Synthese (2007) 154:401–415
DOI 10.1007/s11229-006-9126-9

ORIGINAL PAPER

Informal versus formal mathematics

Francisco Antonio Doria

Published online: 25 January 2007

Abstract We discuss Kunen's algorithmic implementation of a proof for the Paris–Harrington theorem, and the author's and da Costa's proposed "exotic" formulation for the $P = NP$ hypothesis. Out of those two examples we ponder the relation between mathematics within an axiomatic framework, and intuitive or informal mathematics.

Keywords Informal mathematics · Formalized mathematics · Consistency of arithmetic · P vs. NP

1 Introduction

We will be interested here in the differences between arguments that are made within formal mathematics, and their informal counterparts.

- By *formal mathematics* we understand mathematics done within axiomatic systems, based on some formal language, and with proofs that follow the usual practice in logic. (For an example see Scholium 3.22 in this paper.)
- *Informal mathematics* is mathematics done with the usual standards of rigor, sometimes with verbose arguments. The usual example is the exposition style in Rogers' textbook (Rogers, 1967); in the present paper, informal arguments are presented in Proposition 3.6 and in Remark 3.33.

We sometimes refer to "intuition," "intuitive." This is only to be taken as a synonym for "naïve," without any specific philosophical intent.

The author is Visiting Researcher at IEA/USP, Professor of Communications, Emeritus, at the Federal University in Rio de Janeiro, and a full member of the Brazilian Academy of Philosophy.

F. A. Doria (✉)
Institute for Advanced Studies, University of São Paulo,
Av. Prof. Luciano Gualberto, trav. J, 374,
São Paulo 05655-010,
SP, Brazil
e-mail: fadoria@gmail.com

The questions we wish to consider and illustrate are: how different are the formal and informal approaches? Can we reconcile them?

We first present a short discussion of Kunen's algorithmic implementation (Kunen, 1995) of a proof of the Paris–Harrington theorem (or more pedantically, Ramsey–Paris–Harrington theorem, to emphasize its origin). We then present and elaborate on the author's proposal (with N. C. A. da Costa) of an "exotic" formalization of the $P = NP$ hypothesis (Costa & Doria, 2003) in Zermelo–Fraenkel set theory. Our formalization satisfies some previously given (DeMillo & Lipton, 1979, 1980; Joseph & Young, 1981; Kowalczyk 1982) conditions for a consistency result, but for an unexpected twist, so to say.

Kunen's construction (Kunen, 1995) looks deceptively naïve: Kunen describes an algorithmic theorem prover, the Boyer–Moore prover, and then uses it to verify the Paris–Harrington (PH) theorem (Paris & Harrington, 1977). One immediately jumps and asks: what is going on here? Does that mean that there is an algorithm that proves the consistency of Peano Arithmetic, as the Paris–Harrington theorem implies that consistency? Then, how is that possible?

A different conundrum appears in our second example (Costa & Doria, 2003). Here the first difficulty appears when we go from intuitive to formal, as there apparently are several nonequivalent formal versions of the same intuitive formulation for the $P = NP$ hypothesis. We pick up two of those, the "standard" and "exotic" ones. Both are informally equivalent, but may not be so within even a rich, strong axiomatic framework such as Zermelo–Fraenkel set theory plus the Axiom of Choice; we only get a weakened equivalence result with respect to ZFC. Namely, we get that the formal sentence that states the equivalence between the exotic version $[P = NP]^{\mathsf{F}}$ and the standard one, $[P = NP]$, turns out to be consistent with ZFC, supposed consistent, and true of standard arithmetic. Moreover, $[P = NP]^{\mathsf{F}}$ itself is consistent with ZFC.

What are we to make out of that? On which grounds can we claim that this and similar results about the exotic formulation are legitimate results, that further our knowledge about the P versus NP question?

1.1 Background

This work is part of an ongoing joint research program with N. C. A. da Costa, friend and mentor, on computational complexity. Main results in this paper are the product of that joint effort, while flaws are the author's responsibility. An approach to the P versus NP question along similar lines is reviewed in Ben David and Halevi (Ben-David & Halevi, 1991); however, they do not show specific consistency results as we do in (Costa & Doria, 2003) and here, and as they are presented in the papers that precede our work. Papers which offer results in the same direction as the ones described here are (DeMillo & Lipton, 1979, 1980; Fortune, Leivant, & O'Donnell, 1983; Joseph & Young, 1980, 1981). A full discussion will appear in (Costa, Doria, & Bir, 2007). The author owes to E. Bir for a detailed review of the literature on P versus NP and metamathematics.

1.2 A note

We must certainly quote Kreisel's remarkable essay on informal rigor (Kreisel, 1969); however, as stated above, our goals are immensely more modest. Also we notice Suppes' observations on naïve set theory (Suppes, 2002, p. 30):

> The only important distinction between axiomatic and naive set theory is that in axiomatic set theory one continually must consider questions about the existence of sets, since it is precisely around such questions that problems of paradox and inconsistency arise. In naive set theory one proceeds as if there were no paradoxes and blithely assumes that all sets exist in the intended intuitive fashion.

However, our discussion will be restricted to the questions made explicit above.

2 The Ramsey–Paris–Harrington–Kunen theorem

Since the Ramsey–Paris–Harrington theorem is a well-known and much discussed result, we will just summarize its main features out of the reference (Smoryński, 1985). The Ramsey theorem is an elaboration on the Pigeonhole Principle:

> If we have k balls and n boxes, $k > n$, and distribute the balls in the boxes, there will be a box with more than one ball. □

For its infinite version:

> If we partition the set of natural number into n disjoint subsets, then at least one subset in the partition must be infinite. □

(For their counterpart in the Kunen paper, see Kunen (Kunen, 1995), Lemmas 5 and 6.) Now let us be given a set X of integers and let $[X]^n$ be the set of all n-element subsets of X. If we understand integer c as $c = \{0, 1, 2, \dots, c-1\}$, then a colouring of $[X]^n$ is a map $C\colon [X]^n \to c$.

