Counting as Integration in Feasible Analysis

Fernando Ferreira[*1] and **Gilda Ferreira**[**2]

[1] Departamento de Matemática, FCUL - Universidade de Lisboa, P-1749-016 Lisboa, Portugal.
[2] CMAF - Universidade de Lisboa, P-1649-003 Lisboa, Portugal.

Received ..., revised ..., accepted ...
Published online ...

Key words Feasible analysis, integration, counting, reverse mathematics.
MSC (2000) 03F35, 03B30, 03F30

Suppose that it is possible to integrate real functions over a weak base theory related to polynomial time computability. Does it follow that we can count? The answer seems to be: *obviously* yes! We try to convince the reader that the severe restrictions on induction in feasible theories preclude a straightforward answer. Nevertheless, a more sophisticated reflection does indeed show that the answer is affirmative.

1 Introduction

The first author defined in [3] a system of second-order arithmetic BTFA as a groundwork for formalizing and studying the strength of mathematical results over a weak basis related to polynomial time computability. This study is in the spirit of work in *reverse mathematics* initiated by S. Simpson and H. Friedman – see the encyclopedic [12] – where the standard base theory RCA_0 is related to the primitive recursive functions (note that the theories of Simpson and Friedman, as well as the theories of this paper, are theories framed in *classical* logic). The weakening of the base theory was motivated by a challenge of W. Sieg (cf. [11]): to find a mathematically significant subsystem of analysis whose class of provable recursive functions consists only of the computationally "feasible" ones. BTFA is a formal system of analysis (thus, with two sorts of variables) which, however, departs from the usual presentations of arithmetic by being based on a language that purports to describe the set $\{0,1\}^*$ of finite sequences of 0s and 1s (binary strings), instead of the natural numbers. The option for this setting is mainly a matter of taste and convenience, where the latter lies in the fact that in the absence of (the totality) of exponentiation a distinction must be made between numbers $\mathbb{N}_2$ as given by their binary notation (dyadic natural numbers) and mere *tally* numbers envisaged as being given by finite sequences of 1s. This distinction is notationally very clear in systems based on $\{0,1\}^*$.

The fundamental numerical systems of mathematics (up to, and including, the real numbers) were introduced in [5] within the base theory BTFA, as well as the notion of a continuous function of real variable. Therein, it is proved (within BTFA) *Bolzano's intermediate value theorem*, a result which shall be used below. Riemann integration does not seem to have a decent formalization within BTFA, but the second author has recently shown in [4] how to formalize the Riemann integral (for continuous functions with a modulus of uniform continuity) up to the *fundamental theorem of calculus* over a certain weak theory *extending* BTFA. The necessity of going to an extension of BTFA is strongly suggested by general considerations and, more specifically, by work of H. Friedman and K. Ko in the context of the complexity theory of real functions (cf. [9], [8]), where they relate Riemann integration with the class of the polynomial space computable functions. It must be observed that our game, although related to the study of the complexity of real functions, is nevertheless different – namely, a game on *formalizing* mathematics within weak second-order theories (in particular, induction cannot be freely used). G. Ferreira's extended formal theory (see also the forerunner theory [7]) lies in strength between $\mathsf{I}\Delta_0 + \Omega_1$ and S. Buss' theory U^1_2 (see [6] and [1] for these theories). The latter theory is related to polynomial space computability,

* e-mail: ferferr@cii.fc.ul.pt
** e-mail: gildafer@cii.fc.ul.pt

and Riemann integration can indeed be formalized in (a suitable second-order formulation of) it (this follows from G. Ferreira's work). G. Ferreira's theory is (in a precise sense) weaker than U_2^1. As a matter of fact, its provably total functions are exactly those in the *hierarchy of counting functions*, a class defined by K. Wagner in [13] based on an iteration procedure which stems from L. Valiant's well-known counting class $\#P$. The functions of this hierarchy are all polynomial space computable.

The main principle of G. Ferreira's theory is the *counting principle* that guarantees the existence of a counting function up to a for a given set X:

