Brief introduction to unprovability

Andrey Bovykin [*]

Abstract

The article starts with a brief survey of Unprovability Theory as of autumn 2006. Then, as an illustration of the subject's model-theoretic methods, we re-prove exact versions of unprovability results for the Paris-Harrington Principle and the Kanamori-McAloon Principle using indiscernibles. In addition, we obtain a short accessible proof of unprovability of the Paris-Harrington Principle. The proof employs old ideas but uses only one colouring and directly extracts the set of indiscernibles from its homogeneous set. We also present modified, abridged statements whose unprovability proofs are especially simple. These proofs were tailored for teaching purposes.

The article is intended to be accessible to the widest possible audience of mathematicians, philosophers and computer scientists as a brief survey of the subject, a guide through the literature in the field, an introduction to its model-theoretic techniques and, finally, a model-theoretic proof of a modern theorem in the subject. However, some understanding of logic is assumed on the part of the readers.

The intended audience of this paper consists of logicians, logic-aware mathematicians and thinkers of other backgrounds who are interested in unprovable mathematical statements. The paper starts with a brief survey, listing many important achievements and directions of the subject. Most of the results speak for themselves and we omit a discussion of how they are interrelated as well as the story of the subject's big questions, goals, exciting conjectures and dreams which is presumed to be partly known to the readers. The survey is biased towards the Paris-Harrington Principle and its exact versions (understanding this topic is an excellent first step for anyone who decides to study unprovability). In the second part of the paper the reader will find full and very accessible unprovability proofs developed for teaching purposes: the optimal unprovability proof of the Paris-Harrington Principle, two abridged statements whose unprovability proofs are simplest possible and, finally, a model-theoretic proof of some threshold results (exact unprovability results).

1 Brief survey

Peano Arithmetic

Peano Arithmetic (PA) is a first-order theory in the language of arithmetic $\mathcal{L} = \{+, \times, <, 0, 1\}$ that consists of the following axioms: associativity and commutativity of $+$ and $\times$, their neutral elements are 0 and 1 respectively, distributivity, discrete linear order axioms for $<$ (total order, there is a first element 0, no last element, every element has an immediate successor, every nonzero element has an immediate predecessor), 1 is the successor of 0, $x < y \rightarrow x + z < y + z$ and the induction scheme: for every $\mathcal{L}$-formula $\varphi(x, \overline{y})$, we have an axiom $\forall \overline{y}[\ \varphi(0, \overline{y}) \wedge \forall x(\varphi(x, \overline{y}) \rightarrow \varphi(x + 1, \overline{y})) \rightarrow \forall x \varphi(x, \overline{y})\ ]$. In this article we shall deal only with Peano Arithmetic and some of its subsystems but for further study of the subject,

[*] The author was supported by Netherlands travel grant NWO 613.080.000 to participate in a series of workshops in Utrecht University.

beyond this article, the readers will eventually need to explore the whole story surrounding the notions of logical strength and consistency strength of theories and understand the picture of the scale of consistency strength. Although not required to understand the current article, a list of theories that are important in the subject (stretching all way from $I\Delta_0$ and $I\Delta_0(\exp)$ to the strongest extensions of ZFC) can be found on pages 39-40 in [37].

All concrete mathematics of the past can be conducted in Peano Arithmetic

It is common to identify theorems of PA with 'finite mathematics', that is the world of mathematical theorems that can be formulated in $\mathcal{L}$ and whose proof does not require the use of any notion of 'infinite set' in an essential way. Theorems of finite mathematics include the table of derivatives, the table of integrals of elementary functions, for every arithmetically definable complex function f, "if f is analytic on $\mathbb{C}$ and bounded then f is a constant", "in $\mathbb{R}^{24}$, there is a way to place 196560 non-overlapping unit spheres that touch the unit sphere", "the sum $\sum_{\substack{p,p+2\\ \text{both prime}}} \frac{1}{p}$ converges", etc. (It is an exercise to check that all these statements can be formulated in the language $\mathcal{L}$ and their usual proofs can be conducted in Peano Arithmetic.) We give these examples in order to illustrate the extent of what is meant by 'finite mathematics' and show that in this understanding, 'finite mathematics' embraces not only finite combinatorial manipulations but all imaginable mathematics whose objects can be somehow finitely approximated or finitely encoded, including everyday 'continuous' mathematics and many branches of mathematics that seem to use notions beyond that of a natural number but will usually have a way to avoid it by approximations and coding.

Paris-Harrington Principle

Since Gödel's Incompleteness Theorems, for almost half a century logicians did not have examples of PA-unprovable statements that would not refer to diagonalisation or other logicians' tricks. The first PA-unprovable statements of 'mathematical' character (not referring to arithmetisation of syntax and provability) appeared in 1976 in the work of J. Paris (building upon joint work with L. Kirby [52]) and led to the formulation in [69] of the Paris-Harrington Principle (denoted PH): "for any numbers m, n and c, there exists a number N such that for every colouring f of m-subsets of $\{0, 1, \ldots, N-1\}$ into c colours, there is an f-homogeneous $H \subset \{0, 1, \ldots, N-1\}$ of size n such that $|H| > \min H$". This statement PH is not provable in Peano Arithmetic.

