Outline

Edward Nelson
Department of Mathematics
Princeton University

Against finitism. The induction axiom schema of Peano arithmetic P is usually justified as follows. Assume the basis $A_x(0)$ and the induction step $\forall x[A \to A_x(Sx)]$. For any numeral d, a special case of the induction step is $A_x(d) \to A_x(Sd)$. Then we can prove $A_x(d)$, starting with the basis and using detachment (modus ponens) as many times as there are occurrences of S in d.

This justification leaves much to be desired. The argument applies only to numerals, and shows that there is no need at all to postulate induction for numerals. By an unspoken conflation of the concrete concept of numeral with the formal concept of number, induction is postulated for numbers.

Finitism is usually regarded as the most conservative standpoint for the foundations of mathematics. Induction is justified by appeal to the *finitary credo*: for every number x there exists a numeral d such that x is d. It is necessary to make this precise. We cannot express it as a formula of arithmetic because "there exists" in "there exists a numeral d" is a metamathematical existence assertion, not an arithmetical formula beginning with $\exists$.

The finitary credo can be formulated precisely using the concept of the standard model of arithmetic: for every element ξ of $\mathbb{N}$ there exists a numeral d such that it can be proved that d is equal to the name of ξ, but this brings us into set theory. The finitary credo has an infinitary foundation.

The use of induction goes far beyond the application to numerals. It is used to create new kinds of numbers (exponential, superexponential, and so forth) in the belief that they already exist in a completed infinity. If there were a completed infinity $\mathbb{N}$ consisting of all numbers, then the axioms of P would be valid assertions about numbers and P would be consistent.

It is not a priori obvious that P can express combinatorics, but this is well known thanks to Gödel's great paper [Gö31] on incompleteness. As a consequence, exponentiation $\uparrow$ and superexponentiation $\Uparrow$ can be defined in P so that we have

(1) $x \uparrow 0 = S0$

(2) $x \uparrow Sy = x \cdot (x \uparrow y)$

(3) $x \Uparrow 0 = S0$

(4) $x \Uparrow Sy = x \uparrow (x \Uparrow y)$

and similarly for primitive-recursive functions in general. For the schema of primitive recursion to be coherent, it is necessary that the values of the functions always reduce to numerals, since they are defined only for 0 and iterated successors of numbers for which they have been previously defined.

Finitists believe that primitive recursions always terminate; for example, that applying (1)–(4) (together with the open axioms of P) a sufficient number of times,

$$SS0 \Uparrow SS0 \Uparrow SS0 \Uparrow SS0 \Uparrow SS0 \Uparrow SS0 \Uparrow SS0 \Uparrow SS0 \Uparrow SS0 \Uparrow SS0 \Uparrow SS0 \Uparrow SS0 \Uparrow SS0 \Uparrow SS0 \Uparrow SS0$$

reduces to a numeral. But the putative number of times these rules must be applied can only be expressed by means of a superexponential expression—the argument is circular.

The problem is not simply that some primitive recursions are too long. The problem is structural: there is a sharp division between two classes of primitive recursions. This results from work of Bellantoni and Cook [BeCo]; see also Bellantoni's thesis [Be]. They indicate that their work was strongly influenced by previous work of Leivant, who subsequently [Le] gave a more elegant form to the characterization, and it is appropriate to call the result the BCL theorem.

While a number is being constructed by recursion, it is only potential, and when the recursion is complete, it is actual. What Bellantoni and Cook, and Leivant, do is restrict recursions so that they occur only on actual numbers. Then the theorem is that the class of functions computable in this way is precisely the class of polynomial-time functions. This is an astonishing result, since the characterization is qualitative and conceptually motivated, saying nothing whatever about polynomials.

Bellantoni, Cook, and Leivant have revealed a profound difference between polynomial-time recursions and all other recursions. The recursions constructed by the BCL schema enjoy a different ontological status from recursions in general. In the former, recursions are performed only on objects that have already been constructed. In the latter, for example in a superexponential recursion, one counts chickens before they are hatched (and the chicks that they produce as well).