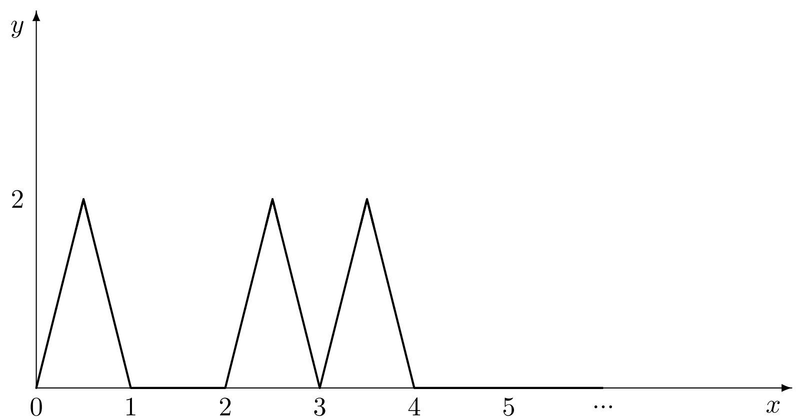
$$\forall a \in \mathbb{N}_2 \forall X \exists f (f(0) = 0 \wedge \forall x \leq a((x \notin X \rightarrow f(x+1) = f(x)) \wedge (x \in X \rightarrow f(x+1) = f(x)+1))).$$

In this paper we address the *reverse* question: does a good theory of integration for real functions imply the counting principle? The answer seems to be: *obviously* yes! Given $X \subseteq \mathbb{N}_2$, it is possible to associate (within BTFA) a continuous total function, Φ_X, from $[0, +\infty[$ into $\mathbb{R}$, with a modulus of uniform continuity, such that:

$$\begin{cases} \Phi_X(\alpha) = 0, & \text{if } m \notin X; \\ \Phi_X(\alpha) = 4\alpha - 4m, & \text{if } m \in X \wedge \alpha \leq m + \frac{1}{2}; \\ \Phi_X(\alpha) = -4\alpha + 4(m+1), & \text{if } m \in X \wedge m + \frac{1}{2} < \alpha; \end{cases}$$

where $\alpha \geq 0$ is a real number and m is the dyadic natural number such that $m \leq \alpha < m+1$.

For instance, if $X = \{0, 2, 3\}$ then Φ_X is the function:



If we have integration, the function $f(x) =_{\mathbb{R}} \int_0^x \Phi_X(t)dt$ does the counting. We must nevertheless be careful because the counting function f (which we may suppose to be defined only in the dyadic natural numbers not exceeding a certain given value $a+1$) is supposed to be a function taking values in $\mathbb{N}_2$, not in $\mathbb{R}$. *Prima facie*, this seems to be only a technical detail without any importance whatsoever. In fact, since functions are (in our setting) given by sets of ordered pairs, we could just define:

$$f := \{\langle x, n \rangle : x, n \in \mathbb{N}_2, x \leq a+1,\ n \leq x,\ \textstyle\int_0^x \Phi_X(t)dt =_{\mathbb{R}} n_{\mathbb{R}}\},$$

where $\langle\ ,\ \rangle$ is a smooth pairing function and $n_{\mathbb{R}}$ is the real number (second-order entity) corresponding to the dyadic natural number n (henceforth, we will omit the subscript). Unfortunately, this "set" is given by Π_1^0-comprehension and this much comprehension is not available in BTFA. Notwithstanding, f can be given as well by

$$f = \{\langle x, n \rangle : x, n \in \mathbb{N}_2, x \leq a+1,\ n \leq x,\ n - \tfrac{1}{2} <_{\mathbb{R}} \textstyle\int_0^x \Phi_X(t)dt <_{\mathbb{R}} n + \tfrac{1}{2}\},$$

and this definition is Σ_1^0. Since comprehension for Δ_1^0-definable sets is available in BTFA, the issue is apparently disposed of. However, a more basic problem looms. How do we know, within BTFA, that the above two definitions of f coincide? We are presupposing that, for $x \leq a+1$, $\int_0^x \Phi_X(t)dt \in \mathbb{N}_2$. The reader may protest and say this is an obvious fact which should follow from any decent theory of integration. We agree that it is obvious, but *no thanks* to integration. A decent theory of integration should indeed show that

$\int_0^{x+1} \Phi_X(t)dt =_{\mathbb{R}} \delta +_{\mathbb{R}} \int_0^x \Phi_X(t)dt,$

for δ the real number 1 or 0 (according to whether $x \in X$ or not). In particular, the implication $f(x) \in \mathbb{N}_2 \rightarrow f(x+1) \in \mathbb{N}_2$ holds. However, it is *not* integration that allows the inference from this implication to $f(x) \in \mathbb{N}_2$, for all $x \leq a+1$. It is *induction*! It is induction on x applied to the formula $\exists n \leq x(\int_0^x \Phi_X(t)dt =_{\mathbb{R}} n)$. This amounts to induction with respect to a Π^0_1-formula, and this much induction is simply not available in BTFA.

We hope that we have convinced the reader that there is something to be said for the implication $integration \rightarrow counting$ over a weak base theory (over BTFA). In the remaining part of the paper we show that, under natural conditions, this implication does indeed hold within BTFA.