Many statements equivalent to PH have been studied: the Hercules-Hydra battle and termination of Goodstein sequences by L. Kirby and J. Paris [72], the flipping principle of L. Kirby [54], the arboreal statement by G. Mills [66], P. Pudlák's Principle ([73], [42]), the kiralic and regal principles by P. Clote and K. McAloon [26].

An important PA-unprovable statement was introduced in [47] by A. Kanamori and K. McAloon. A function f in m arguments is called regressive if $f(x_0, x_1, \ldots, x_{m-1}) \leq x_0$ for all $x_0 < x_1 < \cdots < x_{m-1}$. For regressive functions of m arguments, we cannot guarantee existence of a homogeneous set of size $(m+1)$, e.g., for $f(x_0, x_1, \ldots, x_{m-1}) = x_0 - 1$, every set of size $(m+1)$ is not homogeneous. However, we can talk about min-homogeneous sets: a set H is called min-homogeneous if for all $c_0 < c_1 < \cdots < c_{m-1}$ and $c_0 < d_1 < \cdots < d_{m-1}$ in H, $f(c_0, c_1, \ldots, c_{m-1}) = f(c_0, d_1, \ldots, d_{m-1})$. Now, KM is the following statement: "for any numbers m, a and n with $n \geq m$, there exists $b > a$ such that for every regressive function f defined on m-subsets of $[a, b]$, there is a min-homogeneous set $H \subset [a, b]$ of size at least n". The statement KM is unprovable in PA and is equivalent to PH.

Indicators

First independence results grew out of indicator theory (best references are [53], [52] and

[71]). In a model $M \vDash I\Sigma_1$, an initial segment I is called semi-regular if for every $a \in I$ and every function $f\colon [0,a] \to M$ (i.e., an M-coded set of pairs), the image of f is bounded in I. Every semi-regular initial segment satisfies $I\Sigma_1$. An initial segment I is called strong if for every M-coded partition $P\colon [I]^3 \to 2$, there is an M-coded I-unbounded P-homogeneous subset. Every strong initial segment is a model of PA. Other important classes of initial segments (closed under certain externally-described combinatorial operations) include regular (satisfies $B\Sigma_2$), n-extendible (corresponds to $I\Sigma_{n+1}$) and n-Ramsey (corresponds to $I\Sigma_{n+1}$) initial segments. Now, if a Σ_1-definable in $I\Sigma_1$ Skolem function $Y(x,y)$ is such that for all $a < b$ in M, $Y(a,b) > \mathbb{N}$ if and only if there is a strong initial segment between a and b then the statement "for all a and c, there is $b > a$ such that $Y(a,b) > c$" is unprovable in Peano Arithmetic. The function $Y(x,y)$ is called an indicator for strong initial segments because it indicates whether the set $[a,b]$ is large enough to accommodate a strong initial segment. (Similarly for all other kinds of initial segments above.) All early Ramsey-style independence results can be described in this setting. In case of the Paris-Harrington Principle, the function $Y(x,y) =$ the maximal c such that for every colouring $P\colon [x,y]^c \to c$, there is a homogeneous subset H of size $2c$ such that $|H| > \min H$ is an indicator for strong initial segments.

Many early indicator proofs were conducted in terms of a game between two players where Player I tries to ensure that the final initial segment between a and b is, say, strong, and Player II tries to prevent it. The game of finite (nonstandard) length is determined, so it turns out that if a set is large enough then Player I has a winning strategy, otherwise Player II has a winning strategy. Original sources are [52] and [71]. For another example connecting games and independence results, see the general idea and the Peano Arithmetic section in a recent article by P. Pudlák [74]. There is much more left to say about games and independence results, e.g., about unprovability of existence of a winning strategy in certain games.

Indicator theory ideas are also useful in the model-theoretic approach to Reverse Mathematics, in the spirit of the model-theoretic proof by J. Paris and L. Kirby in [52] that RT^3_2 (the infinite Ramsey Theorem for triples and two colours) implies all of PA, thus providing an alternative to the recursion-theoretic approach to these matters. Modern-day examples of such model-theoretic proofs in the style of indicator theory can be found in the articles [14] and [15] by the author and A. Weiermann.

PH **is an arithmetical version of large cardinals**

The historical prototypes of the Paris-Harrigton Principle and the earlier PA-unprovable statements of [70] are large cardinal axioms. In the case of arithmetic, closedness properties postulated by large cardinal axioms correspond to closedness properties of initial segments of models of arithmetic under the (external) combinatorial properties described above: semi-regularity, regularity, strength, extendibility and Ramseyness. This analogy eventually led to the Paris-Harrington principle and was important in the early days of the subject but was later abandoned and almost forgotten. It may be very fruitful to have a fresh look at this analogy, especially having in mind the modern advances in the study of large cardinals. J. Ketonen's manuscripts [49], [51] provide an alternative way to view such analogy, using ordinals, a theme that is closely connected with the fundamental Ketonen-Solovay article [50].