Remark 2.1 The *Infinite Ramsey Theorem* is: given n, c, positive integers,

> For any colouring $C\colon [X]^n \to c$ there is an infinite $Y \subset \omega$ so that C is constant in Y. □

Remark 2.2 The *Finite Ramsey Theorem* is: given positive integers s, n, c, with $s \geq n + 1$, then there is a number $R(s, n, c)$ so that:

> For all $r \geq R(s, n, c)$ and for all colourings $C\colon [r]^n \to c$ there is a set $Y \subseteq r$ of cardinality s so that C is constant in Y. □

Remark 2.3 For the *Paris–Harrington theorem* one requires a single modification: we say that a $X \subset \omega$ is large if the cardinality of $X \geq \min X$, that is, X has at least as many elements as its smallest element. Then, given positive integers s, n, c, $s \geq n + 1$, there is a number $H(s, n, c)$ so that:

> For all $h \geq H(s, n, c)$ and all colourings $C\colon [h]^n \to c$, there is a large $Y \subset h$ of cardinality at least s so that C is constant on Y. □

Let's now take a look at those successively more complex statements. The two Pigeonhole Principles are trivial and highly intuitive. The infinite Ramsey theorem, less so. Smoryński notices (see Smoryński, 1985) that the diagonal function $R(x + 1, x, x)$ from the finite Ramsey theorem is bounded by a function that grows as F_3 in the Kreisel–Wainer hierarchy (Kreisel, 1951, 1952; Wainer, 1970). So, it can be proved in Primitive Recursive Arithmetic (PRA).

As it is well-known, the diagonal function $H(x + 1, x, x)$ grows about the same as F_{ϵ_0}, that is, Peano Arithmetic (PA) doesn't prove the Paris–Harrington theorem.

We notice two things about the preceding statements, in Remarks 2.1–2.3:

- Each one can be formalized as a Π^0_2 sentence, that is, they fit neatly and precisely within a particular slot in the arithmetic hierarchy.
- Statements in Remarks 2.2 and 2.3 are each equivalent to the existence of a fast growing, total recursive function.

2.1 The Kunen theorem

In his paper (Kunen, 1995), Kunen starts from a simplified proof of the Paris–Harrington theorem. That proof can be fully implemented in PRA but for a single inductive step (Lemma 14) which asserts that, for each ordinal $\alpha < \epsilon_0$ there exists a certain set parametrized by α. The fact that one uses constructive ordinals, as it is well-known, doesn't mean that we have directly introduced infinite quantities in our reasoning—actually we are dealing with a recursive notation system (see Rogers, 1967) as we handle them with the help of ordinal notation systems, which essentially turn out to be quite simple recursive functions for the $\alpha < \epsilon_0$.

The point here is: Kunen's algorithm establishes a proof for the Paris–Harrington theorem. We know that Peano Arithmetic proves the implication:

$$[\text{Paris–Harrington}] \rightarrow [\text{PA is consistent}].$$

Then: does that mean that we have an algorithmic proof for the consistency of PA? Kunen in fact comments in his paper that "it would be interesting to implement [Gentzen's proof of the consistency of PA] on Nqthm," that is, using the Boyer–Moore algorithmic procedures. But isn't that in contradiction with the fact that Peano Arithmetic cannot prove its own consistency?

No. The algorithm works—we can run it in the real world of real, concrete computers, and see that it stops and outputs Paris–Harrington, from which result one derives the consistency of PA. However, PA cannot prove that the algorithm converges.

We are going to elaborate on the consequences of this fact in Sect. 4.

3 Exotic versus standard definitions in computational complexity

The main result we exhibit in this section can be best summarized by the following; for the exotic formulation $[P < NP]^{\mathsf{F}}$ one has:

$$\text{ZFC} \vdash [P < NP]^{\mathsf{F}} \rightarrow [\text{ZFC is consistent}].$$

(Cf. the Paris–Harrington example we just discussed.)

Our axiomatic framework is Zermelo–Fraenkel set theory with the Axiom of Choice (ZFC) plus tools that allow for the introduction of constants. We present in this section both the "standard" and "exotic" formalizations for $P = NP$ and its negation $P < NP$ (standard); $[P = NP]^{\mathsf{f}}$ and its negation $[P < NP]^{\mathsf{f}}$ (exotic), f an adequate strictly increasing total recursive function whose meaning will later be clarified. Sat is the set (coded as usual through binary words) of all satisfiable Boolean expressions in cnf (Machtey & Young, 1979). We recall that Sat $\subset \omega$ is a primitive recursive subset of the set ω of all positive integers. Thus we can code Sat by ω through a primitive recursive coding, which is supposed here.

Developments in this section are based on our work (Costa & Doria, 2003) which benefited from many exacting criticisms and suggestions by Marcel Guillaume (Guillaume, 2000–2002). We can informally state $P = NP$ as:

> There is a time-polynomial Turing machine M_m that correctly "guesses" a satisfying line of truth-values for every input $x \in$ SAT.

That intuitive formulation can be made stricter in several different ways.

Remark 3.1 For example, the *intuitive version of the standard formalization*:

> There is a Turing machine M_m of Gödel number m, and there are positive integers a, b so that for every $x \in$ SAT, the output $\mathsf{M}_m(x)$ is a satisfying line for x, and the number of cycles of M_m over x, $t_m(x) \leq |x|^a + b$. □

Most mathematicians would accept the statement in Remark 3.1 as a legitimate formal definition. Yet we have to go a step further to reach the rigorous presentation of the standard formalization within ZFC:

Definition 3.2 (Standard formalization for $P = NP$.)
$[P = NP] \leftrightarrow_{\text{Def}} \exists m, a, b \in \omega \, \forall x \in \omega \, [(t_m(x) \leq |x|^a + b) \wedge R(x, m)]$. □

$R(x, y)$ is a polynomial predicate; as its interpretation we can say that it formalizes a kind of "verifying machine" that checks whether or not x is satisfied by the output of Turing machine m. Then:

Definition 3.3 $[P < NP] \leftrightarrow_{\text{Def}} \neg[P = NP]$. □

Remark 3.4 $[P < NP]$ is a Π^0_2 sentence. We now impose that all formulations we deal here for the $P < NP$ hypothesis are given as Π^0_2 sentences. The reason is: we are going to explore the connection between Π^0_2 sentences and fast-growing total recursive functions (Kleene, 1936, 1967; Kreisel, 1951, 1952). □