We will not specify precisely what do we mean by integration. Instead, we just use in the proof properties that any *decent* theory of integration of real functions should have. Beyond these, we only appeal to results that stem from the very framework in which we are working. Let us be totally clear about these latter results (details can be found in [5]). In the framework of second-order arithmetic, we are integrating continuous functions (with a modulus of uniform continuity) defined on closed bounded intervals of the real line. These continuous functions are given by certain *sets* of quintuples (following the standard work in reverse mathematics). We will use the fact that if $g : [\alpha, \beta] \mapsto \mathbb{R}$ is a continuous function with a modulus of uniform continuity, then the map $f : [\alpha, \beta] \mapsto \mathbb{R}$ defined by $f(\gamma) = \int_\alpha^\gamma g(t)\, dt$ is also a continuous function and, hence, given by a suitable set of quintuples. This has the consequence that relations of the form $\int_\alpha^a g(t)dt =_{\mathbb{R}} b$ or $\int_\alpha^a g(t)dt <_{\mathbb{R}} b$ between dyadic natural numbers a and b are, respectively, given by Π^0_1 and Σ^0_1 formulas. More precisely, they can be given by formulas of the form $\forall x[\langle x, a, b\rangle \in Z]$ and $\exists x[\langle x, a, b\rangle \in Z]$, for a suitable set Z. These facts will be used in the sequel.

2 A lemma on minimization

We need a preliminary result in order to prove the implication $integration \rightarrow counting$. We remind the reader that the amount of induction present in BTFA is induction *on notation* for Σ^b_1-formulas. However, in the presence of $integration$, stronger forms of induction are available. For instance, plain *induction for Σ^b_1-formulas*:

Lemma 2.1 BTFA $+ \, integration \vdash \varphi(0) \wedge \forall x \in \mathbb{N}_2(\varphi(x) \rightarrow \varphi(x+1)) \rightarrow \forall x \in \mathbb{N}_2\, \varphi(x)$*, where* φ *is a* Σ^b_1*-formula, possibly with second-order parameters.*

This is the amount of induction that characterizes Buss' theory T^1_2 (cf. [1]). It is known (see [10] or [2]) that this form of induction follows from the *minimization principle*:

$(\star)\ \forall a \in \mathbb{N}_2 \forall X(a \in X \rightarrow \exists b \in \mathbb{N}_2(b \in X \wedge \forall x \in \mathbb{N}_2(x < b \rightarrow x \notin X))).$

Therefore, it is sufficient to prove:

Sublemma *Over* BTFA, $integration$ *implies the minimization principle* $(\star)$.

Proof. We reason inside BTFA. Suppose that we have $integration$. Let X be a set of elements in $\mathbb{N}_2$ and $a \in X$. Consider Φ_X the continuous function associated to the set X according to the introductory section. Since Φ_X is a continuous total function on $[0, a+1]$ with a modulus of uniform continuity, we know that the function $f : [0, a+1] \mapsto \mathbb{R}$, defined by $f(\gamma) = \int_0^\gamma \Phi_X(t)dt$, is well-defined and continuous. We have:

$f(0) = 0$ and

$f(a+1) = \int_0^{a+1} \Phi_X(t)dt = \int_0^a \Phi_X(t)dt + \int_a^{a+1} \Phi_X(t)dt \geq 1.$

Note that $\int_a^{a+1} \Phi_X(t)dt = 1$ because a is in X. By *Bolzano's intermediate value theorem* (cf. [5]), there is a real number α such that $\alpha \in]0, a+1] \wedge f(\alpha) = 1$. Let $m \in \mathbb{N}_2$ be such that $m < \alpha \leq m+1$. We claim that:

($) There is at least one dyadic natural number less than or equal to m in X.

Suppose (in order to get a contradiction) that $\forall k \in \mathbb{N}_2(k \leq m \rightarrow k \notin X)$. Under these circumstances, $\int_0^{m+1} \Phi_X(t)dt = \int_0^{m+1} 0 = 0$ and $\int_0^{m+1} \Phi_X(t)dt = \int_0^\alpha \Phi_X(t)dt + \int_\alpha^{m+1} \Phi_X(t)dt \geq 1$, a contradiction. Now, take $k \leq m$ in X. This element is, in fact, the least element of X. For, suppose that there is also $x < k$ in X. Then $x+1 \leq k \leq m < \alpha$. Hence,

$$1 = \int_0^\alpha \Phi_X(t)dt = \int_0^x \Phi_X(t)dt + \int_x^{x+1} \Phi_X(t)dt + \int_{x+1}^k \Phi_X(t)dt + \int_k^\alpha \Phi_X(t)dt > 1,$$

because $\int_x^{x+1} \Phi_X(t)dt = 1$ and $\int_k^\alpha \Phi_X(t)dt > 0$ (since $k \in X$ and $k < \alpha$). □

3 Integration implies counting

Proposition 3.1 BTFA $\vdash$ *integration* $\to$ *counting*.