Fragments of Peano Arithmetic

If we restrict the induction scheme to Σ_n-formulas, that is, arithmetical formulas of the form $\exists x_1 \forall x_2 \exists x_3 \ldots \varphi(x_1, x_2, \ldots, x_n, \overline{y})$, where φ is preceded by no more than n quantifiers and itself contains no unbounded quantifiers, then the theory obtained is denoted by $I\Sigma_n$ ("induction for Σ_n-formulas"). Clearly, for every n, $I\Sigma_n \subseteq I\Sigma_{n+1}$ and $\mathrm{PA} = \bigcup_{n=1}^{\infty} I\Sigma_n$.

Also, it is known that for every n, $I\Sigma_n \neq I\Sigma_{n+1}$ since $I\Sigma_{n+1}$ proves $\mathrm{Con}(I\Sigma_n)$.

A useful alternative axiomatisation of $I\Sigma_n$ uses an instance of the least-number principle $\forall\overline{y}[\ \exists x\varphi(x,\overline{y}) \to \exists z(\varphi(z,\overline{y}) \wedge \forall w < z\neg\varphi(w,\overline{y})]$ for every Σ_n-formula $\varphi(x,\overline{y})$.

It is widely believed that all $\mathcal{L}$-theorems of existing mathematics (apart from logicians' discoveries we are talking about in this article) can be proved not only in PA but even in $I\Sigma_2$. It would be surprising if someone managed to find an existing theorem in mathematics that can be formulated in $\mathcal{L}$ but does not have a proof formalisable in $I\Sigma_2$.

For every $k \in \mathbb{N}$, the statement $\mathrm{PH}^{(k+1)}$ defined as "for all n and c, there exists N such that for every colouring f of $(k+1)$-subsets of $\{0,1,\dots,N-1\}$ into c colours, there is an f-homogeneous $H \subset \{0,1,\dots,N-1\}$ of size at least n such that $|H| > \min H$" is $I\Sigma_k$-unprovable [71] and is equivalent to $\mathrm{KM}^{(k+1)}$ (the Kanamori-McAloon Principle for $(k+1)$-subsets) and to $\mathrm{RFN}_{\Sigma_1}(I\Sigma_k)$, the 1-consistency of $I\Sigma_k$: $\forall\varphi \in \Sigma_1\ (\mathrm{Pr}_{I\Sigma_k}(\varphi) \to \varphi)$. (It says "for all Σ_1-statements φ, if $I\Sigma_k$ proves φ then φ holds". In order to write it as a formula, it is necessary to use the satisfaction predicate for Σ_1-formulas. Unprovability of $\mathrm{RFN}_{\Sigma_1}(I\Sigma_k)$ in $I\Sigma_k$ easily follows from Gödel's Second Incompleteness Theorem: put φ to be $\exists x\ x \neq x$ to observe that $\mathrm{RFN}_{\Sigma_1}(I\Sigma_k)$ implies $\mathrm{Con}_{I\Sigma_k}$.) A good exposition of 1-consistency and reflection principles is in the old Smorynski's paper [85], a good exposition of the satisfaction predicate is in Kaye's textbook [48].

Threshold results for PH and KM

Let $\log^{(n)}(x) = \underbrace{\log_2(\log_2 \dots \log_2}_{n \text{ times}}(x))$ and the tower-function $2_n(x)$ be defined as $2_0(x) = x$, $2_{n+1}(x) = 2^{2_n(x)}$. Also, define $\log^*(m)$ as the minimal n such that $2_n(2) \geq m$. The Ramsey number $R(k,c,m)$ is defined as the minimal number such that for any colouring of k-subsets of $\{0,1,2,\dots,R(k,c,m)-1\}$ in c colours, there is a monochromatic subset of size m. For any set X, we write $[X]^n$ for the set of its n-subsets. The set of all n-subsets of $\{a,a+1,\dots,b\}$ will be denoted by $[a,b]^n$, without any confusion. As usual, a natural number N is identified with the set of its predecessors $\{0,1,\dots,N-1\}$.

For every function $F(x)$, define $\mathrm{PH}_F^{(k)}$ as the statement "for all n and c there exists N such that for every $f\colon [N]^k \to c$, there is a homogeneous $H \subseteq N$ of size at least n and such that $F(\min H) < |H|$". We say that f is F-regressive if for all $x_0 < x_1 < \dots < x_{k-1}$, we have $f(x_0,x_1,\dots,x_{k-1}) \leq F(x_0)$. Now define $\mathrm{KM}_F^{(k)}$ as the statement "for all n there exists N such that for every F-regressive f defined on $[N]^k$, there is a min-homogeneous subset of N of size at least n". Also, define PH_F as $\forall k\ \mathrm{PH}_F^{(k)}$ and KM_F as $\forall k\ \mathrm{KM}_F^{(k)}$. It is easy to see that for every strictly increasing F, PH_F implies PH and KM_F implies KM thus making these statements PA-unprovable. A. Weiermann [89] proved that for every n, $\mathrm{PH}_{\log^{(n)}}$ is PA-unprovable but $\mathrm{PH}_{\log^*}$ is PA-provable. In the case of fixed dimension, the story is more complex. The following interesting result was first proved by Gyesik Lee [59] for all $n \neq k-1$ and later completed by L. Carlucci, G. Lee and A. Weiermann [22]: if $k \geq 2$ then

1. if $n \leq k-1$ then $\mathrm{KM}_{\log^{(n)}}^{(k+1)}$ is $I\Sigma_k$-unprovable;

2. if $n > k-1$ then $I\Sigma_1$ proves $\mathrm{KM}_{\log^{(n)}}^{(k+1)}$.

