THE JOURNAL OF SYMBOLIC LOGIC
Volume 68, Number 1, March 2003

AN APPLICATION OF GRAPHICAL ENUMERATION TO PA*

ANDREAS WEIERMANN†

Abstract. For α less than ε_0 let $N\alpha$ be the number of occurrences of ω in the Cantor normal form of α. Further let $|n|$ denote the binary length of a natural number n, let $|n|_h$ denote the h-times iterated binary length of n and let $\mathrm{inv}(n)$ be the least h such that $|n|_h \leq 2$. We show that for any natural number h first order Peano arithmetic, PA, does not prove the following sentence: For all K there exists an M which bounds the lengths n of all strictly descending sequences $\langle \alpha_0, \dots, \alpha_n \rangle$ of ordinals less than ε_0 which satisfy the condition that the Norm $N\alpha_i$ of the i-th term α_i is bounded by $K + |i| \cdot |i|_h$.

As a supplement to this (refined Friedman style) independence result we further show that e.g., primitive recursive arithmetic, PRA, proves that for all K there is an M which bounds the length n of any strictly descending sequence $\langle \alpha_0, \dots, \alpha_n \rangle$ of ordinals less than ε_0 which satisfies the condition that the Norm $N\alpha_i$ of the i-th term α_i is bounded by $K + |i| \cdot \mathrm{inv}(i)$. The proofs are based on results from proof theory and techniques from asymptotic analysis of Polya-style enumerations.

Using results from Otter and from Matoušek and Loebl we obtain similar characterizations for finite bad sequences of finite trees in terms of Otter's tree constant 2.9557652856....

§1. Introduction and motivation. A fascinating result of ordinal analysis is the classification of the provably recursive functions of first order Peano arithmetic PA in terms of the Hardy–Wainer hierarchy $(H_\alpha)_{\alpha<\varepsilon_0}$. If PA proves $\forall x \exists y T(e, x, y)$ for some natural number e, then there exists some $\alpha < \varepsilon_0$ such that $\{e\}$ is elementary recursive in H_α. Moreover, if $\{e_0\} = H_{\varepsilon_0}$ then PA does not prove $\forall x \exists y T(e_0, x, y)$. These classical results can be reformulated neatly in terms of purely combinatorial independence results as follows. For a binary number-theoretic function f let $A(f)$ be the assertion $\forall K \exists M \forall n \forall \alpha_0, \dots, \alpha_n < \varepsilon_0 [\alpha_0 > \dots > \alpha_n \ \& \ \forall i \leq n [N\alpha_i \leq f(K, i)] \implies n \leq M]$ where $N\alpha$ denotes the number of occurrences of ω in the Cantor normal form of α. Then, by the preceding, PA $\nvdash A(f)$ where $f(k, i) := k \cdot i!$. From the mathematical point of view it seems quite natural to investigate whether this result can be sharpened by using functions f which grow slower than $k, i \mapsto k \cdot i!$. According to Simpson [13] (or Smith [14]) Friedman already showed PA $\nvdash A(f)$ where $f(k, i) := k \cdot (i + 1)$ (or even $f(k, i) := k + i$). In this paper we characterize the class of functions f with PA $\nvdash A(f)$ in a nearly optimal way. The proof combines methods from proof theory with methods from pure mathematics[1]. To

Received May 23, 2001; accepted February 17, 2002.

*Research supported by a Heisenberg-Fellowship of the Deutsche Forschungsgemeinschaft.

†The main results of this paper were obtained during the author's visit of T. Arai in Hiroshima. The author would like to thank T. Arai for his generous support.

[1] For carrying out the calculations we have profited from the asymptotic analysis of integer partitions and the hints on asymptotic properties of trees of height less than or equal to 3 given in [7].

0022-4812/03/6801-0002/$2.20

the author it has been a surprise that analytical methods from infinitesimal calculus can be applied to metamathematical issues like unprovability assertions.

Our investigation is inspired by [6] where a related problem in the context of finite trees has been solved. The main result of [6] is strengthened in Section 4.

§2. A proof of the unprovability result.

Conventions. Throughout this paper small Greek letters range over ordinals less than ε_0 and small Latin letters range over non negative integers. By log $(\ln, \log_3)$ we denote the logarithm with respect to base 2 $(\mathrm{e}, 3)$, where e denotes the Euler number $2.71828\ldots = \sum_{n=0}^{\infty} \frac{1}{n!}$. The least natural number greater than or equal to a given non negative real number x is denoted by $\lceil x \rceil$. The greatest natural number smaller than or equal to a given non negative real number x is denoted by $\lfloor x \rfloor$. The binary length $|n|$ of a natural number n is defined by $|n| := \lceil \log(n+1) \rceil$. The h-times iterated length function $|\cdot|_h$ is defined recursively as follows $|x|_0 := x$ and $|x|_{h+1} := ||x|_h|$. Further let $\mathrm{inv}(n)$ be the least natural number h such that $|n|_h \leq 2$. As usual we assume that the ordinals less than ε_0 are available in PA via a standard coding.

In this section we prove the following result.

THEOREM 2.1. *For all natural numbers h,* PA $\nvdash \forall K \exists M \forall n \forall \alpha_0, \ldots, \alpha_n < \varepsilon_0 [\alpha_0 > \ldots > \alpha_n \;\&\; \forall i \leq n [N\alpha_i \leq K + |i| \cdot |i|_h] \Longrightarrow n \leq M]$.

For this purpose it is convenient for us to recall an independence result from [15].

DEFINITION 1. For $x < \omega$ and $\alpha < \varepsilon_0$ let

$$A_\alpha(x) := \max\{A_\beta(x) + 1 : \beta < \alpha \;\&\; N\beta \leq N\alpha + x\}.$$

As usual put $\omega_0(\alpha) := \alpha$ and $\omega_{n+1}(\alpha) := \omega^{\omega_n(\alpha)}$. Further let $\omega_m := \omega_m(1)$.

LEMMA 2.2. 1. $A_\alpha(x) = \max\{n : (\exists \alpha_0, \ldots, \alpha_n < \varepsilon_0)[\alpha = \alpha_0 > \ldots > \alpha_n \;\&\; [(\forall i < n) N\alpha_{i+1} \leq N\alpha_i + x]]\}$.