We reason inside BTFA. Suppose that there is $integration$ and fix $X \subseteq \mathbb{N}_2$ and $a \in X$. Take n a tally number such that $a + 2 \leq 2^{n-2}$ and consider Φ_X the continuous function associated to X according to the introductory section. Let f be the continuous total function defined in $[0, a+1]$ by $f(\gamma) = \int_0^\gamma \Phi_X(t)dt$. Such a function exists by hypothesis. Since every real number can be approximated to within any degree of accuracy by a dyadic rational number (these were suitably defined in [5] as certain strings, and are denoted as elements of $\mathbb{D}$), we have

$$\forall x \leq a + 1 \exists q \in \mathbb{D} \, | \int_0^x \Phi_X(t)dt - q| < \tfrac{1}{2^{n+1}}.$$

It *follows* from the discussion at the end of the introductory section that the relation between x and q given by $| \int_0^x \Phi_X(t)dt - q| < \frac{1}{2^{n+1}}$ is given by a Σ^0_1-formula, i.e., it is equivalent to $\exists w[\langle w, x, q\rangle \in Z]$, for some set Z. Therefore, $\forall x \leq a + 1 \exists q \exists w(q \in \mathbb{D} \wedge \langle w, x, q\rangle \in Z)$. By the *bounded collection scheme* (a principle of BTFA), there is $s \in \mathbb{N}_2$ such that

$$(\star) \quad \forall x \leq a + 1 \exists q \preccurlyeq s \exists w \leq s \, (q \in \mathbb{D} \wedge \langle w, x, q\rangle \in Z).$$

Here $q \preccurlyeq s$ means that q has (string) length less than or equal to the (string) length of s. Now, consider the following Σ^b_1-formula:

$$\varphi(x, j) \,:\equiv\, x \in \mathbb{N}_2 \wedge j \in \mathbb{N}_2 \wedge x \leq a + 1 \wedge \exists q \preccurlyeq s \exists w \leq s(q \in \mathbb{D} \wedge \langle w, x, q\rangle \in Z \wedge |q - j| < \tfrac{x+1}{2^n}).$$

Fact 1 BTFA $+$ $integration$ *proves:*

(a) $\forall x \leq a + 1 \, \exists j \leq x \, \varphi(x, j)$ *(existence);*

(b) $x \leq a + 1 \wedge \varphi(x, j) \wedge \varphi(x, i) \to j = i$ *(uniqueness).*

If this is shown, then $\varphi(x, j)$ is equivalent to the Π^b_1-formula $x \leq a + 1 \wedge \forall i \leq x(\varphi(x, i) \to j = i)$. By Δ^0_1-comprehension (actually, less than that), within BTFA it is ensured the existence of the set $\{\langle x, j\rangle : \varphi(x, j)\}$. Of course, this set is a *function* h defined on the dyadic natural numbers less than or equal to $a + 1$ and taking values in $\mathbb{N}_2$.

Fact 2 *h is the counting function of the set X up to a.*

Proof of Fact 1. Existence is proved by *plain* induction on x. Observe that we have argued in the previous section that this form of induction is available to us. Firstly, we must show that $\varphi(0, 0)$. By ($\star$), take $q \in \mathbb{D}$ and $w \in \mathbb{N}_2$ such that $q \preccurlyeq s$, $w \leq s$ and $\langle w, 0, q\rangle \in Z$. Thus, $| \int_0^0 \Phi_X(t)dt - q| < \frac{1}{2^{n+1}}$, implying $|q - 0| < \frac{1}{2^n}$.