Similar theorems hold for the family $\mathrm{PH}_{\log^{(n)}}^{(k)}$:

1. if $n \leq k$ then $I\Sigma_k$ does not prove $\mathrm{PH}_{\log^{(n)}}^{(k+1)}$ (A.Weiermann [89]);

2. if $n > k$ then $I\Sigma_1$ proves $\mathrm{PH}_{\log^{(n)}}^{(k+1)}$ (G.Lee [59]).

There are even slightly sharper threshold results for this case in [22]. Here is the picture in the case $k = 1$ (see [56]): if A^{-1} is the inverse of the Ackermann function and $\{F_m\}_{m\in\omega}$ is the Grzegorczyk hierarchy of primitive recursive functions then:

1. $I\Sigma_1 \nvdash \mathrm{PH}^{(2)}_{\frac{\log}{A^{-1}}}$;
2. for every $m \in \omega$, $I\Sigma_1 \vdash \mathrm{PH}^{(2)}_{\frac{\log}{F_m^{-1}}}$;
3. $I\Sigma_1 \nvdash \mathrm{KM}^{(2)}_f$, where $f(x) = x^{\frac{1}{A^{-1}(x)}}$;
4. for every $m \in \omega$, $I\Sigma_1 \vdash \mathrm{KM}^{(2)}_{f_m}$, where $f_m(x) = x^{\frac{1}{F_m^{-1}(x)}}$.

In particular, $\mathrm{KM}^{(2)}_{\log}$ is provable but $\mathrm{PH}^{(2)}_{\log}$ is unprovable.

The reason for $I\Sigma_1$-provability of $\mathrm{PH}^{(k+1)}_{\log^{(n)}}$ and $\mathrm{KM}^{(k+1)}_{\log^{(n)}}$ for large n comes from the Erdös-Rado theorem [28], which implies that an upper bound for the Ramsey number $R(m, k+1, m)$ is $2_n(m)$ for some large enough n depending only on k. Given m and c, let $\ell = \max\{m, c\}$. Consider any colouring $f\colon [0, 2_n(\ell)]^{k+1} \to c$. By the Ramsey Theorem, there is $H \subset [0, 2_n(\ell)]$ of size at least ℓ which is f-homogeneous. Also, $\log^{(n)}(\min H) < \log^{(n)}(2_n(\ell)) = \ell \leq |H|$. Thus $I\Sigma_1 \vdash \mathrm{PH}^{(k+1)}_{\log^{(n)}}$.

A similar argument for $I\Sigma_1$-provability of $\mathrm{KM}^{(k+1)}_{\log^{(n)}}$ for large n goes as follows: consider n such that $R(m, k+1, m) < 2_n(m)$ for all m. Let f be a $\log^{(n)}$-regressive function defined on $[0, 2_n(m)]$. Then the image of f is contained in $[0, m]$. Hence there is a homogeneous (thus also min-homogeneous) subset $H \subseteq [0, 2_n(m)]$ of size at least m.

Kruskal's Theorem and well-quasi-orders

Very often an unprovable statement can be viewed as a 'miniaturisation' of an infinitary theorem. A spectacular example of miniaturisation is H. Friedman's Theorem [79] on unprovability of a finite version of Kruskal's Theorem. Define a (nonplane, rooted) tree as a partially ordered set with the least element and such that the set of all predecessors of every point is linearly ordered. Infinite Kruskal's Theorem says: if $T_1, T_2, \ldots$ is an infinite sequence of finite trees then there are $i < j$ such that $T_i \trianglelefteq T_j$, i.e., there is an inf-preserving embedding from T_i into T_j (that is, trees are well-quasi-ordered by $\trianglelefteq$). Friedman's Theorem says that neither the Infinite Kruskal Theorem nor its finite version ('miniaturisation') "for all k there is N such that whenever $\langle T_i\rangle_{i=1}^N$ is a sequence of finite trees such that for all $i \leq N$ we have $|T_i| \leq k + i$ then there are $i < j \leq N$ such that $T_i \trianglelefteq T_j$" is provable in ATR_0, a theory stronger than Peano Arithmetic. Here, $|T|$ is the number of vertices in T. Also, Kruskal's theorem restricted to binary trees is unprovable in ACA_0 (and its finite version unprovable in PA) and Kruskal's theorem for binary trees and two labels unprovable in ATR_0. (There are also important results by L. Gordeev about the strength of Kruskal's theorem with gap-condition and other unprovability results in [39, 40, 41].) It was later shown by M. Loebl and J. Matoušek [60] that if the condition $|T_i| \leq k + i$ is replaced by $|T_i| \leq k + \frac{1}{2}\log_2 i$ then the statement becomes $I\Sigma_1$-provable but for the condition $|T_i| \leq k + 4\log_2 i$, the statement is PA-unprovable. What happens between $\frac{1}{2}$ and 4 was recently resolved by A. Weiermann [88]. Let α be Otter's constant ($\alpha = \frac{1}{\rho}$, where ρ is the radius of convergence of $\sum_{i=0}^{\infty} t_i z^i$, where t_i is the number of finite trees of size i), $\alpha \approx 2.955765\ldots$. Then for any primitive recursive real number r,

1. if $r \leq \frac{1}{\log_2 \alpha}$ then the statement with the condition $|T_i| \leq k + r\log i$ is $I\Sigma_1$-provable;
2. if $r > \frac{1}{\log_2 \alpha}$ then the statement is PA-unprovable.