3.1 The weak $P < NP$ hypothesis

Notice that there are sensible, non-Π^0_2 formulations for $P < NP$. For instance, given a recursive enumeration (Baker, Gill, & Solovay, 1975) of the polynomial Turing machines, if predicate $G(m, x)$ formalizes "polynomial machine coded by m correctly guesses a satisfying line for x," we can say that the infinite set of sentences:

$$\exists x \, \neg G(0, x), \exists x \, \neg G(1, x), \exists x \, \neg G(2, x), \ldots$$

intuitively translates the idea expressed in the $P < NP$ hypothesis. However, while this formulation is implied by the Π^0_2 formulation, the converse doesn't hold. We may call it *the weak $P < NP$ hypothesis.*

3.2 Exotic formalization, I

From here on we follow our work (Costa & Doria, 2003). We have called "exotic" the formalization in Definition 3.5 as we have an intuitive equivalence between each one of those, and the standard formalization (see Proposition 3.6) that however doesn't always hold in several interesting axiomatic systems, such as PA or ZFC itself. Let f be a 1-variable, strictly increasing, total recursive function.

Definition 3.5 (An exotic formalization for $P = NP$, I.)
$[P = NP]^{\mathsf{f}} \leftrightarrow_{\text{Def}} \exists m \in \omega, a, b \, \forall x \in \omega \, [(t_m(x) \leq |x|^{\mathsf{f}(a)} + \mathsf{f}(b)) \wedge R(x, m)]$. □

3.3 Main theorem, informal version

The main result is proved within informal mathematics:

Proposition 3.6 *If* g *is total and strictly increasing,* $[P = NP]^{\mathsf{g}} \leftrightarrow [P = NP]$.

Proof within informal mathematics: It is enough to consider the bounding term.

- $[\Rightarrow]$. There are $a, b \in \omega$ so that $[t_m(x) \leq |x|^{\mathsf{g}(a)} + \mathsf{g}(b)]$. As g has an infinite domain and is strictly increasing, we have that for any a, b, there are $\mathsf{g}(a) = a', \mathsf{g}(b) = b'$. Thus the bounding term becomes $[t_m(x) \leq |x|^{a'} + b']$.
- $[\Leftarrow]$. There are $a', b' \in \omega$ so that $[t_m(x) \leq |x|^{a'} + b']$. Since g is strictly increasing, there are $a, b \in \omega$ so that $a' \leq \mathsf{g}(a)$, $b' \leq \mathsf{g}(b)$. Therefore there exist $a, b \in \omega$ such that $[t_m(x) \leq |x|^{\mathsf{g}(a)} + \mathsf{g}(b)]$. □

The argument above only requires that g have an infinite domain, and be strictly increasing on its domain; see Remark 3.33. We will later elaborate on the meaning of "informal mathematics" for this proof. The main reason why it doesn't go through in ZFC is because that theory cannot "see" all total recursive functions that exist in the standard model for arithmetic (Kleene, 1936, 1967).

Remark 3.7 From here on we drop the "$\in \omega$" in quantifiers, that is, $\exists x$ will mean $\exists x \in \omega$, and the same for $\forall x$. Exceptions will be dealt with accordingly. □

3.4 Exotic formalization, II

Let f be in general a (possibly partial) recursive function, and let e_{f} be the Gödel number of an algorithm that computes f. Let $p(\langle e_{\mathsf{f}}, b, c\rangle, x_1, x_2, \ldots, x_k)$ be an universal Diophantine polynomial with parameters e_{f}, b, c; for convenience we may take p to be positive definite.

The primitive recursive predicate $\neg Q$ given below in Definition 3.9 is actually dependent on a and b. However, to simplify things we substitute those for a single variable a, that is, we consider a bounding term of the form $|x|^a + a$.

Definition 3.8 $M_{\mathsf{f}}(x, y) \leftrightarrow_{\mathrm{Def}} \exists x_1, \ldots, x_k\, [p(\langle e_{\mathsf{f}}, x, y\rangle, x_1, \ldots, x_k) = 0]$. □

Actually $M_{\mathsf{f}}(x, y)$ stands for $M_{e_{\mathsf{f}}}(x, y)$, or better, $M(e_{\mathsf{f}}, x, y)$, as dependence is on the Gödel number e_{f}.

Definition 3.9 $\neg Q(m, a, x) \leftrightarrow_{\mathrm{Def}} [(t_m(x) \leq |x|^a + a) \rightarrow \neg R(x, m)]$. □

From here on we agree that all quantified variables range over the whole of ω unless specifically noted.

Definition 3.10 (Standard formalization, II.)

$$[P < NP] \leftrightarrow_{\mathrm{Def}} \forall m, a\, \exists x\, \neg Q(m, a, x). \qquad \square$$

Definition 3.11 $[P = NP] \leftrightarrow_{\mathrm{Def}} \neg[P < NP]$. □

From Definition 3.8:

Definition 3.12 $\neg Q_{\mathsf{f}}(m, a, x) \leftrightarrow_{\mathrm{Def}} \exists a'\, [M_{\mathsf{f}}(a, a') \wedge \neg Q(m, a', x)]$. □

This is the way we introduce function composition, to handle cases when f isn't total. We will sometimes write $\neg Q(m,\mathsf{f}(a),x)$ for $\neg Q_{\mathsf{f}}(m,a,x)$, whenever f is total. However, one should always keep in mind that actual dependence is on the Gödel number e_{f}, and not on f.

Definition 3.13 (Exotic formalization, II.)

$$[P<NP]^{\mathsf{f}} \leftrightarrow_{\mathrm{Def}} \forall m,a\,\exists x\,\neg Q_{\mathsf{f}}(m,a,x). \qquad \square$$

Notice that this is a Π_2^0 sentence:

$$\forall m,a\,\exists x,a',x_1,\ldots,x_k\,\{[p(\langle e_{\mathsf{f}},a,a'\rangle,\ldots,x_1,\ldots,x_k)=0]\wedge\neg Q(m,a',x)\}.$$

Q is primitive recursive. Since we decided to keep $[P<NP]^{\mathsf{f}}$ a Π_2^0 sentence, this requirement leads to Definition 3.12.