Not only is induction as a general axiom schema lacking any justification other than an appeal to $\mathbb{N}$ as a completed infinity, but its application to specific variable-free primitive-recursive terms lacks a cogent justification.

I shall exhibit a superexponential recursion and prove that it does not terminate, thus disproving Church's Thesis from below and demonstrating that finitism is untenable.

Relativization. Robinson's theory [rRo] is the theory whose nonlogical axioms are those of P without the induction axiom schema but with Robinson's axiom

$$x \neq 0 \to \exists y[x = \mathrm{S}y]$$

It may be reformulated as an open theory Q_0 (Robinson arithmetic) by introducing a unary function symbol P (*predecessor*) and replacing Robinson's axiom by

$$\mathrm{P}0 = 0$$

$$x \neq 0 \to x = \mathrm{SP}x$$

The theory Q_0 is quite weak, but by relativization schemata of Solovay [Slv] (see also [Ne]) it can be greatly strengthened by relativization. We construct a theory schema Q_0^* that is locally relativizable in Q_0 and is such that we can use induction on bounded formulas, where "bounded" means that there is a polynomial bound on the length. It is striking that exponentiation lies just the other side of the BCL division of recursive functions into two classes, and just the other side of the class of functions that can be proved total by relativization.

Remarkably, Q_0 proves the quasitautological consistency of its arithmetization, as was proved in [Ne]. That is,

(5) $\vdash_{\mathsf{Q}_0^*} \ulcorner \mathsf{Q}_0 \urcorner$ is quasitautologically consistent

which is equivalent, using the terminology and notation of [Sh §4.3], to

(6) $\vdash_{\mathsf{Q}_0^*} \neg\exists s[s$ encodes a special sequence in $\Delta_0(\ulcorner \mathsf{Q}_0 \urcorner)]$

The Hilbert-Ackermann consistency theorem. This theorem states that a quasitautologically consistent open theory is consistent. As the proof is presented in [Sh §4.3], it is an algorithm for elimination of special constants in a special sequence. Call a special sequence in $\Delta(\mathrm{T})$, where T is an open theory, a (ρ, λ, κ)-*special sequence* in case the maximal rank of the special constants in it is at most ρ, their maximal level is at most λ, and the maximal level of those special constants of rank ρ is at most κ (so $\kappa \leq \lambda$), and call it a (ρ, λ)-*special sequence* in case it is a (ρ, λ, κ)-special sequence for some κ. The key step in the proof of the consistency theorem is Lemma 2 of [Sh §4.3]. This lemma eliminates one special constant of maximal level κ among those of maximal rank, but the proof goes through unchanged if all such special constants are eliminated together. The result is that a (ρ, λ, κ)-special sequence with $\kappa > 0$ can be replaced by a $(\rho, 2 \cdot \lambda - \kappa, \kappa - 1)$-special sequence, with a quadratic bound on the increase in length of the special sequence. This result can be arithmetized as a theorem (*) of Q_0^*. The full proof of the consistency theorem involves iterating Shoenfield's Lemma 2 κ times, leading to a $(\rho - 1, \lambda^{2^\lambda}, \lambda^{2^\lambda})$-special sequence, and then iterating this ρ times. This cannot be arithmetized in Q_0^*.

But suppose that ρ and λ are given specific numbers. Let us see what follows by finitary reasoning. It must be emphasized that we are not adjoining any finitary assumptions to Q_0^*; we are only seeing what finitary reasoning says about what can be proved in Q_0^*.

Refer to Lemma 1 of [Sh §4.3] and call π a (ρ, λ)-*proof* of A in T in case π is a sequence of formula in $\Delta(\mathrm{T})$ in which the maximal rank of the special constants is at most ρ and their maximal level is at most λ, and the closure of A is a tautological consequence of them. Let $\bar{\xi}$ be the numeral S...S0 with ξ occurrences of S. According to finitism, for any ρ and λ