Similar results have been proved by A. Weiermann for plane trees, binary plane trees, Higman's lemma [92], ordinal notations [90], [88]. Several general theorems have recently been proved by the author [12], giving an answer to the following question: if X is a well-quasi-ordered combinatorial class (i.e., there is a fixed notion of size of objects and for every n there are finitely-many objects of size n) such that for some function f, the statement "for all K there is N such that whenever $X_1, X_2, \ldots, X_N$ are in X and for all $i \leq N$, $|X_i| < K + f(i)$ then for some $i < j \leq N$, $X_i \trianglelefteq X_j$" is unprovable in some theory T, what is the exact threshold function that separates provable and T-unprovable instances of the corresponding well-quasi-orderedness statements for $\mathrm{Seq}(X)$ (the set of all finite sequences of elements of X), $\mathrm{Cyc}(X)$ (finite cycles of elements of X), $\mathrm{Mult}(X)$ (finite multisets of X), $\mathrm{Trees}(X)$ (plane rooted trees labeled by X) and some other compound combinatorial classes that inherit well-quasi-orderedness from X by a usual minimal bad sequence argument? The proofs use Weiermann-style compression techniques and ideas from analytic combinatorics. There should be ordinal-theoretic results closely connected to these theorems (roughly: how the multiset-, sequence-, cycle-, tree- and other constructions transform the maximal order-types of well-ordered linearisations of original well-quasi-orders) but these results, to the author's knowledge, haven't yet been written down by anyone.

Another important example is a theorem by H. Friedman, N. Robertson and P. Seymour on unprovability of the Graph Minor Theorem [32]. For multigraphs G and H, we say that H is a minor of G if H can be obtained from G by a succession of three elementary operations: edge removal, edge contraction and removal of an isolated vertex. The first-order version of their theorem states that the statement "for all k there is N such that whenever $G_1, \ldots, G_N$ is a sequence of finite multigraphs such that for all $i \leq N$ we have $|G_i| \leq k + i$ then for some $i < j \leq N$, G_i is a minor of G_j" is not provable in Π^1_1-CA_0, a very strong subsystem of second-order arithmetic. Multigraphs can well be replaced by simple graphs in this formulation.

For a function f, let the statement GM_f be "for all K there is N such that for any sequence of simple graphs $G_1, G_2, \ldots, G_N$ such that $|G_i| < K + f(i)$, there are $i < j$ such that G_i is a minor of G_j". It is now easy to conjecture that for every primitive recursive real number a, if $a \leq \sqrt{2}$ then $\mathrm{GM}_{a \cdot \sqrt{\log}}$ is $I\Sigma_1$-provable but if $a > \sqrt{2}$ then $\mathrm{GM}_{a\sqrt{\log}}$ is unprovable. So far it has been proved by the author in [12] that if $a \leq \sqrt{2}$ then $\mathrm{GM}_{a \cdot \sqrt{\log}}$ is $I\Sigma_1$-provable and $\mathrm{GM}_{7 \log}$ is PA-unprovable. A certain lemma on graph enumeration is currently missing for the proof of the full conjecture to go through. There are similar conjectures about unprovability thresholds for Kruskal's theorem with gap condition and for graph minor theorem for subcubic graphs and for classes of graphs omitting certain minors. This enterprise has to use analytic combinatorics, graph minor theory and Pólya theory.

Well-quasi-order theory has been extensively studied in the framework of Reverse Mathematics ([83]), the study of logical strength and consistency strength of second-order arithmetical assertions. Reverse Mathematics is the closest relative of first-order unprovability theory and there is exchange of ideas flowing both ways. There are many results, methods and ideas in Reverse Mathematics that are very relevant to the subject but we omit them in this brief survey. It is difficult to draw a strict line between the two subjects and any future exposition of first-order unprovability may have to incorporate some discussion of relevant parts of Reverse Mathematics.