Definition 3.14 $[P=NP]^{\mathsf{f}} \leftrightarrow_{\mathrm{Def}} \neg[P<NP]^{\mathsf{f}}$. □

Remark 3.15 For the universal polynomial $p(\langle e_{\mathsf{f}},a,b\rangle,x_1,x_2,\ldots,x_k)$, if e_{f} is the index of a Turing machine that computes f:

1. $[\mathsf{f}$ is total$] \leftrightarrow_{\mathrm{Def}} \forall a\,\exists b,x_1,\ldots,x_k\,[p(\langle e_{\mathsf{f}},a,b\rangle,x_1,\ldots,x_k)=0]$.
2. $[\mathsf{f}$ is total$] \leftrightarrow \forall a\,\exists b\,M_{\mathsf{f}}(a,b)$.
3. If $\forall a\,\exists b\,M_{\mathsf{f}}(a,b)$ then:

   (a) $\mathsf{f}(a) =_{\mathrm{Def}} \mu_b M_{\mathsf{f}}(a,b)$.

   (b) $\forall a\,M_{\mathsf{f}}(a,\mathsf{f}(a))$. □

There are two ways we can show that a function g grows faster than function f: either by direct comparison, as when we check that an exponential function overtakes all polynomial functions; or by "cloning" or "embedding" some adequately fast-growing function into the function we are considering. That last procedure is used here in an indirect way; we will exhibit a more direct construction in (Costa et al., 2007).

3.5 Function F, $[P<NP]^{\mathsf{F}}$ and $[P<NP]$

We use here a well-known recursive function that is diagonalized over all ZFC—provably total recursive functions. We note it F in this section. See (Costa & Doria, 2003), and (Kaye, 1991), pp. 51–52 on it.

Remark 3.16 For each n, $\mathsf{F}(n)=\max(\{e\}(k))+1$, that is is the sup of those $\{e\}(k)$ such that:

1. $k\le n$.
2. $\ulcorner \mathrm{Pr}_{\mathrm{ZFC}}(\ulcorner\forall x\,\exists z\,T(e,x,z)\urcorner)\urcorner \le n$.

$\mathrm{Pr}_{\mathrm{ZFC}}(\ulcorner\xi\urcorner)$ means, there is a proof of ξ in ZFC, where $\ulcorner\xi\urcorner$ means: the Gödel number of ξ. So, $\ulcorner\mathrm{Pr}_{\mathrm{ZFC}}(\ulcorner\xi\urcorner)\urcorner$ means: "the Gödel number of sentence 'there is a proof of ξ in ZFC.' "

Condition 2 above translates as: there is a proof of $[\{e\}$ is total$]$ in ZFC whose Gödel number is $\le n$. (T is Kleene's predicate (Kleene, 1967).) □

Proposition 3.17 *We can explicitly compute a Gödel number* e_{F} *so that* $\{e_{\mathsf{F}}\}=\mathsf{F}$. □

Proposition 3.18 $[\mathsf{F}$ *is total*$]$ *isn't proved within* ZFC, *supposed consistent.* □

Definition 3.19 $[P < NP]^{\mathsf{F}} \leftrightarrow_{\mathrm{Def}} \forall m, a\, \exists x\, \neg Q_{\mathsf{F}}(m, a, x)$. □

Lemma 3.20 *If* $I \subseteq \omega$ *is infinite and* $0 \in I$*, then:*

$$\mathrm{ZFC} \vdash \{[\forall m\, \forall a \in I\, \exists x \neg Q(m, a, x)] \to [\forall m\, \forall a \in \omega\, \exists x\, \neg Q(m, a, x)]\}. \qquad \square$$

This is the "size of gaps doesn't matter" result. Its meaning is: as long as we have an infinite succession of ever larger bounds that make the Turing machines polynomial, our standard definitions hold. The size of the intermediate gaps between each pair of bounds doesn't matter.

This also shows that our exotic formalizations $[P < NP]^{\mathsf{F}}$ and $[P = NP]^{\mathsf{F}}$ can be viewed as reasonable, informally equivalent, versions for $[P < NP]$ and $[P = NP]$.

3.6 Main theorem, formalized version

Now follows our main result. Proof is computational:

Proposition 3.21 $\mathrm{ZFC} \vdash [P < NP]^{\mathsf{F}} \leftrightarrow \{[\mathsf{F}\ \textit{is total}\,] \wedge [P < NP]\}$.

Proof See (Costa & Doria, 2003). □

We quote and prove a scholium due to its interest:

Scholium 3.22 $\mathrm{ZFC} \vdash [P < NP]^{\mathsf{F}} \to [\mathsf{F}\ \textit{is total}]$. □

Remark 3.23 The following informal argument clarifies the meaning of the scholium and gives a proof for it: let $\mathsf{f}_{\mathsf{F}}(\langle m, a\rangle) = \min_x [\neg Q(m, \mathsf{F}(a), x)]$, where we can here look at F as a (partial) recursive function. (The brackets $\langle \ldots, \ldots \rangle$ note the usual 1–1 pairing function.) Now if f_{F} is total, then $\mathsf{F}(a)$ has to be defined for all values of the argument a, that is, F must be total. The function f_{F} is the so-called *exotic counterexample function* to $[P = NP]^{\mathsf{F}}$.

We can similarly define a *standard counterexample function*:

$$\mathsf{f}(\langle m, a\rangle) = \min_x [\neg Q(m, a, x)]. \qquad \square$$

Proof of the scholium The actual argument we present is here given in full as a sample of the techniques used in Costa & Doria (2003).