(7) $\vdash_{\mathsf{Q}_0^*} \neg\exists p[p$ encodes a $(\bar{\rho}, \bar{\lambda})$-proof of $\ulcorner 0 \neq 0 \urcorner]$

i.e.,

(8) $\vdash_{\mathsf{Q}_0^*} \neg\exists s[s$ encodes a $(\bar{\rho}, \bar{\lambda})$-special sequence in $\ulcorner \mathsf{Q}_0 \urcorner]$

For suppose that there is such an encoded special sequence. It is an encoded $(\bar{\rho}, \bar{\lambda}, \bar{\kappa})$-special sequence for some κ with $\kappa \leq \lambda$, so by the arithmetized theorem (*), it is a theorem of Q_0^* that there is an encoded $(\bar{\lambda}, 2 \cdot \bar{\lambda} - \bar{\kappa}, \bar{\kappa} - 1)$-special sequence. Applying this theorem κ times, conclude that Q_0^* proves that there is an encoded $(\bar{\rho} - 1, \bar{\lambda}_1, \bar{\lambda}_1)$-special sequence, where $\lambda_1 = \lambda^{2^\lambda}$. Now apply this result ρ times, and conclude (by finitary reasoning) that Q_0^* proves that there is an encoded special sequence in $\Delta_0(\ulcorner \mathsf{Q}_0 \urcorner)$, contradicting (6).

Incompleteness without diagonalization. The *Kolmogorov complexity* of a number ξ, denoted by $\mathcal{K}(\xi)$, is the length (number of bits in the program) of the shortest Turing machine that halts and outputs ξ. The notion was introduced by Solomonoff [Slm] and then independently by Kolmogorov [Ko].

The most important theorem in the subject was proved by Chaitin [Ch]. Let T be a theory capable of arithmetizing itself and expressing combinatorics, such as P or Q_0. Let K be the arithmetization of $\mathcal{K}$. Choose a number λ and construct a machine as follows. It systematically searches through strings, ordered first by length and then lexicographically. If it finds one that is a proof in T of $\mathrm{K}(\bar{\xi}) > \bar{\lambda}$, it outputs ξ and halts. Now λ can be referred to by a term of T of length logarithmic in λ , so for λ large enough the machine is itself of length less than λ. Fix such a λ and denote it by ℓ. Call the corresponding machine the *Chaitin machine for* T.

We have

(9) if $\mathcal{K}(\xi) \leq \ell$, there is a π such that $\pi \vdash_{\mathrm{T}} \mathrm{K}(\bar{\xi}) \leq \bar{\ell}$

simply by verifying that the steps of the Turing machine in question are followed. Chaitin's theorem is

(10) if T is consistent, there do not exist ξ and π such that $\pi \vdash_{\mathrm{T}} \mathrm{K}(\bar{\xi}) > \bar{\ell}$

Otherwise, the Chaitin machine would find a proof in T of some $\mathrm{K}(\bar{\eta}) > \bar{\ell}$ (where η may or may not be the same as ξ) and output η. Then we would have $\mathcal{K}(\eta) \leq \ell$ by the definition of Kolmogorov complexity, giving a contradiction in T by (9). Gödel's first incompleteness theorem is a consequence. There is no diagonalization or self-reference in this proof; Chaitin remarks that it is a version of the Berry paradox.

Kritchman and Raz [KrRa] have given a stunning proof without diagonalization or self-reference of Gödel's second incompleteness theorem.

There are strictly fewer than $2^{\ell+1}$ Turing machines with at most ℓ bits, so by the pigeonhole principle there is at least one ξ with $\xi < 2^{\ell+1}$ such that $\mathcal{K}(\xi) > \ell$. Let μ be the number of such ξ, so

(11) $\mu > 0$

There are $2^{\ell+1}$ days left in the course and the teacher announces that there will be an examination on one of those days, but it will come as a surprise. Think of μ as being the number of days remaining after the surprise examination. The examination is not given on the last day, by (11). Now suppose that $\mu = 1$ (the surprise examination occurs on the penultimate day of classes). Then there is a unique ξ with $\xi < 2^{\ell+1}$ such that $\mathcal{K}(\xi) > \ell$, and T proves $\mathrm{K}(\bar{\eta}) \leq \bar{\ell}$ for every other η with $\eta < 2^{\ell+1}$, by (9). Hence if $\mu = 1$, T proves $\mathrm{K}(\bar{\xi}) > \bar{\ell}$, which is impossible if T is consistent, by Chaitin's theorem. Consequently, if T proves the consistency of its arithmetization then T proves $\bar{\mu} \geq 2$. Now suppose that $\mu = 2$ and argue in the same way, and continue up to $\mu = 2^{\ell+1}$. In this way, if T proves the consistency of its arithmetization, it proves a contradiction, since $\mu \leq 2^{\ell+1}$. This yields Gödel's second incompleteness theorem.