2. PA $\nvdash \forall K \exists n A_{\omega_K}(1) = n$. *Moreover $K \mapsto A_{\omega_K}(1)$ eventually dominates every provably recursive function of* PA.

PROOF. See, for example, [15]. ⊣

DEFINITION 2. For natural numbers k and h define

$$S_k^h := \{\alpha < \omega_h : N\alpha = k\}$$

and let s_k^h be the number of elements in S_k^h. Moreover let

$$S_{\leq k}^h := \{\alpha < \omega_h : N\alpha \leq k\}$$

and let $s_{\leq k}^h$ be the number of elements in $S_{\leq k}^h$.

Then $s_{\leq k}^h = \sum_{l \leq k} s_l^h$ and we have $s_k^h \leq s_l^h$ for $k \leq l$ and $h > 0$ since if $N\alpha = k$ then $N(\alpha + l - k) = l$ for $l \geq k$. The following lemma (which is provable in RCA_0) yields a partial asymptotic analysis of s_k^h.

LEMMA 2.3 (RCA_0). *For any $h \geq 3$ there exist a constant $C_h > 0$ and a natural number K_h such that $s_k^h \geq 2^{C_h \cdot \frac{k}{|k|_{h-2}}}$ for $k \geq K_h$.*

Using Lemma 2.2 and Lemma 2.3 we can show Theorem 2.1 as follows.

PROOF OF THEOREM 2.1. The idea of the proof is to construct a slowed down long sequence (α_i) from a given long sequence (α_i') which witnesses the definedness of $A_{\omega_m}(1)$ for an appropriate m. The details are as follows.

Let h be given. Let $h' := h + 3$. Since $h' \geq 3$ we may pick $K_{h'}$ and $C_{h'}$ according to Lemma 2.3. Let D be a constant such that

$$|i|_{h'-2} \geq \frac{1}{C_{h'}} \cdot ||i| \cdot |i|_{h'-2}|_{h'-2} \tag{1a}$$

$$|i| \cdot |i|_{h'-2} \geq K_{h'} \tag{1b}$$

and

$$|i|_{h'-2} + 1 \leq |i|_{h'-3} \tag{1c}$$

hold for $i \geq D$.

Let an arbitrary number K be given. Without loss of generality we may assume that $m := m(K) := \lfloor \frac{K-D}{2} \rfloor - 1 \geq h'$.

Assume that $\omega_m = \alpha_0' > \ldots > \alpha_M'$ is a sequence with $M = A_{\omega_m}(1)$, $N\alpha_0' = m+1$ and $N\alpha_{i+1}' \leq N\alpha_i' + 1$ for $0 \leq i < M$. Consider

$$M_i := S^m_{\leq |i| \cdot |i|_{h'-2}}$$

for $i \geq D$. Assume that enum_i is the enumeration function for M_i, i.e., $\mathrm{enum}_i(l)$ is the l-th (with respect to $\leq$) member of M_i. Let $\alpha_i := \omega_m(\alpha_{|i|}') + \mathrm{enum}_i(2^{|i|} - i)$ for $M \geq i > D$ and $\alpha_i := \omega_{m+m} + D - i$ for $i \leq D$. Then $(\alpha_i)_{i \leq M}$ is well-defined. Indeed, by (1a), (1b) and Lemma 2.3 there are at least

$$2^{C_{h'} \frac{|i| \cdot |i|_{h'-2}}{||i| \cdot |i|_{h'-2}|_{h'-2}}} \geq 2^{|i|} \geq i$$

elements in $S^{h'}_{|i| \cdot |i|_{h'-2}}$ hence in M_i for $i \geq D$. Moreover, we have $N\alpha_{|i|}' \leq N\alpha_0' + |i| = m + 1 + |i|$ for $1 \leq i \leq M$. Now (1c) and the definition of m yield $N\alpha_i \leq K + |i| \cdot (|i|_{h'-2} + 1) \leq K + |i| \cdot |i|_{h'-3} = K + |i| \cdot |i|_h$ for $D < i \leq M$. The definition of m further yields $N\alpha_i \leq K + |i| \cdot |i|_h$ for $1 \leq i \leq D$. Thus $N\alpha_i \leq K + |i| \cdot |i|_h$ for $1 \leq i \leq M$. Further we have $\alpha_i < \alpha_j$ for $i > j$. For if $|i| > |j|$ then this holds due to $\alpha_{|j|}' > \alpha_{|i|}'$ and if $|i| = |j|$ then $M_i = M_j$ and $2^{|i|} - i < 2^{|i|} - j$. Finally, since $K \mapsto A_{\omega_{m(K)}}(1)$ eventually dominates every provably recursive function of PA, the lengths M of the sequences $(\alpha_i)_{i \leq M}$ as a function of K cannot be proved to exist in PA either. ⊣

We are left with proving Lemma 2.3. This will be done in a sequel of sublemmas.

LEMMA 2.4 (RCA_0). *There is a natural number K_2 such that $s_k^2 \geq \mathrm{e}^{2\sqrt{k}}$ for $k \geq K_2$.*

PROOF. Let p_k be the number of integer partitions of k, i.e., the number of ordered tuples $(i_1, \ldots, i_m)$ such that $i_1 \geq \ldots \geq i_m \geq 1$ and $\sum_{n=1}^m i_n = k$. Then $p_k = s_k^2$. Indeed, each integer partition $(i_1, \ldots, i_m)$ of k corresponds to an element $\omega^{i_1 - 1} + \cdots + \omega^{i_m - 1} \in S_k^2$ and vice versa. Now use the partion theorem

$$\lim_{k \to \infty} \frac{p_k \cdot 4\sqrt{3}k}{\mathrm{e}^{\pi\sqrt{\frac{2}{3}k}}} = 1.$$

(See, for example, [4] or Section 2 of [8] for a proof). ⊣

For $h \geq 3$ and natural numbers p, q let $R^h(p, q)$ be the set of ordinals $\alpha < \omega_h$ which have a Cantor normal form $\alpha = \omega^{\alpha_1} + \cdots + \omega^{\alpha_p}$ of length p where $N\alpha_i = q$ for $1 \leq i \leq p$. Further let $r^h(p, q)$ be the number of elements in $R^h(p, q)$. Then $r^h(p, q) \leq s^h_{p\cdot(q+1)}$.