1. $\neg Q_{\mathsf{F}}(m, a, x) \leftrightarrow_{\mathrm{Def}} \exists b\, [M_{\mathsf{F}}(a, b) \wedge \neg Q(m, b, x)]$. (Definition of $\neg Q_{\mathsf{F}}$.)
2. Assume $\neg Q_{\mathsf{F}}(m, a, x)$. From step 1.:
   (a) $\exists b\, [M_{\mathsf{F}}(a, b) \wedge \neg Q(m, b, x)]$. (Assumed.)
   (b) $[(\exists b M_{\mathsf{F}}(a, b)) \wedge (\exists b \neg Q(m, b, x))]$. (From step 2a., and modus ponens.)
3. $\neg Q_{\mathsf{F}}(m, a, x) \to [(\exists b M_{\mathsf{F}}(a, b)) \wedge (\exists b \neg Q(m, b, x))]$. (The assumption in step 2a. implies step 2b.)
4. Recall that the following is a propositional theorem: $[A \to (B \wedge C)] \to [A \to B]$.
5. $[\neg Q_{\mathsf{F}}(m, a, x)] \to [\exists b (M_{\mathsf{F}}(a, b)]$. (From step 4. applied to step 3. plus modus ponens.)
6. Generalization rule applied to the unquantified variable x: $\forall x\, \{[\neg Q_{\mathsf{F}}(m, a, x)] \to [\exists b (M_{\mathsf{F}}(a, b)]\}$.
7. Internalization of $\forall x$, and modus ponens: $\{[\exists x \neg Q_{\mathsf{F}}(m, a, x)] \to [\exists b (M_{\mathsf{F}}(a, b)]\}$.
8. Generalization Rule applied to the unquantified variables m, a: $\forall m, a\, \{\exists x [\neg Q_{\mathsf{F}}(m, a, x)] \to [\exists b\, M_{\mathsf{F}}(a, b)]\}$.

9. $[\forall m, a\, \exists x[\neg Q_{\mathsf{F}}(m,a,x)] \rightarrow [\forall a \exists b\, M_{\mathsf{F}}(a,b)]$. (Operations, modus ponens, and elimination of empty quantifier.)
10. That is, $[P < NP]^{\mathsf{F}} \rightarrow [\mathsf{F}$ is total]. □

From the properties of F and from the the scholium:

Corollary 3.24

$$\mathrm{ZFC} \vdash [P < NP]^{\mathsf{F}} \rightarrow \mathrm{Consis}\,(\mathrm{ZFC}),$$

where Consis (ZFC) *is the usual sentence that asserts the consistency of* ZFC. □

This gives an idea of the—if we may say so—conceptual depth of the objects we are dealing with here. We handle sentences that imply the consistency of the axiomatic framework they are embedded into.

3.7 Results

As we have seen, intuitively and informally $[P < NP]^{\mathsf{F}}$ and $[P < NP]$ are the same thing. The chief result we have is that ZFC, if consistent, doesn't prove $[P < NP]^{\mathsf{F}}$, that is, it doesn't prove our exotic version for $P < NP$. The reason is clear: as we have "cloned" F into $[P < NP]^{\mathsf{F}}$, then it cannot be proved in ZFC, as ZFC cannot prove that F is total. However, we also show that if ZFC is consistent, then so is $\mathrm{ZFC}^* = \mathrm{ZFC} + [P < NP] \leftrightarrow [P < NP]^{\mathsf{F}}$. And the point is, how strong is this new theory? Namely: $\mathrm{ZFC} \nvdash [P < NP]^{\mathsf{F}}$. What about ZFC* ? Formally:

Proposition 3.25 *If* ZFC *is consistent, then* ZFC *doesn't prove* $[P < NP]^{\mathsf{F}}$.

Proof $\mathrm{ZFC} \vdash [[P < NP]^{\mathsf{F}} \rightarrow (\mathsf{F}$ is total)]. (Scholium 3.22.) So, ZFC cannot prove $[P < NP]^{\mathsf{F}}$. □

Corollary 3.26 $[P = NP]^{\mathsf{F}}$ *is consistent with* ZFC. □

From the viewpoint of naïve wisdom, this result is unexpected: it asserts that it is consistent with ZFC to assume that there is a poly algorithm—an easy, fast, algorithm. A different construction that led to a similar result for Peano Arithmetic was presented in May 2000 (Doria, 2000). We also note that such a result is in the line of the previous results by DeMillo and Lipton (1979, 1980) and by Joseph and Young (1980, 1981).

Remark 3.27 Guillaume (2000–2002) pointed out to N. C. A. da Costa and the author that the results for the P versus NP question we present here hold in any first-order theory with a recursively enumerable set of theorems where theory $I\Sigma_1$ can be interpreted. That is to say, formal sentences like $[P < NP]^{\mathsf{F}}$ that intuitively translate as $P < NP$ cannot be derived in a whole family of very reasonable, strong, formal systems. The author is aware that this kind of result goes against the current expectations in the field (Aronson, 2003; Ben-David & Halevi, 1991), as it is usually believed that some formal sentence that intuitively translates as $P < NP$ will eventually be proved within perhaps Peano Arithmetic. This is certainly not the case with the exotic formalization for $P < NP$ presented here, while the possibility remains open that there will be another reasonable formalization for $P < NP$ which isn't equivalent to the ones in the paper, and that can be proved in arithmetic. This wouldn't be an unusual situation, after all, the existence of reasonable multiple nonequivalent formalizations for the same intuitive concept (Franzen, 2004). □

3.8 From an informal to a formal equivalence result

We start from a reformulation of Proposition 3.6:

Proposition 3.28 *If* ZFC *has a model* **N** *with standard arithmetic, that is, if* ZFC *is* arithmetically sound, *then* ZFC $+ [P < NP]^{\mathsf{F}} \leftrightarrow [P < NP]$ *is consistent and holds of* **N**. □

Σ_1-soundness is quite natural: from the intuitive viewpoint it asks that, for a Σ_1 sentence ϕ in the language of a theory T like the ones considered here, if $T \vdash \phi$, then ϕ (ϕ is true of the standard model for arithmetic). That requirement can be formalized as a Reflection Principle (Beklemishev, 1997; Feferman, 1960, 1962; Franzen, 2004; Smoryński, 1989), but we will not require its actual explicit formulation here; suffices to abbreviate (with some abuse of language) it as [T is Σ_1-sound]. The next result is standard; it is proved in the case of Peano Arithmetic in (Paris & Harrington, 1977); the generalization to ZFC is straightforward (Beklemishev, 1997) :

Lemma 3.29 ZFC $\vdash$ [F *is total*] $\leftrightarrow$ [ZFC *is* Σ_1*–sound*]. □