This proof is radically different from Gödel's self-referential proof. The latter derives a contradiction from the consideration of proofs of an entirely unrestricted kind, whereas the former derives a contradiction from the consideration of proofs of quite specific kinds, of bounded complexity. We shall exploit this difference.

A contradiction in finitism. Let us examine what finitism says about what Q_0^* can prove concerning the Kritchman-Raz proof. Consider the proofs π of (9). The formulas

are all of rank at most ρ_0 for a specific ρ_0, and even though the π may be immensely long, there is a fixed bound λ_0 on the level. This is because each proof operates on the given record of the working of the Turing machine, repeatedly citing the same theorems to verify that each step of the computation accords with the machine's program. The proofs of the arithmetizations of (10) and (11) are of a certain specific rank and level, so for certain specific ρ_1 and λ_1

$$\vdash_{\mathsf{Q}_0^*} \bar{\mu} = 1 \to \exists p[p \text{ encodes a } (\bar{\rho}_1, \bar{\lambda}_1)\text{-proof of } \ulcorner 0 \neq 0 \urcorner]$$

By finitary reasoning, from (7) we have

$$\vdash_{\mathsf{Q}_0^*} \bar{\mu} \geq 2$$

The proofs of $\bar{\mu} \geq 3$, $\ldots$, $\bar{\mu} \geq \bar{\ell}_1$, where $\ell_1 = 2^{\ell+1} + 1$, increase in rank and level, but in a manner that can be bounded explicitly. The upshot is that finitism proves that Q_0^*, and hence Q_0, is inconsistent. But there is a well-known finitary argument, based on the Hilbert-Ackermann consistency theorem, showing that Q_0 is consistent. Hence finitary reasoning leads to a contradiction.

Peano arithmetic expresses finitary reasoning and so it is inconsistent. In fact, we have a contradiction in a small fragment of primitive-recursive arithmetic (PRA). How small such a fragment can be will be investigated.

Qea. If this were normal science, the proof that P is inconsistent could be written up rather quickly. But since this work calls for a paradigm shift in mathematics, it is essential that all details be developed fully. At present, I have written just over 100 pages beginning this. The current version is posted as a work in progress at http://www.math.princeton.edu/~nelson/books.html, and the book will be updated from time to time. The proofs are automatically checked by a program I devised called qea (for *quod est absurdum*, since all the proofs are indirect). Most proof checkers require one to trust that the program is correct, something that is notoriously difficult to verify. But qea, from a very concise input, prints out full proofs that a mathematician can quickly check simply by inspection. To date there are 733 axioms, definitions, and theorems, and qea checked the work in 93 seconds of user time, writing to files 23 megabytes of full proofs that are available from hyperlinks in the book.

Consistency of Q_0. Since finitary methods are no longer available, the consistency of Q_0 becomes an open question.

It is hopeless, and unnecessary, to try to prove the consistency of arithmetized Q_0 within a theory, due to the obstacle of Gödel's second theorem. What we can hope to do is grade the proofs in Q_0 by complexity and for each grade prove, in a system we trust, that there is no proof of a contradiction in Q_0 of the given grade. For the trusted system, I take a quantifier-free system with polynomial-time functions, as in the BCL theorem. I shall investigate whether something even simpler will suffice.

The classical Q_0 is equiconsistent with its intuitionistic version, by Gödel's short paper [Gö33]. The logical operators that present a problem are $\forall$ and the intuitionistic $\to$. The intuitionistic semantics of these operators is infinitary. I shall explore a recursive notion

of "verifiability", somewhat similar to Kleene's realizability [Kl §82], and attempt to show that for each grade of complexity nothing verifies $0 \neq 0$.