LEMMA 2.5 (RCA_0). *There exists a natural number K_3 such that $s^3_k \geq 2^{\frac{k}{|k|}}$ for all $k \geq K_3$.*

PROOF. [2] For any choice of p and q with $p \cdot (q+1) \leq k$ we have $r^3(p, q) \leq s^3_k$. Thus it suffices to find a lower bound for $r^3(p, q)$ for appropriate p and q.

Let $p := p(k) := \lfloor \frac{k}{|k|^2} \rfloor$ and $q := q(k) := |k|^2 - 1$. Then, of course, $p\cdot(q+1) \leq k$ and $s^3_k \geq s^3_{p\cdot(q+1)} \geq r^3(p, q)$. There exists a natural number K_3 such that for $k \geq K_3$ the following holds

$$\sqrt{q} \cdot p \geq \frac{1}{\log \mathrm{e}} \cdot \frac{k}{|k|} \tag{3a}$$

since $\lim_{k\to\infty} \frac{\sqrt{q}\cdot p}{\frac{k}{|k|}} = 1$,

$$(\log(\mathrm{e}) - 1) \cdot \sqrt{q} \cdot p \geq |p| \tag{3b}$$

since $\lim_{k\to\infty} p(k) = +\infty$

$$\sqrt{q} \cdot p \geq p \cdot |p| \tag{3c}$$

and

$$s^2_q \geq \mathrm{e}^{2\sqrt{q}} \tag{3d}$$

by Lemma 2.4.

We have $r^3(p, q) \geq \frac{(s^2_q)^p}{p!}$ since for fixed p there are at least $(s^2_q)^p$ sequences of length p with entries in S^2_k. Since we have to consider only ordered sequences we have to divide this number by $p!$.

Since $p! \leq (\frac{p}{\mathrm{e}})^p \cdot p \cdot \mathrm{e}$ we obtain by (3) that $r^3(p, q) \geq \frac{(\mathrm{e}^{2\sqrt{q}})^p}{p!} \geq \frac{\mathrm{e}^{2\cdot\sqrt{q}\cdot p}\cdot \mathrm{e}^{p-1}}{p^{p+1}} \geq 2^{2\sqrt{q}\cdot p \cdot \log(\mathrm{e}) - (p+1)\cdot\log(p)} \geq 2^{\log \mathrm{e} \cdot \sqrt{q} \cdot p} \geq 2^{\frac{k}{|k|}}$. ⊣

PROOF OF LEMMA 2.3. By induction on $h \geq 3$. The case $h = 3$ is done in Lemma 2.5. Assume now that the assertion holds for $h - 1 \geq 3$. For any choice of p and q with $p \cdot (q+1) \leq k$ we have $r^h(p, q) \leq s^h_k$. Thus it suffices to find a lower bound for $r^h(p, q)$ for appropriate p and q. Let $p := p(k) := \lfloor \frac{k}{|k|^2} \rfloor$ and $q := q(k) := |k|^2 - 1$. Then, of course, $p \cdot (q+1) \leq k$. Let $r := r^h(p, q)$. There exists a natural number K_h such that for $k \geq K_h$ the following holds

$$p \cdot \frac{q}{|q|_{h-3}} \geq \frac{3}{4} \frac{k}{|k|_{h-2}}, \tag{4a}$$

$$C_{h-1} \cdot \frac{1}{8} \cdot \frac{q}{|q|_{h-3}} \cdot p \geq |p|, \tag{4b}$$

[2] In this proof we follow a hint to exercise 10.7.6 (e) on p.397 in [7] where a bound on the number of trees of height less than or equal to three which have k leaves is obtained.

(4c) $$C_{h-1}\cdot\frac{1}{8}\cdot\frac{q}{|q|_{h-3}}\geq|p|$$

since $\lim_{k\to\infty}\frac{|k|}{||k|^2-1|+1}=+\infty$ and

(4d) $$s_q^{h-1}\geq 2^{C_{h-1}\cdot\frac{q}{|q|_{h-3}}}$$

due to the induction hypothesis since $\lim_{k\to\infty} q=+\infty$.

The proof has now a similar structure as the proof of the previous lemma. First we have $r\geq\frac{(s_q^{h-1})^p}{p!}$ by a similar reasoning as in the previous proof. Since $p!\leq(\frac{p}{\mathrm{e}})^p\cdot p\cdot\mathrm{e}$ we obtain by (4) that $r\geq\frac{(2^{C_{h-1}\frac{q}{|q|_{h-3}}})^p}{p!}\geq 2^{C_{h-1}\cdot\frac{p\cdot q}{|q|_{h-3}}}\cdot\frac{\mathrm{e}^{p-1}}{p^{p+1}}\geq 2^{C_{h-1}\frac{p\cdot q}{|q|_{h-3}}-(p+1)\log p}\geq 2^{C_{h-1}\cdot\frac{3}{4}\cdot\frac{p\cdot q}{|q|_{h-3}}}\geq 2^{C_{h-1}\cdot\frac{3}{4}\frac{3}{4}\cdot\frac{k}{|k|_{h-2}}}\geq 2^{C_h\cdot\frac{k}{|k|_{h-2}}}$ where $C_h:=\frac{C_{h-1}}{2}$. ⊣

The proof shows that we may put $C_h:=(\frac{1}{2})^{h-3}$.

§3. **Proof of the provability assertion.** In this section we show the following theorem. (Recall that $\mathrm{inv}(i)$ is the least h such that $|i|_h\leq 2$.)

THEOREM 3.1. PRA $\vdash\forall K\exists M\forall n\forall\alpha_0,\dots,\alpha_n<\varepsilon_0\big[\alpha_0>\dots>\alpha_n\ \&\ \forall i\leq n\ [N\alpha_i\leq K+|i|\cdot\mathrm{inv}(i)]\implies n\leq M\big]$.

COROLLARY 3.2. PRA $\vdash\forall K\exists M\forall n\forall\alpha_0,\dots,\alpha_n<\varepsilon_0\big[\alpha_0>\dots>\alpha_n\ \&\ \forall i\leq n\ [N\alpha_i\leq K+|i|\cdot K]\implies n\leq M\big]$.

Theorem 3.1 follows from the following Lemma. (Recall that $s^h_{\leq k}$ is the number of elements in $S^h_{\leq k}$. Moreover let $\log_3^{n+1}(k)=\log_3(\log_3^n(k))$ where $\log_3^1(k)=\log_3(k)$) and similarly let $\ln^{n+1}(k)=\ln(\ln^n(k))$ where $\ln^1(k)=\ln(k)$).