Proposition 3.30 ZFC + [ZFC *is* Σ_1*–sound*] $\vdash [P < NP] \leftrightarrow [P < NP]^{\mathsf{F}}$. □

Follows as a corollary Proposition 3.28. Then, going further:

Corollary 3.31

1. ZFC + [F *is total*] $+ [P = NP]^{\mathsf{F}} \vdash [P = NP]$.
2. *If theory* ZFC + [F *is total*] $+ [P = NP]^{\mathsf{F}}$ *is consistent, then so is theory* ZFC $+ [P = NP]$.
3. *If* ZFC $+ [P = NP]^{\mathsf{F}}$ *is ω-consistent, then* ZFC $+ [P = NP]$ *is consistent.*

Proof For the last assertion see that a consistent theory $T + \neg$ [F is total] cannot be ω-consistent. □

Remark 3.32 We may insist (Costa & Doria, 2003) that it still isn't obvious why a theory like ZFC $+ [P = NP]^{\mathsf{F}}$ should be ω-consistent. From the intuitive viewpoint, our condition is equivalent to the following: *there is a model for T where all bounds* $|x|^{\mathsf{F}(a)} + \mathsf{F}(b)$ *in the clocks that regulate our polynomial machines do in fact converge.* So, we are here dealing with a reasonable property. But—does it hold? Is there one such model for ZFC? □

3.9 The formal viewpoint meets the informal

Let us summarize our discussion up to this point:

- $[P < NP]^{\mathsf{F}}$ and $[P < NP]$ both translate in a reasonable intuitive way the same concept.
- ZFC $\nvdash [P < NP]^{\mathsf{F}}$, while we still don't know what happens to $[P < NP]$.
- ZFC$^* =$ ZFC $+ [P < NP] \leftrightarrow [P < NP]^{\mathsf{F}}$ is consistent, if so is ZFC. But can we repeat with ZFC* our results for ZFC ?

The next discussion will appear in full and in a formal version in (Costa et al., 2007), but we sketch it here. We essentially use Proposition 3.6 and Proposition 3.28.

Remark 3.33 We know, from the previous results, but also from Proposition 3.28, that theory:

$$\text{ZFC}^* = \text{ZFC} + [P < NP] \leftrightarrow [P < NP]^{\mathsf{F}}$$

is consistent, if so is ZFC. Moreover, if we examine the proof of that proposition, or of Proposition 3.6, we notice that it is *sufficient* to suppose that F has an infinite domain. Therefore, one naïvely sees that ZFC* cannot prove [F is total].

Informally always, as we derive in the same theory the equivalence $[P < NP] \leftrightarrow [P < NP]^{\mathsf{F}}$, and again as $[P < NP]^{\mathsf{F}} \rightarrow$ [F is total], we cannot prove $[P < NP]$ in ZFC*, and a fortiori in ZFC.

Notice that this is an *informal* argument, based on Propositions 3.6 and 3.28. Can we make it formal? We discuss that matter in (Costa et al., 2007).

Finally: do we have an F' such that it has an infinite domain, grows—over its domain—as F, but such that it is undecidable whether [F' is total] ? We can use here another variant of a trick developed in 1991 to prove the undecidability of chaos (Costa & Doria, 1991, Proposition 3.28; see also Costa & Doria, 2005, Propositions 12 and 13, and Costa et al., 2007). □

4 Informal versus formal mathematics

Our main tool in this paper are the fast-growing functions that cannot be proved to be total by some quite strong theories like ZFC and its nonrecursive extension ZFC_1. We are now going to take a closer look at them.

One of the main trends in mathematical thought, to which the author adheres, is that no axiomatization like those currently available for arithmetic or set theory can exhaust, or even approximate, the wealth of intuitive mathematics. (This idea lurks behind (Kreisel, 1969).) Yet axiomatic systems are pointers that show the way to that wealth in the intuitive domain of mathematics, due to the incompleteness phenomenon. This section wishes to present the case for that claim.

We start from the following example: suppose that we are given a prescription so that, for an integer n, we can compute a finite set of numbers S_n. Then put: $\mathsf{F}^*(n) = \max S_n + 1$. Most mathematicians would immediately agree that F^* is both computable and total. But is that really so?

Can we construct that function within ZFC? Sure—but we won't then be able to prove that similar functions are total, in the general case. It is enough to consider the construction and properties of F in Remark 3.16.

Let's take a closer look at F.

4.1 Why can't we prove in ZFC that F is total?

Let's substitute the construction of F in Remark 3.16 by the following one, for a function F' that can be proved to be primitive recursive in F:

- We suppose that ZFC is consistent; therefore, it has a nontrivial recursively enumerable set of theorems. Suppose that we have built the Turing machine {ZFC} that outputs all theorems of ZFC, and only those.

  This machine operates as follows: if x is the Gödel number of sentence ξ_x in the language of ZFC, $\{\text{ZFC}\}(x) \downarrow$, that is, it stops if and only if ξ_x is a theorem of ZFC, and diverges ($\{\text{ZFC}\}(x) \uparrow$) otherwise.

- Turn on $\{\mathrm{ZFC}\}$. Enumerate by dovetailing all theorems of ZFC.
- Out of the list of theorems of ZFC, pick up those of the form $\forall x\, \exists y\, T(e,x,y)$, where T is Kleene's predicate and e the Gödel number of $\{e\}$.
- If e appears in the n-th theorem of the form $\forall x\, \exists y\, T(e,x,y)$, define $\mathsf{F}'(n) = \{e\}(n) + 1$.
- F' is computable—it has an algorithm, just described; it is total in the intuitive way, and cannot be proved to be total in ZFC, as it is diagonal over all provably total $\{e\}$.

The nonprovability of [F' is total] in ZFC follows from a diagonal argument (see Remark 3.16). Crucial to the diagonal argument is the recursive enumerability of the sentences—and thus of the Gödel numbers e—that provably mean [$\{e\}$ is total]. This idea goes back to Kleene in 1936: (Kleene, 1936), (Kleene, 1967), p. 257.