Single tree, single sequence and boolean relation theory

Here is a new way to obtain unprovable statements from existing statements that talk about long sequences of objects by assembling elements of sequences of objects into one single object to talk about, with order (or embeddability) relation between objects in the sequence becoming order (or embeddability) relation between chunks of this single object. For a tree T, let $T[i]$ be the tree of nodes of T of height $\leq i$. It has been proved by H. Friedman in [35] that the statement "for all n and k there is K such that whenever T is a rooted nonplane tree labeled by $\{1, 2, \ldots, n\}$ and every non-leaf has degree k then there are $i < j \leq K$ such that $T[i]$ is inf-preserving, label-preserving and leaf-preserving embeddable into $T[j]$" is equivalent to 1-consistency of Π^1_2-TI_0, a system much stronger than Peano Arithmetic, and even the statement with $k = 2$ already implies 1-consistency of ATR_0. There are also versions that involve a gap-condition. Another application of similar ideas deals with Higman's lemma. The statement "for all m there is K such that for every K-sequence $x_1, x_2, \ldots, x_K$ of elements of $\{1, 2, \ldots, m\}$, there are $i < j < \frac{K}{2}$ such that $x_i, x_{i+1}, \ldots, x_{2i}$ is a subsequence of $x_j, x_{j+1}, \ldots, x_{2j}$" is unprovable in $I\Sigma_2$ (see [36]).

Another series of H. Friedman's results under the general name of "boolean relation theory" can be found in [38]: if we list all (second-order) statements of a certain simple shape (all of them a priori equally simple and natural) and try to classify them according to their truth, some of them turn out to be unprovable in some theories stronger than ZF, e.g. the following statement from [38] is provably in ACA' equivalent to 1-consistency of the theory $\mathrm{ZFC} + \{\text{there is an } n\text{-Mahlo cardinal}\}_{n\in\omega}$: "for any two functions f, $g\colon \mathbb{N}^k \to \mathbb{N}$ such that there are two constants $c, d > 1$ with $c \cdot |\overline{x}| < f(\overline{x}), g(\overline{x}) < d \cdot |\overline{x}|$ for all but finitely many $\overline{x} \in \mathbb{N}^k$, there exist infinite sets $A, B, C \subseteq \mathbb{N}$ such that $A \cup .f(A) \subseteq C \cup .g(B)$ and $A \cup .f(B) \subseteq C \cup .g(C)$", where $|\overline{x}|$ is the maximal element of the k-tuple $\overline{x}$, $f(A)$ is the image of f on k-tuples of A and $A \cup .D$ means the union of A and D together with the statement that they are disjoint.

Sine, zeta-function, diophantine approximation and universality

The results in this section spring from H. Friedman's sine-principle [11]. For every $n \geq 1$ and every function F of one argument, let us introduce the statement SP^n_F: "for all m, there is N such that for any sequence $a_1, a_2, \ldots, a_N$ of rational numbers, there is $H \subseteq A$ of size m such that for any two n-element subsets $a_{i_1} < a_{i_2} < \cdots < a_{i_n}$ and $a_{i_1} < a_{k_2} < \cdots < a_{k_n}$ in H, we have $|\sin(a_{i_1} \cdot a_{i_2} \cdot \ldots \cdot a_{i_n}) - \sin(a_{i_1} \cdot a_{k_2} \cdot \ldots \cdot a_{k_n})| < F(i_1)$". For $n \geq 2$ and any function $F(x)$ eventually dominated by $(\frac{2}{3})^{\log^{(n-1)}(x)}$, the principle SP^{n+1}_F is not provable in $I\Sigma_n$. The proof in [11] uses the Rhin-Viola theorem on irrationality measure of π.

The sine-principle led to a series of exciting developments. What other functions can be taken instead of sine so that the statement would still be unprovable? The following theorem by the author and A. Weiermann [16] describes a large class of such functions. Let f, g and h be three functions such that for any $N \in \mathbb{N}$ and any small $\varepsilon > 0$,

1. h is a periodic function with period a, continuous on its period;

2. f is such that for any $b_1, b_2, \ldots, b_N$, linearly independent over $a\mathbb{Q}$ and any $c_1, c_2, \ldots, c_N \in [0, a)$, there is $x \in \mathbb{Q}$ such that for all $i \in \{1, 2, \ldots, N\}$, $|f(b_i \cdot x) \; mod \; a \; - \; c_i| < \varepsilon$;

3. g is any continuous function on a subset of $\mathbb{R}$ whose image contains $[d, +\infty)$ for some $d \in \mathbb{R}$.

Then the statement "for all m and n, there is N such that for any sequence $\langle a_i \rangle_{i=1}^N$ of rational numbers, there is $H \subseteq N$ of size m such that for any two n-sequences $i_1 < i_2 < \ldots < i_n$ and $i_1 < k_2 < \ldots < k_n$ of natural numbers smaller than N,

$$|h(f(g(a_{i_1} \cdot a_{i_2} \cdot \ldots \cdot a_{i_n}))) - h(f(g(a_{i_1} \cdot a_{k_2} \cdot \ldots \cdot a_{k_n})))| < 2^{-i_1}\text{"}$$

is unprovable in Peano Arithmetic and the versions for fixed n are unprovable in $I\Sigma_{n-1}$. This class of functions includes $\sin(p(x_1 \cdot x_2 \cdot \ldots \cdot x_n))$, $\{p(x_1 \cdot x_2 \cdot \ldots \cdot x_n)\}$, $\left\{\frac{1}{p(x_1 \cdot x_2 \cdot \ldots \cdot x_n)}\right\}$ for any non-constant polynomial p (using H. Weyl's theorem on simultaneous diophantine approximation). In particular, the proof of this general theorem shows that unprovability of the sine-principle can now be demonstrated using simultaneous diophantine approximation instead of an argument involving irrationality measure of π. We have several conjectures that suggest versions of these results in p-adic setting.