LEMMA 3.3. *Let $h\geq 3$. There exists a constant $C_h>0$ such that for all k with $\log_3^{h-2}(k)\geq 1$ we have $s^h_{\leq k}\leq 2^{C_h\cdot\frac{k}{\log_3^{h-2}(k)}}$.*

PROOF OF THEOREM 3.1. We argue informally in PRA while assuming that the proof of Lemma 3.3 can be formalized in RCA_0 so that the assertion of Lemma 3.3 holds in PRA. Let $3_0(k):=k$ and $3_{m+1}(k):=3^{3_m(k)}$. Assume that K is given. Choose C_K according to Lemma 3.3. Let $N:=3_K(K+C_K)$. Assume that we have given a sequence $\alpha_0>\dots>\alpha_n$ with $N\alpha_i\leq K+|i|\cdot\mathrm{inv}(i)$ for $i\leq n$. We claim that $n\leq N$. Otherwise $\omega_{K-1}>\alpha_1>\dots>\alpha_{N+1}$ would be a sequence with $N\alpha_i\leq K+|N+1|\cdot\mathrm{inv}(N+1)$ for $1\leq i\leq N+1$. By Lemma 3.3 $N+1$ is bounded by

$$2^{C_K\cdot\frac{K+|N+1|\cdot\mathrm{inv}(N+1)}{\log_3^{K-2}(K+|N+1|\cdot\mathrm{inv}(N+1))}}\leq 2^{C_K\cdot\frac{K+(3_{K-1}(K+C_K)+1)\cdot 2\cdot(K+C_K)}{3^{K+C_K}}}<N.$$

Contradiction. ⊣

PROOF OF LEMMA 3.3. Let t^h_k ($t^h_{\leq k}$) be the number of finite rooted trees which have height bounded by h and which have k (at most k) nodes. It is easily seen that the number of elements in $S^h_{\leq k}$ is bounded by $t^h_{\leq k}+1$. Indeed, to any α in this set we define inductively a tree as follows. If $\alpha=0$ then $T(\alpha)$ consist of a singleton tree. Assume that α has the Cantor normal form $\omega^{\alpha_1}+\dots+\omega^{\alpha_n}$. Assume that we assigned inductively trees $T(\alpha_1),\dots,T(\alpha_n)$ to $\alpha_1,\dots,\alpha_n$. Then we assign to α the rooted tree with immediate subtrees $T(\alpha_1),\dots,T(\alpha_n)$. For different ordinals we obtain different non isomorphic trees. If $\alpha < \omega_K$ then the height of $T(\alpha)$ is bounded by K and if $N\alpha \leq k$ then $T(\alpha)$ has at most $k+1$ nodes.

Now we want to obtain non trivial bounds on $t^h_{\leq k}$. For this we first compute bounds on t^h_k.[3] Let T^h be the generating function for the sequence $(t^h_k)_{k=0}^\infty$. Thus $T^h(x) = \sum_{k=0}^\infty t^h_k \cdot x^k = \sum_{n=1}^\infty t^h_k \cdot x^k$ since $t^h_0 = 0$. Let p_j denote the number of integer partitions of j, i.e., the number of sequences $(i_1, \dots, i_k)$ with $i_1 \geq \dots \geq i_k \geq 1$ and $i_1 + \cdots + i_k = j$. Then, $T^2(x) = x \cdot \sum_{j=1}^\infty p_j \cdot x^j + \frac{x}{1-x}$ since trees of height 2 correspond to integer partitions in a unique fashion and trees of height 1 correspond uniquely to natural numbers.

According to [11] we have

$$T^{h+1}(x) = \sum_{n=1}^\infty t^{h+1}_n \cdot x^n = x \cdot e^{\sum_{j=1}^\infty \frac{T^h(x^j)}{j}} = x \cdot \prod_{j=1}^\infty \frac{1}{(1-x^j)^{t^h_j}}. \tag{5}$$

for all $x \in]0,1[$.

Let $e_0(k) := k$ and $e_{m+1}(k) := e^{e_m(k)}$. We prove by induction on h that for any $h \geq 2$ there is a constant D_h such that for every $x \in]0,1[$

$$\ln\left(\frac{T_h(x)}{x}\right) \leq e_{h-2}\left(\frac{D_h}{1-x}\right) \tag{6}$$

and extract bounds on t^h_k from this afterwards. The assertion holds for $h = 2$ since as shown in [7] we have $\ln(\sum_{j=1}^\infty p_j \cdot x^j) \leq \frac{\pi^2}{6} \cdot \frac{x}{1-x}$. Hence $\ln(\frac{T_2(x)}{x}) \leq \frac{3}{1-x}$ and we may put $D_2 := 3$.

By induction hypothesis assume that $\ln(\frac{T_h(x)}{x}) \leq e_{h-2}(D_h \cdot \frac{1}{1-x})$. Then $T^h(x) \leq x \cdot e_{h-1}(D_h \cdot \frac{1}{1-x})$, hence by taking logarithms and expanding $-\ln(1-x^j)$ into its power series we obtain by (5) for $x \in]0,1[$

$$\begin{aligned}
\ln\left(\frac{T^{h+1}(x)}{x}\right) &= \sum_{j=1}^\infty t^h_j(-\ln(1-x^j)) \\
&= \sum_{j=1}^\infty t^h_j \sum_{n=1}^\infty \frac{x^{jn}}{n} \\
&= \sum_{n=1}^\infty \frac{1}{n} \sum_{j=1}^\infty t^h_j x^{nj} \\
&= \sum_{n=1}^\infty \frac{1}{n} T^h(x^n) \\
&\leq \sum_{n=1}^\infty \frac{1}{n} x^n e_{h-1}\left(\frac{D_h}{1-x^n}\right) \\
&\leq \sum_{n=1}^\infty \frac{1}{n} x^n e_{h-1}\left(\frac{D_h}{1-x}\right)
\end{aligned}$$

[3] In what follows we utilize formulas from [11] and some hints provided on p.328 and p.396 in [7].

$$\leq \frac{1}{1-x} \mathrm{e}_{h-1}(\frac{D_h}{1-x})$$
$$\leq \mathrm{e}_{h-1}(\frac{D_h+1}{1-x}).$$

By positivity of the summands involved all calculations are legitimate a posteriori. We then may put $D_{h+1} := D_h + 1$ and the induction is finished. (Note that the radius of convergence of $T^h(x)$ is not less than 1.)