To prove the induction step, fix $x \leq a$ and suppose, by induction hypothesis, that there is $j \leq x$ such that $\varphi(x, j)$. As a consequence, there is $q' \in \mathbb{D}$ such that $| \int_0^x \Phi_X(t)dt - q'| < \frac{1}{2^{n+1}}$ and $|q' - j| < \frac{x+1}{2^n}$. By ($\star$), let $q \in \mathbb{D}$ and $w \in \mathbb{N}_2$ such that $q \preccurlyeq s$, $w \leq s$ and $\langle w, x + 1, q\rangle \in Z$. There are two cases to consider: either $x \in X$ or $x \notin X$. In the first case, $\int_x^{x+1} \Phi_X(t)dt = 1$. Therefore,

$$\begin{aligned}
|q - (j + 1)| &\leq |q - \int_0^{x+1} \Phi_X(t)dt| + | \int_0^{x+1} \Phi_X(t)dt - (q' + 1)| + |q' - j| \\
&< \frac{1}{2^{n+1}} + | \int_0^x \Phi_X(t)dt - q'| + |q' - j| \\
&\leq \frac{1}{2^{n+1}} + \frac{1}{2^{n+1}} + \frac{x + 1}{2^n} = \frac{x + 2}{2^n}.
\end{aligned}$$

We have shown that $\varphi(x + 1, j + 1)$. In case $x \notin X$, a similar argument shows that $\varphi(x + 1, j)$.

To prove the uniqueness condition, suppose that $\varphi(x, j)$ and $\varphi(x, i)$ with $x \leq a+1$. Hence, there are $q, q' \in \mathbb{D}$ such that

$|\int_0^x \Phi_X(t)dt - q| < \frac{1}{2^{n+1}}$; $|q-j| < \frac{x+1}{2^n}$; $|\int_0^x \Phi_X(t)dt - q'| < \frac{1}{2^{n+1}}$; and $|q'-i| < \frac{x+1}{2^n}$.

We get,

$$\begin{aligned} |j-i| &\leq |j-q| + |q - \int_0^x \Phi_X(t)dt| + |\int_0^x \Phi_X(t)dt - q'| + |q'-i| \\ &< \frac{x+1}{2^n} + \frac{1}{2^{n+1}} + \frac{1}{2^{n+1}} + \frac{x+1}{2^n} = \frac{2x+3}{2^n} \leq \\ &\leq \frac{2(a+1)+3}{2^n} = \frac{2(a+2)}{2^n} + \frac{1}{2^n} \leq \frac{2^{n-1}}{2^n} + \frac{1}{2^n} < 1. \end{aligned}$$

Note that n was taken so that $a+2 \leq 2^{n-2}$. Since $j, i \in \mathbb{N}_2$, it follows that $i = j$.

Proof of Fact 2. This is a consequence of the existence proof of the previous fact.

Acknowledgements Fernando Ferreira was partially supported by CMAF, POCI/FCT and FEDER. Gilda Ferreira was supported by a PhD grant from FCT, POCI/FCT, FEDER and CMAF.

References

[1] S. Buss. *Bounded arithmetic*. PhD thesis, Princeton University, New Jersey, 1985. A revision of this thesis was published by Bibliopolis (Naples) in 1986.

[2] F. Ferreira. *Polynomial time computable arithmetic and conservative extensions*. PhD thesis, Pennsylvania State University, Pennsylvania, 1988.

[3] F. Ferreira. A feasible theory for analysis. *The Journal of Symbolic Logic*, vol. 59 (3), pp. 1001–1011, 1994.

[4] G. Ferreira. *Sistemas de análise fraca para a integração*. PhD thesis (in Portuguese), Universidade de Lisboa, Portugal. Forthcoming.

[5] A. Fernandes and F. Ferreira. Groundwork for weak analysis. *The Journal of Symbolic Logic*, vol. 67 (2), pp. 557–578, 2002.

[6] P. Hájek and P. Pudlák. *Metamathematics of first-order arithmetic*. Perspectives in Mathematical Logic, Springer-Verlag, 1993.

[7] J. Johannsen and C. Pollett. On proofs about threshold circuits and counting hierarchies (extended abstract). *Thirteenth Annual IEEE Symposium on Logic in Computer Science*, pp. 444–452,1998.

[8] K.-I. Ko. *Complexity theory of real functions*. Birkhäuser, Boston, 1991.

[9] K.-I. Ko and H. Friedman. Computational complexity of real functions. *Theoretical Computer Science*, vol. 20, pp. 323–352, 1982.

[10] J. Krajíček. *Bounded arithmetic, propositional logic, and complexity theory*, Encyclopedia of Mathematics and its Applications, vol. 60, Cambridge University Press, 1995.

[11] W. Sieg. Hilbert's Program sixty years later. *The Journal of Symbolic Logic*, vol. 53, pp. 338–348, 1988.

[12] S. Simpson. *Subsystems of second order arithmetic*. Perspectives in Mathematical Logic, Springer, Berlin, 1999.

[13] K. Wagner. Some observations on the connection between counting and recursion. *Theoretical Computer Science*, vol. 47, pp. 131–147, 1986.