Another interesting example deals with the Riemann zeta-function [16]: for any $\sigma \in \mathbb{R}$, consider the statement "for all m and n, there is N such that for any sequence $\langle a_i \rangle_{i=1}^{N}$ of rational numbers, there is $H \subseteq N$ of size m such that for any two n-sequences $i_1 < i_2 < \ldots < i_n$ and $i_1 < k_2 < \ldots < k_n$ of natural numbers smaller than N, $|\zeta(\sigma + i \cdot a_{i_1} \cdot a_{i_2} \cdot \ldots \cdot a_{i_n}) - \zeta(\sigma + i \cdot a_{i_1} \cdot a_{k_2} \cdot \ldots \cdot a_{k_n})| < 2^{-i_1}$". For $\sigma > 1$, the unprovability proof is similar to that for $\sin(x)$, using almost periodicity. For $\sigma \in (\frac{1}{2}, 1]$, the unprovability proof uses a probabilistic argument.

The reason behind unprovability of these statements is that Ramsey-style assertions (in this case the Kanamori-McAloon Principle) can be concealed in this setting by replacing quantification over all possible colourings by the quantifier over all possible sequences of rational numbers. E.g. in the case of the sine-function, what we need is that every function is approximated by sine on some subset: for any $\varepsilon > 0, K, n$ and any function $g\colon [K]^n \to [-1, 1]$, there is a sequence of rational numbers $\langle a_1, a_2, \ldots, a_K \rangle$ such that for any $i_1 < \ldots < i_n \leq K$, $|g(i_1, \ldots, i_n) - \sin(a_{i_1} \cdot \ldots \cdot a_{i_n})| < \varepsilon$. This situation where all possible patterns (e.g. functions, colourings, etc) are already present in one complex object is called universality and there is more to say about using universal objects to reformulate Ramsey-style statements.

The following unprovability results have been proved in [16], using discrete analogues of S. Voronin's universality theorem about the Riemann zeta-function. The statement "for all n there is N such that whenever $\langle a_i \rangle_{i=1}^{N}$ are natural numbers, there is a subset $H \subset N$ of size n such that for all $k < l < m$ in H,

$$\mid \zeta^{(a_k)}(\frac{3}{4} + i \cdot a_l) - \zeta^{(a_k)}(\frac{3}{4} + i \cdot a_m) \mid < \frac{1}{2^k}\text{''}$$

is unprovable in $I\Sigma_1$. The statement "for all n there is N such that whenever $\langle a_i \rangle_{i=1}^{N}$ are natural numbers, there is a subset $H \subset N$ of size n such that for all $k < l < m$ in H,

$$\mid \zeta(\frac{1}{a\,k} + i \cdot a_l) - \zeta(\frac{1}{a\,k} + i \cdot a_m) \mid < \frac{1}{2^k}\text{''}$$

is unprovable in $I\Sigma_1$. Similar results will hold for a wide range of zeta- and L-functions.

Universality, in the broad understanding of the word, will be a rich source of unprovable statements in the future. We want to encode and treat a random colouring of n-tuples, so assertions that involve all possible patterns (or some 'random' patterns) of finite configurations should attract our attention. Nowadays, it should be possible to show unprovability of some (versions of) already existing strong conjectures in, say, number theory. A promising example that begs for an independence proof is Schinzel's hypothesis H (that is rooted in the work of 19th century mathematicians, e.g. Bunyakovskiy): "for any finite collection of irreducible polynomials $P_1(x), P_2(x), \ldots, P_n(x)$ with integer coefficients and such that $\prod_{i \leq n} P_i$ has no fixed prime divisor, there exist infinitely-many integers m such that for all $i \leq n$, $P_i(m)$ are prime". This conjecture is extremely strong (implying the twin-prime conjecture and infinity of the set of primes of the form $n^2 + 1$) and its formulation already provides some necessary ingredients for the unprovability proof. ($I\Sigma_1$-unprovability of hypothesis H was already conjectured in the PhD thesis of A. Woods [93]).

Braids

Braids are very popular and interesting objects in today's mainstream mathematics. The n-strand braid group B_n is a group with the following presentation:

$$B_n = \langle \sigma_1, \ldots, \sigma_{n-1};\ \sigma_i \sigma_j = \sigma_j \sigma_i \quad \text{for } |i - j| \geq 2,\ \sigma_i \sigma_j \sigma_i = \sigma_j \sigma_i \sigma_j \quad \text{for } |i - j| = 1 \rangle.$$

A braid is called positive if it has a representation without σ_i^{-1} for any i.

There are several independence results about braids, relying on P. Dehornoy's left-invariant ordering $\prec$ of positive braids as ω^{ω^ω} and S. Burckel's ordering of n-strand positive braids as $\omega^{\omega^{n-2}}$, see [27]. Some unprovability results immediately follow from $I\Sigma_2$-unprovability of transfinite induction up to ω^{ω^ω}. The statement "for every K there is N such that for any sequence $B_1, B_2, \dots, B_N$ of positive braids such that $|B_i| < K + i$, there are $i < j \leq N$ such that $B_i \prec B_j$" is not provable in $I\Sigma_2$. (Here, $|B|$ is the smallest number of letters in a braid word representing B.) It is also possible to do a Friedman-style argument to turn a statement about a sequence of braids into a statement about comparing internal segments of one long braid.