Now let $C_h > D_{h+1}$. Let

$$x := x(n) := 1 - \frac{C_h}{\ln^{h-2}(n)}$$

for large enough n such that $x \in]0, 1[$. Since the coefficients of $T_n^{h+1}(x)$ are all non negative, we obtain by (6)

$$t_n^h \leq \frac{1}{x^n} T^h(x) \leq \frac{1}{x^{n-1}} \mathrm{e}_{h-1}(\frac{D_h}{1-x}).$$

Hence

$$\ln(t_n^h) \leq (-n+1) \cdot \ln(x) + \mathrm{e}_{h-2}(\frac{D_h}{1-x}).$$

Since $\lim_{x \downarrow 0} \frac{-\ln(1-x)}{x} = 1$ we obtain

$$(-n+1) \cdot \ln(x) = (-n+1) \cdot \ln(1 - \frac{C_h}{\ln^{h-2}(n)})$$
$$= \frac{n-1}{n} \frac{-\ln(1 - \frac{C_h}{\ln^{h-2}(n)})}{\frac{C_h}{\ln^{h-2}(n)}} \frac{nC_h}{\ln^{h-2}(n)}$$
$$\sim \frac{nC_h}{\ln^{h-2}(n)}. \tag{7}$$

Moreover

$$\mathrm{e}_{h-2}(\frac{D_h}{1-x}) \leq n^{\frac{D_h}{C_h}} \tag{8}$$

for large n. Hence $t_n^h \leq \mathrm{e}^{(C_h+1)\frac{n}{\ln^{h-2}(n)}}$ for large n by (7) and (8) since $\frac{D_h}{C_h} < 1$.

Let E be a natural number such that $\ln^{h-2} k \geq 1$ for $k \geq E$. From the calculation above we know that for a suitable constant C which does not depend on k. $t_k^h \leq \mathrm{e}^{C\frac{k}{\ln^{h-2}(k)}}$ for all $k \geq E$. Then $t_{\leq k}^h \leq t_{\leq E}^h + \sum_{l=E+1}^{k} t_l^h \leq t_{\leq E}^h + k \cdot \mathrm{e}^{C\frac{k}{\ln^{h-2} k}} \leq \mathrm{e}^{C'\frac{k}{\ln^{h-2}(k)}}$ for a suitable constant C' which does not depend on k. Since $\ln(x) \geq \log_3(x)$ we finally obtain the assertion. ⊣

By refining the the previous calculations one obtains refined Friedman style independence results for the fragments $\mathrm{I}\Sigma_n$ of Peano arithmetic. Using multiplicative number theory it is also possible to obtain related results for PA and $\mathrm{I}\Sigma_n$ in the style of Friedman and Sheard [3] where the ordinals are represented via a Schütte style prime number coding [12]. For familiar theories like ATR_0, ID_1 $\Pi_1^1 - (\mathrm{CA})_0$ one can obtain corresponding theorems. These results will be reported elsewhere.

Notes added in proof. 1. Using deep methods from complex analysis the asymptotic behaviour of t_k^h has been determined in more detail by Yamashita in [16].
2. After having seen this manuscript T. Arai proved in [1] the following refinement of Theorem 2.1 and 4.8. Let $a_\alpha(K, i) := K + |i| \cdot |i|_{H_\alpha(i)^{-1}}$ where $H_\alpha(i)^{-1} := \min\{k : H_\alpha(k) \geq i\}$. Then, for $\alpha \leq \varepsilon_0$, PA $\vdash \forall K \exists M \forall n \forall \alpha_0, \dots, \alpha_n < \varepsilon_0$ $[\alpha_0 > \dots > \alpha_n \,\&\, \forall i \leq n[N\alpha_i \leq a_\alpha(K, i)] \implies n \leq M]$ if and only if $\alpha = \varepsilon_0$.

§4. A related unprovability result concerning finite trees. In this section we show that the methods used in the proof of Theorem 2.1 together with results of Otter [9] and Loebl and Matoušek [6] can easily be adapted to prove a related unprovability result concerning the embeddability relation on the set of finite trees. Recall that a finite rooted tree T (with outdegree bounded by a natural number l) is a nonvoid set of nodes such that there is one distinguished node, root(T), called the root of T and the remaining nodes are partitioned into $m \geq 0$ ($l \geq m \geq 0$) disjoint sets $T_1, \dots, T_m$, and each of these sets is a finite rooted tree (with outdegree bounded by l). The trees $T_1, \dots, T_m$ are called the immediate subtrees of T. The cardinality of T is denoted by $|T|$. We say that a finite rooted tree T^1 is embeddable into a finite rooted tree T^2, $T^1 \trianglelefteq T^2$, if either T^1 is embeddable into an immediate subtree of T^2 or if there exist listings $(T_i^1)_{i \leq m}$, $(T_j^2)_{j \leq n}$ of the (multiset of) immediate subtrees of T^1 and T^2 and natural numbers $j_1 < \dots < j_m \leq n$ such that T_k^1 is embeddable into $T_{j_k}^2$ for $1 \leq k \leq m$. Then $\trianglelefteq$ is transitive and $S \trianglelefteq T$ yields $|S| \leq |T|$.

Kruskal's famous tree theorem is as follows.

THEOREM 4.1 (cf. [5]). *For any ω-sequence $(T^i)_{i<\omega}$ of finite rooted trees there exist natural numbers i and j such that $i < j$ and $T^i \trianglelefteq T^j$.*

Using König's Lemma one easily proves the following Lemma.

LEMMA 4.2. *Let f be a binary number-theoretic function. For any K there is an N such that for all sequences $(T^i)_{i \leq N}$ of finite rooted trees with $|T^i| \leq f(K, i)$ for $1 \leq i \leq N$ there exist natural numbers i and j such that $1 \leq i < j \leq N$ and $T^i \trianglelefteq T^j$.*

Assume that the set of finite rooted trees is coded as usual primitive recursively into the set of natural numbers. For a binary function f let $B(f)$ be the following statement (formula) about finite rooted trees:

$$\forall K \exists N \forall T^1, \dots, T^N \big((\forall i \leq N)|T^i| \leq f(K, i) \implies \exists i, j[i < j \,\&\, T^i \trianglelefteq T^j]\big).$$

Then Friedman's celebrated miniaturization result is as follows.