Modern mathematics. To demonstrate that sophisticated modern mathematics can be done in so primitive a theory as Q_0^* is an open-ended project, and I plan to make a bare beginning in the book. It may result that one can prove in Q_0^* that exponentiation is not total, but as already pointed out in [Ne] it is consistent to adjoin this assumption. This is not so restrictive as it may seem. One can construct iterated relativizations of Q_0^*, leading to worlds of arbitrary depth δ, such that exponentiation is well defined from each depth $\delta + 1$ to δ. (And superexponentiation rarely occurs in mainstream mathematics.) A number that is not exponentiable is like a nonstandard number in Abraham Robinson's nonstandard analysis [aRo]. The material on probability theory in [Ne2] is an example of what can be done in this way, in a kind of (Raphael + Abraham) Robinson mathematics. I plan to develop others.

References

[aRo] Abraham Robinson, *Non-standard Analysis*, American Elsevier, New York, 1966.

[Be] Stephen Bellantoni, *Predicative Recursion and Computational Complexity*, ftp://ftp.cs.toronto.edu/pub/reports/theory/cs-92-264.ps.Z

[BeCo] Stephen Bellantoni and Stephen Cook "A new recursion-theoretic characterization of poly-time functions", *Computational Complexity* 2 (1992), 97-110. http://www.cs.toronto.edu/~sacook/homepage/ptime.pdf

[Ch] G. J. Chaitin, "Computational complexity and Gödel's incompleteness theorem", *ACM SIGACT News* 9, pp. 11-12, 1971.

[Gö31] Kurt Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwanter Systeme", *Monatshefte für Mathematik und Physik* 38 (1931) 173-198.

[Gö33] Kurt Gödel, "Zur intuitionistischen Aritmetik und Zahlentheorie", *Ergebnisse eines mathematischen Kolloquiums* 4 (1933) 34-38.

[Le] Daniel Leivant, "Ramified recurrence and computational complexity I: Word recurrence and poly-time", in P. Cole and J. Remmel, eds., *Feasible Mathematics II*, Perspectives in Computer Science, Birkhauser-Boston, (1994) 320-343.

[Kl] S. C. Kleene, *Introduction to Metamathematics*, North-Holland, New York, 1952.

[Ko] A. N. Kolmogorov, "Three approaches to the quantitative definition of information", *Problems of Information Transmission* 1 (1965) 1-7.

[KrRa] Shira Kritchman and Ran Raz, "The surprise examination paradox and the second incompleteness theorem", *Notices of the AMS* 57 (2010) 1454-1458.
www.ams.org/notices/201011/rtx101101454p.pdf

[Ne] Edward Nelson, *Predicative Arithmetic*, Mathematical Notes 32, Princeton University Press, 1986. http://math.princeton.edu/~nelson/books/pa.pdf

[Ne2] Edward Nelson, *Radically Elementary Probability Theory*, Annals of Mathematics Studies 117, Princeton University Press, 1987.
http://math.princeton.edu/~nelson/books/rept.pdf

[rRo] Raphael M. Robinson, "An essentially undecidable system", *Proceedings of the International Congress of Mathematicians* 1, Cambridge, Massachusetts, (1950) 729-730.

[Sh] Joseph R. Shoenfield, *Mathematical Logic*, Association for Symbolic Logic, A K Peters, Ltd., Natick, Massachusetts, 1967.

[Slm] R. J. Solomonoff, "A preliminary report on a general theory of inductive inference", *Report V-131*, Zator Co., Cambridge, Massachusetts, Feb. 1960, revised Nov. 1960.
http://world.std.com/~rjs/z138.pdf

[Slv] Robert M. Solovay, Letter to P. Hájek, August 1976. Cited by Samuel R. Buss in *Diffusion, Quantum Theory, and Radically Elementary Mathematics*, William G. Faris, ed., Mathematical Notes 47, Princeton University Press, Princeton, New Jersey, 2006.