4.2 Σ_1-soundness again

The infinite set of formal sentences (Feferman, 1960, 1962; Smoryński, 1985) abbreviated as [ZFC is Σ_1-sound] can be proved to be equivalent to [F is total]. Σ_1-soundness translates as an intuitive fact—if ZFC proves the sentence $\exists x\, P(x)$, for primitive recursive P, then there is an a so that $P(a)$ holds. Σ_1-unsound theories are easy to construct (see Smoryński, 1991, p. 342ff), but the concept is again definitively non-intuitive, when applied to the natural numbers.

Now, given a function F like the one in Remark 3.16, we can establish the following, as noticed before:

Proposition 4.1 ZFC $\vdash$ [F *is total*] $\leftrightarrow$ [ZFC *is* Σ_1*–sound*]. □

We wouldn't be able to get a glimpse of Σ_1-unsound theories (or of ω-inconsistent theories) if it weren't for the incompleteness phenomenon; in the present case, if those "boundary" functions like F or F'—functions that overtake infinitely many times all provably total recursive functions in ZFC and sort of mark the "boundary" between what we can prove and what we cannot prove in ZFC—didn't exist with all its problems. They remind us of G. Chaitin's algorithmic complexity version of the Gödel incompleteness theorem (Chaitin, 1987a, b), which also offers a bound beyond which there is formal incompleteness and to which they may perhaps be related.

Σ_1-unsound but consistent theories, as well as consistent, ω-inconsistent theories, point towards a mathematical world that defies many of our conventional views, like e.g. an implicit structure for the standard natural numbers that a priori gets rid of infinitely large naturals that appear in nonstandard constructions (or that simply give a more precise meaning for the ... that appear in the enumeration of the naturals that we can actually name, such as $0, 1, 2, 3, \ldots$).

So here we go from a formally correct, but unexpected, phenomenon like ω-inconsistency or Σ_1-unsoundness, to a new, potentially richer, intuition about the mathematical world.

4.3 Hierarchy of the F's

Wainer (1970) gives an explicit construction for a hierarchy of fast-growing total recursive functions along lines that first appear in Kreisel (1951, 1952). He stops at F_{ϵ_0}, but we can go much further, as long as we have a notation for the ordinal that is ascribed to the fast-growing function. Some of those have specific meanings:

- Ackermann's Function has the growth-rate of F_{ω}, and marks the "outer boundary," if we may say so, of Primitive Recursive Arithmetic (as every primitive recursive function is dominated by F_{ω}).
- F_{ϵ_0} has the same rate of growth as the diagonal Paris–Harrington function, and indicates the external boundary of PA.
- Since Ackermann required an induction up to η to prove the consistency of an axiomatized theory of the real numbers, $\mathsf{F}_{\eta} > \mathsf{F}_{\epsilon_0}$ indicates the external boundary of that theory.

  And so on.

All those functions can be given explicit programs, and even rather short ones (Wainer, 1970). However the point is: while they look intuitively total, we need progressively stronger theories to show that fact. This point must be stressed: the concept of "total recursive function" is only safely applied with reference to a given axiomatic framework. The stronger (or perhaps more restrictive) the framework, the larger will be the corresponding class of total recursive functions.

This again relativizes (with respect to axiomatic systems) a concept that might seem to be safe, that of total recursive function.

And, as the cherry on top of the pie, if we intuitively accept that [F is total] holds, then we must intuitively accept that [PA is consistent] also holds—very much like the Kunen–Boyer–Moore algorithm that leads—informally—to [PA is consistent] whenever it verifies the Paris–Harrington theorem.

4.4 A remark on ZFC

The situation for ZFC is more involved, as we have no knowledge of an ordinal α so that our function $\mathsf{F} = \mathsf{F}_{\alpha}$. Such an ordinal $\alpha > \omega_1$, where ω_1 is the first nonconstructive ordinal (Rogers, 1967). Therefore we cannot argue that F is intuitively total out of a hierarchy of total functions based on constructive ordinals.

4.5 From the intuitive to the counterintuitive

Of course we have theories that prove that all intuitively total recursive F are total: the Turing—Feferman theorem ensures that (Beklemishev, 1997; Feferman, 1960, 1962; Franzen, 2004; Turing, 1939); actually it is enough to add Shoenfield's recursive ω-rule (Shoenfield, 1959) to Peano Arithmetic to ensure it (Costa & Doria, 2006). Such theories are so close to being recursively enumerable theories like PA or ZFC that we may call them near-recursive theories. (See also Longo, 2002.)

But how about models for some axiomatic system like ZFC where F isn't a total function? First, one easily sees that any consistent theory like ZFC $+ \neg$ [F is total] is an ω-inconsistent theory, as we have already pointed out.

But how about the behavior of F in those theories? We have several alternatives:

- Does F have a finite domain?
- Does F have an infinite domain?
- How are we to distinguish the two possible situations: our theory proves, for some $a \in \omega$, that $\exists y\, y = \mathsf{F}(a)$, but we can find no constant $b \in \omega$ so that $b = \mathsf{F}(a)$.
  (So we may say—intuitively? in such a weird situation?—that one goes on computing $\mathsf{F}(a)$ but one never manages to find a b that fits the place of $\mathsf{F}(a)$.)

Or we may say that the elements of ω which aren't in the domain of F have no names.

These, if we may say so, highly counterintuitive possibilities appear not to have been extensively considered in the literature on the subject. We may wonder which magnificent mathematical worlds may be uncovered when we start to explore them.

Remark 4.2 The author thanks an anonymous referee for his detailed examination of this work plus several interesting comments on it. □

Acknowledgements It is a pleasure to dedicate this paper to Pat Suppes on his 80th anniversary. This paper is part of the Research Project on Complexity and the Foundations of Computation at the Institute for Advanced Studies, University of São Paulo (IEA–USP); we thank them for support. Regretfully no support from Brazilian federal agencies can be acknowledged.

Very special thanks are due to M. Guillaume, for his careful and exacting review and criticism of the work in Sect. 3 since the very early version that was presented at the Brazilian Logic Symposium in May 2000. The author owes him many corrections, suggestions and criticisms which were duly incorporated here; we can say for sure that Guillaume's help was decisive to place the present work upon firm soil.

References

Aronson, S. (2003). Is *P* vs. *NP* formally independent? In L. Fortnow (Ed.), *The computational complexity column*. The paper can be found at http://www.cs.uchicago.edu/~fortnow/beatcs

Baker, T., Gill, J., & Solovay, R. (1975). Relativizations of the *P* =?*NP* question. *SIAM Journal of Computing, 4*, 431.