THEOREM 4.3 (cf. [13, 14]). *Let $f(K, i) := K + i$. Then* PA $\nvdash B(f)$. (*In fact we even have* $\mathrm{ATR}_0 \nvdash B(f)$.)

This result has later been sharpended considerably by Loebl and Matoušek as follows.

THEOREM 4.4 (cf. [6]). *Let $f(K, i) := K + 4 \cdot \log(i)$. Then* PA $\nvdash B(f)$.

This result is rather sharp since Loebl and Matoušek obtained the following lower bound.

THEOREM 4.5 (cf. [6]). *Let $f(K, i) := K + \frac{1}{2} \cdot \log(i)$. Then* PRA $\vdash B(f)$.

For a real number r let $f_r(K, i) := K + r \cdot \log(i)$. Then the rational numbers r for which PA $\nvdash B(f_r)$ form a Dedekind cut and one might be interested in the real number c which is represented by this cut. In this section we are going to show that $c = \frac{1}{\log(\alpha)}$ where $\alpha = 2.9557652856\ldots$ is Otter's tree constant (cf. [9]). The real number α is defined as follows. Let $t(0) := 0$, $t(1) := 1$ and $t(i+1) = \frac{1}{i} \cdot \sum_{j=1}^{i}(\sum_{d|j} d \cdot t(d) \cdot t(i-j+1)$. Then $t(i)$ is the number of finite trees with i nodes. Let ρ be the convergence radius of $\sum_{i=0}^{\infty} t(i) \cdot z^i$. Then $\alpha := \frac{1}{\rho}$. Similarly let $t_l(i)$ be the number of finite trees with i nodes and with outdegree bounded by l and let ρ_l be the convergence radius of $\sum_{i=0}^{\infty} t_l(i) \cdot z^i$ and $\alpha_l := \frac{1}{\rho_l}$. Moreover let $t(\leq n)$ $(t_l(\leq n))$ be the number of finite trees (with outdegree bounded by l) with at most n nodes.

THEOREM 4.6 (cf. [9]). 1. *There is a* $\beta > 0$ *such that* $\lim_{n\to\infty} \frac{t(n)}{\alpha^n \cdot n^{-\frac{3}{2}}} = \beta$.

2. *For any* $l \geq 2$ *there is a* $\beta_l > 0$ *such that* $\lim_{n\to\infty} \frac{t_l(n)}{\alpha_l^n \cdot n^{-\frac{3}{2}}} = \beta_l$.

In addition to Otter's result we need the following technical result.

THEOREM 4.7. $\lim_{N\to\infty} \rho_N = \rho$.

PROOF. Obviously we have $\rho_M \geq \rho_N$ for $M \leq N$. Thus $\rho_\infty := \lim_{N\to\infty} \rho_N$ exists and $\rho_\infty \geq \rho$. Assume for a contradiction that $\rho_\infty > \rho$. Then we obtain $\sum_{i=0}^{\infty} t(i) \cdot \rho_\infty^i = +\infty$, hence

$$\sum_{i=0}^{N} t(i) \cdot \rho_\infty^i > 1 \tag{9}$$

for some N.

Otter's paper [9], more precisely equation (11) on page 592 in that paper, yields

$$\sum_{i=0}^{\infty} t_N(i) \cdot \rho_{N+1}^i \leq 1. \tag{10}$$

Thus

$$\sum_{i=0}^{\infty} t_N(i) \cdot \rho_\infty^i \leq 1. \tag{11}$$

This yields by (9) $1 < \sum_{i=0}^{N} t(i) \cdot \rho_\infty^i = \sum_{i=0}^{N} t_N(i) \cdot \rho_\infty^i \leq \sum_{i=0}^{\infty} t_N(i) \cdot \rho_\infty^i \leq 1$. Contradiction. ⊣

THEOREM 4.8 (cf. [2]). *Let* $U(z) = \sum_{i=0}^{\infty} u_i z^i$ *and* $V(z) = \sum_{i=0}^{\infty} v_i z^i$ *be two power series such that for some* $\rho \geq 0$ $\lim \frac{v_{i-1}}{v_i} = \rho$ *and the radius of convergence of* $U(z)$ *is greater than* ρ. *Let* $U(z) \cdot V(z) = \sum_{i=0}^{\infty} w_i z^i$. *Then* $\lim_{i\to\infty} \frac{w_i}{v_i} = U(\rho)$.

THEOREM 4.9 (RCA_0). *Let* $c := \frac{1}{\log(\alpha)}$ *where* α *is Otter's tree constant. Let* r *be a primitive recursive real number and let* f_r *be defined by* $f_r(K, i) := K + r \cdot \log(i)$.

1. *If* $r > c$ *then* $\mathrm{PA} \nvdash B(f_r)$.
2. *If* $r \leq c$ *then* $\mathrm{PRA} \vdash B(f_r)$.

Adapting ideas from the previous section we give a proof of Theorem 4.9 which is based on Otter's result, Theorem 4.6 and the result of Loebl and Matoušek, Theorem 4.4.

For a real number r let $F_r(K)$ be the least N such that for all sequences $(T^i)_{1\leq i\leq N}$ of finite rooted trees with $|T^i| \leq K + r \cdot \log(i)$ for $1 \leq i \leq N$ there exist natural numbers i and j such that $1 \leq i < j \leq N$ and $T^i \trianglelefteq T^j$ and let $F_{\mathrm{LM}} := F_4$. Then the proof of Theorem 4.4 provided in [6] shows that F_{LM} eventually dominates every function which is provably recursive in PA.

We now prove Theorem 4.9.

