Logic, logic, and logic*, Harvard University Press, Cambridge, MA, 1998, With introductions and an afterword by John P. Burgess, With a preface by Burgess and Richard Jeffrey, Edited by Jeffrey. MR 1675856 (2000b:03005)

[2] John P. Burgess, *Fixing Frege*, Princeton Monographs in Philosophy, Princeton University Press, Princeton, NJ, 2005. MR MR2157847 (2006e:03006)

[3] Stanley N. Burris, *Logic for mathematics and computer science*, Prentice Hall, Upper Saddle River, New Jersey, USA, 1998.

[4] Alonzo Church, *Introduction to mathematical logic. Vol. I*, Princeton University Press, Princeton, N. J., 1956. MR 18,631a

[5] Richard Dedekind, *Was sind und was sollen die Zahlen?*, Friedrich Vieweg, Braunschweig, 1893.

[6] Richard Dedekind, *Essays on the theory of numbers. I: Continuity and irrational numbers. II: The nature and meaning of numbers*, authorized translation by Wooster Woodruff Beman, Dover Publications Inc., New York, 1963. MR MR0159773 (28 #2989)

[7] John Dyer-Bennet, *A theorem on partitions of the set of positive integers*, Amer. Math. Monthly **47** (1940), 152–154. MR MR0001234 (1,201b)

[8] Herbert B. Enderton, *A mathematical introduction to logic*, second ed., Harcourt/Academic Press, Burlington, MA, 2001. MR 1801397 (2001h:03001)

[9] Euclid, *Euclid's Elements*, Green Lion Press, Santa Fe, NM, 2002, All thirteen books complete in one volume, the Thomas L. Heath translation, edited by Dana Densmore. MR MR1932864 (2003j:01044)

[10] Carl Friedrich Gauss, *Disquisitiones arithmeticae*, Springer-Verlag, New York, 1986, Translated into English by Arthur A. Clarke, revised by William C. Waterhouse.

[11] Kurt Gödel, *The completeness of the axioms of the functional calculus of logic (1930a)*, From Frege to Gödel (Jean van Heijenoort, ed.), Harvard University Press, 1976, pp. 582–91.

[12] ———, *On formally undecidable propositions of* principia mathematica *and related systems I (1931)*, From Frege to Gödel (Jean van Heijenoort, ed.), Harvard University Press, 1976, pp. 596–616.

[13] Paul R. Halmos, *I want to be a mathematician*, MAA Spectrum, Mathematical Association of America, Washington, DC, 1985, An automathography in three parts. MR 961440 (89e:01044)

[14] Leon Henkin, *The completeness of the first-order functional calculus*, J. Symbolic Logic **14** (1949), 159–166. MR MR0033781 (11,487d)

[15] ———, *On mathematical induction*, Amer. Math. Monthly **67** (1960), 323–338. MR MR0120156 (22 #10913)

[16] Wilfrid Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications, vol. 42, Cambridge University Press, Cambridge, 1993. MR 94e:03002

[17] Casimir Kuratowski, *Sur la notion d'ordre dans la théorie des ensembles*, Fundamenta Mathematicae (1921), 161–71.

[18] Edmund Landau, *Foundations of analysis. The arithmetic of whole, rational, irrational and complex numbers*, third ed., Chelsea Publishing Company, New York, N.Y., 1966, translated by F. Steinhardt; first edition 1951; first German publication, 1929. MR 12,397m

[19] Serge Lang, *Algebra*, third ed., Addison-Wesley, Reading, Massachusetts, 1993, reprinted with corrections, 1997.

[20] Saunders Mac Lane and Garrett Birkhoff, *Algebra*, third ed., Chelsea Publishing Co., New York, 1988. MR MR941522 (89i:00009)

[21] Barry Mazur, *How did Theaetetus prove his theorem?*, `http://www.math.harvard.edu/~mazur/preprints/Eva.pdf`, accessed Sept. 24, 2008.

[22] Giuseppe Peano, *The principles of arithmetic, presented by a new method (1889)*, From Frege to Gödel (Jean van Heijenoort, ed.), Harvard University Press, 1976, pp. 83–97.

[23] Bertrand Russell, *Letter to Frege (1902)*, From Frege to Gödel (Jean van Heijenoort, ed.), Harvard University Press, 1976, pp. 124–5.

[24] Thoralf Skolem, *Some remarks on axiomatized set theory (1922)*, From Frege to Gödel (Jean van Heijenoort, ed.), Harvard University Press, 1976, pp. 290–301.

[25] Alfred Tarski, *Logic, semantics, metamathematics*, second ed., Hackett Publishing Co., Indianapolis, IN, 1983, Papers from 1923 to 1938, Translated by J. H. Woodger, Edited and with an introduction by John Corcoran. MR 736686 (85e:01065)

[26] Alfred North Whitehead and Bertrand Russell, *Principia mathematica*, vol. I, University Press, Cambridge, 1910.

[27] Ernst Zermelo, *Investigations in the foundations of set theory I (1908a)*, From Frege to Gödel (Jean van Heijenoort, ed.), Harvard University Press, 1976, pp. 199–215.