Beklemishev, L. (1997). Provability and reflection. Lecture Notes for ESSLLI' 97.

Ben-David, S., & Halevi, S. (1991). On the independence of *P* vs. *NP*. Technical Report # 699, Technion.

Chaitin, G. (1987a). Information-theoretic limitations of formal systems. In *Information, randomness & incompleteness*. World Scientific.

Chaitin, G. (1987b). *Information, randomness & incompleteness*. World Scientific.

da Costa, N. C. A., & Doria, F. A. (1991). Undecidability and incompleteness in classical mechanics. *International Journal Theoretical Physics, 30*, 1041–1073.

da Costa, N. C. A., & Doria, F. A. (2002). Barebones—notes on complexity and the *P* =?*NP* question. unpublished, IEA–USP.

da Costa, N. C. A., & Doria, F. A. (2003). Consequences of an exotic definition for *P* = *NP*. *Applied Mathematics and Computation, 145*, 655; da Costa, N. C. A., & Doria, F. A. (2005). Addendum to: 'Consequences of an exotic definition for *P* = *NP*'. *Applied Mathematics and Computation* (to appear).

da Costa, N. C. A., & Doria, F. A. (2004). On set theory as a foundation for computer Science. *Bull. Section of Logic, Łodz, 33*, 33–40.

da Costa, N. C. A., & Doria, F. A. (2005). Computing the future. In K. Vela Velupillai (Ed.), *Computability, complexity and constructivity in economic analysis*. Blackwell.

da Costa, N. C. A., & Doria, F. A. (2006). Some thoughts on hypercomputation. To appear in Doria, F. A., & Costa, J. F., Hypercomputation – special issue. *Applied Mathematics and Computation*, *178*, 83–92.

da Costa, N. C. A., Doria, F. A., & Bir, E. (2007). On the metamathematics of *P* vs. *NP*. *Applied Mathematics and Computation* (to appear).

DeMillo, R. A., & Lipton, R. J. (1979). Some connections between computational complexity theory and mathematical logic. In *Proceedings of the 12th Annual ACM Symposium on the Theory of Computing* (pp. 153–159).

DeMillo, R. A., & Lipton, R. J. (1980). The consistency of *P* = *NP* and related problems with fragments of number theory. In *Proceedings of the 12th Annual ACM Symposium on the Theory of Computing* (pp. 45–57).

Doria, F. A. (2000). Talk at the Brazilian logic symposium on his joint work with N. C. A. da Costa on computational complexity.

Feferman, S. (1960) Arithmetization of metamathematics in a general setting. *Fundamental Mathematics, 49*, 35.

Feferman, S. (1962). Transfinite recursive progressions of axiomatic theories. *Journal of Symbolic Logic, 27*, 259.

Fortune, S., Leivant, D., & O'Donnell, M. (1983). The expressiveness of simple and second-order type structures. *Journal of ACM, 38*, 151–185.

Franzen, T. (2004). Transfinite progressions: A second look at completeness. *Bulletin of Symbolic Logic, 10*, 367–389.

Guillaume, M. (Sept. 2000—Aug. 2002). E-mail messages to N. C. A. da Costa and F. A. Doria.

Hartmanis, J., & Hopcroft, J. (1976). Independence results in computer science. *SIGACT News, 13*.

Joseph, D., & Young, P. (1980). Independence results in computer science? *Proceedings of the 12th Annual ACM Symposium on the Theory of Computing* (pp. 58–69).

Joseph, D., & Young, P. (1981). Fast programs for initial segments and polynomial time computation in weak models of arithmetic. *STOC Milwaukee, 1981*, 55–61.

Kaye, R. (1991). *Models of Peano arithmetic*. Clarendon Press.

Kleene, S. C. (1936). General recursive functions of natural numbers. *Mathematische Annalen, 112*, 727.

Kleene, S. C. (1967). *Mathematical logic*. Wiley.

Kowalczyk, W. (1982). A sufficient condition for the consistency of $P = NP$ with Peano Arithmetic. *Fundanemta Informaticae, 5*, 233–245.

Kreisel, G. (1951). On the interpretation of non–finitist proofs, I and II. *Journal of Symbolic Logic, 16*, 241.

Kreisel, G. (1952). On the interpretation of non–finitist proofs, I and II. *Journal of Symbolic Logic, 17*, 43.

Kreisel, G. (1969). Informal rigor. In J. Hintikka (Ed.), *The philosophy of mathematics*. Oxford.

Kunen, K. (1995). A Ramsey theorem in Boyer–Moore logic. *Journal of Automated Reasoning, 15*, 217–230.

Longo, G. (2002). On the proofs of some formally unprovable propositions and prototype proofs in type theory. *Lecture Notes in Computer Science, 2277*, 160–167 (Springer).

Machtey, M., & Young, P. (1979). *An introduction to the general theory of algorithms*. North-Holland.

Paris, J., & Harrington, L. (1977). A mathematical incompleteness in Peano arithmetic. In J. Barwise (Ed.), *Handbook of mathematical logic.* North-Holland.

Rogers Jr., H. (1967). *Theory of recursive functions and effective computability*. McGraw–Hill.

Shoenfield, J. (1959). On a restricted ω-rule. *Bulletin Academie Polonaise des Sciences, Serie Sciences Mathematiques, 7*, 405–407.

Smoryński, C. (1985). Some rapidly growing functions. In L. A. Harrington et al. (Eds.), *Harvey Friedman's research on the foundations of mathematics*. Elsevier.

Smoryński, C. (1989). The incompleteness theorems. In J. Barwise (Ed.), *Handbook of mathematical logic*. North-Holland.

Smoryński, C. (1991). *Logical number theory, I*. Springer.

Suppes, P. (2002). *Representation and invariance of scientific structures*. Stanford, CSLI Publications.

Turing, A. M. (1939). Systems of logic based on ordinals. *Proceedings of the London Mathematical Society Serie, 2, 45*, 161.

Wainer, S. S. (1970). A classification of the ordinal recursive function. *Archiv Mathematik Logik, 18*, 136.