PROOF OF THEOREM 4.9. Ad 2: By Cauchy's formula for the product of two power series we have $\sum_{n=0}^{\infty} t(\leq n)z^n = \frac{1}{1-z}\sum_{i=0}^{\infty} t(i)z^i$. By employing Theorem 4.6 and Theorem 4.8 we find a natural number D so large that

$$t(\leq n) < \frac{1}{1-\alpha^{-1}}\frac{\alpha^n}{n^{\frac{3}{2}}} \cdot \beta \cdot 1.1. \tag{12}$$

for $n \geq D$. Let K be given. Put

$$M := 2^{8^{K+D}}.$$

Assume that $(T^i)_{i=1}^M$ is a sequence of finite trees such that $|T^i| \leq K + c \cdot \log(i)$ for $1 \leq i \leq M$ and that the T^i are pairwise distinct. Then $|T^i| \leq K + c \cdot \log(M) = K + c \cdot 8^{K+D}$. Thus by (12) we arrive at the contradiction

$$M < \frac{1}{1-\alpha^{-1}}\frac{\alpha^{K+c\cdot 8^{K+D}}}{(K+c\cdot 8^{K+D})^{\frac{3}{2}}} \cdot \beta \cdot 1.1 < M. \tag{13}$$

Ad 1: Since $r > c$ and $\lim_{m\to\infty} \alpha_m = \alpha$ we may pick an m such that $r > \frac{1}{\log(\alpha_m)}$. Then we may choose a rational number r' such that $r > r' > \frac{1}{\log(\alpha_m)}$.

According to assertion 2 of Theorem 4.6 we find a natural number E so large that

$$t_m(n) \geq \alpha_m^n \cdot \beta_m \cdot n^{-\frac{3}{2}} \cdot 0.9 \tag{14}$$

for all $n \geq E$. Let D be so large that for $i \geq D$ the following inequalities hold:

$$\lfloor r' \cdot |i| \rfloor \geq E, \tag{15a}$$

$$2^{\lfloor r'\cdot|i|\rfloor\cdot\log(\alpha_m)} \cdot \beta_m \cdot 0.9 \cdot (\lfloor r' \cdot |i|\rfloor)^{-\frac{3}{2}} \geq 2^{|i|} \tag{15b}$$

and

$$4 \cdot \log(|i|) + r' \cdot |i| \leq r \cdot \log(i). \tag{15c}$$

Now assume that K is given. We may assume that $k := \lfloor \frac{K}{3} \rfloor \geq D$ and $k + m + 4 + D \leq K$. Let $S^1, \ldots, S^{N-1}$ be a finite sequence of finite rooted trees where $N = F_{\mathrm{LM}}(k)$ and $|S^i| \leq k + 4 \cdot \log(i)$ for $1 \leq i \leq N-1$ such that there are no indices i, j with $1 \leq i < j \leq N-1$ and $S^i \trianglelefteq S^j$. Let $\leq$ be a primitive recursive extension of the partial ordering $\trianglelefteq$ on the set of finite rooted trees to a linear ordering. (E.g., one may employ the ordering which is induced by the correspondence between finite rooted trees and ordinals less than ε_0.) Let M_d^m be the set of finite trees T such that T has at most d nodes and the outdegree of T does not exceed m. Let $enum_d^m(l)$ be the l-th member of M_d^m with respect to the linear order $\leq$. Define a sequence of finite trees as follows. Let T^i be the finite rooted tree consisting of a root and two immediate subtrees U^i and V^i. The tree V^i is defined as follows. If $i < D$ let V^i be the uniquely defined (linear) tree with $D - i$ nodes sucht that the outdegree does not exceed 1. If $i \geq D$ let V^i be the tree $\mathrm{enum}^m_{\lfloor r'\cdot|i|\rfloor}(2^{|i|} - i)$. The tree U^i consists by definition of a root and two immediate subtrees U_1^i and U_2^i. U_1^i is S^1 for $i < D$ and $S^{|i|}$ for $i \geq D$. The tree U_2^i consists of a root and $m + 1$ immediate subtrees consisting exactly of one root. Then T^i is well-defined. Indeed, by (14) and (15a) the number of elements in $M^m_{\lfloor r'\cdot|i|\rfloor}$ is for $i \geq D$ at least

$$\alpha_m^{\lfloor r'\cdot|i|\rfloor} \cdot (\lfloor r' \cdot |i|\rfloor)^{-\frac{3}{2}} \cdot \beta_m \cdot 0.9 \geq 2^{|i|}.$$

Moreover (15c) yields

$$|T^i| \leq K + r \cdot \log(i)$$

for $1 \leq i \leq N-1$. Indeed for $i \geq D$ (15c) yields $|T^i| = 1 + |V^i| + |U^i| \leq 1 + \lfloor r' \cdot |i|\rfloor + 1 + k + 4 \cdot \log_2(|i|) + m + 2 \leq K + r \cdot \log_2(i)$. For $i < D$ we obtain $|T^i| = 1 + |V^i| + |U^i| \leq 1 + D + k + 1 + m + 2 \leq K$. We claim that

$$T^i \trianglelefteq T^j$$

does not hold for $1 \leq i < j \leq N-1$. Assume for a contradiction that $T^i \trianglelefteq T^j$ for some i, j with $1 \leq i < j \leq N-1$. First we exclude the possibility that T^i is embeddable into an immediate subtree of T^j. Indeed $T^i \trianglelefteq V^j$ is impossible since the outdegree of V^j does not exceed m but the outdegree of T^i does. Now assume that $T^i \trianglelefteq U^j$. Here we have to distinguish again some cases. The case $T^i \trianglelefteq U^j_2$ is impossible since $|T^i| > |U^j_2|$. If $T^i \trianglelefteq U^j_1$ then $U^i_1 \trianglelefteq U^i \trianglelefteq T^i \trianglelefteq U^j_1$. Hence $U^i_1 = U^j_1$ by the choice of the sequence $(S^i)_{i=1}^{N-1}$. But then $|T^i| > |U^j_1|$ contradicting $T^i \trianglelefteq U^j_1$. Therefore $T^i \trianglelefteq U^j$ yields that U^i is embedabble into an immediate subtree of U^j. $U^i \trianglelefteq U^j_2$ is excluded for cardinality reasons. $U^i \trianglelefteq U^j_1$ yields $U^i_1 \triangleleft U^i \trianglelefteq U^j_1$ hence $U^i_1 = U^j_1$ but then $|U^i| > |U^j_1|$ in contradiction to $U^i \trianglelefteq U^j_1$. Thus the case $T^i \trianglelefteq U^j$ does not occur and T^i is not embeddable into an immediate subtree of T^j.

Therefore $T^i \trianglelefteq T^j$ yields that U^i is embeddable into an immediate subtree of T^j. $U^i \trianglelefteq V^j$ is impossible since the outdegree of V^j does not exceed m but the outdegree of U^i does. Therefore $U^i \trianglelefteq U^j$ and hence necessarily $V^i \trianglelefteq V^j$ also. $U^i \trianglelefteq U^j_2$ is impossible since $|U^i| > |U^j_2|$. If $U^i \trianglelefteq U^j_1$ then $U^i_1 \trianglelefteq U^i \trianglelefteq U^j_1$ hence $U^i_1 = U^j_1$ and $|U^i| > |U^j_1|$ in contradiction to $U^i \trianglelefteq U^j_1$. Hence U^i_1 is embeddable into an immediate subtree of U^j. We claim that $U^i_1 \trianglelefteq U^j_1$. Otherwise $U^i_1 \trianglelefteq U^j_2 = U^i_2 \trianglelefteq U^j_1$. Thus $U^i_1 \trianglelefteq U^j_1$ hence $U^i_1 = U^j_1$ by the choice of $(S^i)_{i=1}^{N-1}$. If $U^i_1 = S^1$ then $U^j_1 = S^1$ and necessarily $i < j < D$. By construction in this case $|V^i| > |V^j|$ in contradiction to $V^i \trianglelefteq V^j$. If $U^i_1 \neq S^1$ then necessarily $D \leq i < j \leq N-1$. We have $U^i_1 = S^{|i|}$ and $U^j_1 = S^{|j|}$ hence $|i| = |j|$. Therefore $2^{|i|} - i > 2^{|j|} - j$ and $V^i = \mathrm{enum}_{\lfloor r'\cdot|i|\rfloor}(2^{|i|} - i) > \mathrm{enum}_{\lfloor r'\cdot|i|\rfloor}(2^{|j|} - j) = V^j$ in contradiction to $V^i \trianglelefteq V^j$.

The argument shows that $F_r(K)$ majorizes $F_{\mathrm{LM}}(\lfloor \frac{K}{3} \rfloor)$ for large K. Thus F_r is not provably recursive in PA since F_{LM} eventually dominates every provably recursive function of PA. Thus PA $\nvdash B(f_r)$. ⊣

In view of [10] we conjecture that the proof above can be adapted to show that for $r > c$ even $\mathrm{ACA}_0 + (\Pi^1_2 - \mathrm{BI}) \nvdash B(f_r)$ where $f_r(K, i) = K + r \cdot \log(i)$.

Related independence results can be obtained for binary trees and Friedman's extension of Kruskal's theorem which is based on the gap condition Moreover we obtained related refined versions of the Paris Harrington theorem, the hydra battle and the Goodstein process. These results will be reported elsewhere.

Questions: 1. Is it possible to use the methods of this paper in the context of bounded arithmetic?
2. Is it possible to give a purely proof-theoretic treatment of the unprovability results obtained in this paper?
3. Is it possible to characterize the slow growing hierarchy via a similar bounding function result?

Acknowledgements. The author is grateful to the referee for his detailed comments which led to an improvement of the paper. He would like to thank J. Matoušek and T. Arai for a discussion and I. Lepper and Gye Sik Lee for their careful proof reading.

REFERENCES

[1] T. Arai, *On the slowly well orderedness of ε_0*, ***Mathematical Logic Quarterly***, vol. 48 (2002), pp. 125–130.

[2] S. N. Burris, ***Number theoretic density and logical limit laws***, Mathematical Surveys and Monographs, vol. 86, American Mathematical Society.

[3] H. Friedman and M. Sheard, *Elementary descent recursion and proof theory*, ***Annals of Pure and Applied Logic***, vol. 71 (1995), pp. 1–45.

[4] G. H. Hardy and S. Ramanujan, *Asymptotic formulae for the distribution of integers of various types*, ***Proceedings of the London Mathematical Society***, vol. 16 (1917), pp. 112–132.

[5] J. B. Kruskal *Well-quasi-orderings, the tree theorem, and Vázsonyi's conjecture*, ***Transactions of the American Mathematical Society***, vol. 95 (1960), pp. 210–225.

[6] M. Loebl and J. Matoušek, *On undecidability of the weakened Kruskal theorem*, ***Logic and combinatorics (Arcata, California, 1985)***, American Mathematical Society, Providence, RI, 1987, pp. 275–280.

[7] J. Matoušek and J. Nešetřil, ***Invitation to Discrete Mathematics***, The Clarendon Press, Oxford, New York, 1998.

[8] D. J. Newman, ***Analytic Number Theory***, Springer, New York, 1998.

[9] R. Otter, *The number of trees*, ***Annals of Mathematics***, vol. 49 (1948), pp. 583–599.

[10] M. Rathjen and A. Weiermann, *Proof-theoretic investigations on Kruskal's theorem*, ***Annals of Pure and Applied Logic***, vol. 60 (1993), pp. 49–88.

[11] J. Riordan, *The enumeration of trees by height and diameter*, ***IBM Journal of Research and Development***, vol. 4 (1960), pp. 473–478.

[12] K. Schütte, ***Proof Theory***, Springer, Berlin, 1977.

[13] S. G. Simpson, *Non-provability of certain combinatorial properties of finite trees*, ***Harvey Friedman's research on the foundations of mathematics***, North-Holland, Amsterdam, 1985, pp. 87–117.

[14] R. Smith, *The consistency strength of some finite forms of the Higman and Kruskal theorems*, ***Harvey Friedman's research on the foundations of mathematics***, North-Holland, Amsterdam, 1985, pp. 119–136.

[15] A. Weiermann, *What makes a (point-wise) subrecursive hierarchy slow growing?*, ***Sets and proofs (Leeds, 1997)***, Cambridge University Press, Cambridge, 1999, pp. 403–423.

[16] M. Yamashita, *Asymptotic estimation of the number of trees*, ***Transactions of the Institute of Electronics and Communication Engineering of Japan***, vol. 62-A (1979), pp. 128–135, in Japanese.

INSTITUT FÜR MATHEMATISCHE LOGIK UND GRUNDLAGENFORSCHUNG
DER WESTFÄLISCHEN WILHELMS-UNIVERSITÄT MÜNSTER
EINSTEINSTR. 62, D-48149 MÜNSTER, GERMANY
E-mail: weierma@math.uni-muenster.de