At the dawn of modern logic, early logicians, most notably K. Gödel, wrote about ordinal descent through ω^{ω^ω} with justification why this is an acceptable mathematical principle. This ordinal was perceived by some people at that time as the biggest ordinal that allows for a convincing verbal "justification" of why the corresponding ordinal descent principle is true. This ordinal nowadays turns up in various natural mathematical contexts, see for example Higman's lemma and S. Simpson's article [82] on the Robson Basis Theorem.

One more family of independence results comes from braid-theoretic analogues of the hydra battle. Termination of the following game on positive braids is unprovable in $I\Sigma_2$. (These results have been proved jointly by Lorenzo Carlucci and the author in [13].) In the rest of this section, we write i instead of σ_i. For numbers a, b, define a wave between a and b as the braid word $w_{a,b} = a(a+1)^2 \dots (b-1)^2 b^2 (b-1)^2 \dots (a+1)^2 a$ if $a < b$ and $w_{a,b} = a(a-1)^2 \dots (b+1)^2 b^2 (b+1)^2 \dots (a-1)^2 a$ if $a > b$. A braid word is even if all blocks of its consecutive equal letters are of even length and for neighbouring blocks consist of numbers i and j that differ by 1. (We are temporarily restricting ourselves to even braids in order not to worry about applications of braid relations.) Given a positive even braid word w, we define a reduct $w[k]$ for every $k > 0$ as follows. If w ends in 11 then $w[k]$ is obtained from w by deleting this 11. Otherwise let $b_i^{m_i}$ be the first block of equal letters (counting from right to left) of length $m_i > 2$. Let $\hat{b_i}$ be the closest occurrence of the letter b_i to the right of $b_i^{m_i}$ if it exists and the empty word if there is no such occurrence. Let us then write w as $\dots\ b_i^{m_i}\ a\, u\, \hat{b_i}\ \dots$. If $a < b_i$ then $w[k]$ is defined as

$$\dots\ b_i^{m_i-2}\, (b_i-1)^3\, w^k_{b_i-1,\min u}\, (b_i-1)\, u\, \hat{b_i}\ \dots,$$

otherwise $w[k]$ is $\dots\ b_i^{m_i-1}\, (b_i+1)^3\, w^k_{b_i+1,\max u}\, (b_i+1)\, u\, \hat{b_i} \dots$. The sequence $w[1][2]\dots[k]$ consists of positive even braid words and decreases with respect to the braid ordering $\prec$. Eventually it terminates but its termination is unprovable in $I\Sigma_2$. There are versions of this sequence for 3-strand braids [13]: a simple version, whose termination time is Ackermannian and another one, formulated as a game, whose termination is unprovable in $I\Sigma_2$.

There is some hope to translate statements about braids into geometrical, topological statements with some amount of 'physical' meaning, e.g. using the fact that braid groups are fundamental groups of certain configuration spaces. This is a very rich topic with many new results (all so far based on left-orderability of certain groups) being obtained these days by P. Dehornoy, A. Weiermann, L. Carlucci and the author (for a survey, see [17]).

Long term rewriting

Another interesting twist in the story is the rewrite systems whose termination is PA-unprovable. Good references are short articles by L. Beklemishev [6] and W. Buchholz [20]. An early rewrite system that imitated the hydra was proposed by Dershowitz. Recently, G. Moser [67] provided a full analysis of this system and related systems. Alternative systems have been formulated by W. Buchholz and by H. Touzet [87]. There are also relevant results by I. Lepper. Here is an example by L. Beklemishev [5] presented as a battle between a gardener and a worm. (It is not exactly a rewrite system but there are versions of it formulated as a rewrite system in [6].) A worm is a finite sequence of natural numbers $f\colon \{0,1,\ldots,m\} \to \mathbb{N}$. For every worm $w = f(0), f(1), \ldots, f(m)$, define $w[n]$ as follows. If $f(m) = 0$ then chop it off, i.e., put $w[n] = f(0), f(1), \ldots, f(m-1)$. Otherwise find the maximal place $k < m$ such that $f(k) < f(n)$, define two words $r = f(0), \ldots, f(k)$ and $s = f(k+1), \ldots, f(m-1), f(m) - 1$ and set $w[n]$ to be r followed by $(n+1)$ copies of s. The statement $\forall w \exists n \; w[1][2]\ldots[n] = \varnothing$ ("every worm dies") is unprovable in Peano Arithmetic.

Other examples

Many other examples of unprovable statements and a discussion of the subject in its late 1970s - early 1980s state can be found in the four volumes [80], [81], [68] and [7], devoted entirely or partially to arithmetical unprovability results. In addition, I would like to mention unprovable statements obtained by translating consistency statements for various theories into non-solvability assertions about corresponding Diophantine equations (see [63]), unprovable statements about Diophantine games [63], the Schütte-Simpson treatment of finite sequences of natural numbers with gap-condition [78], K. McAloon's [64], [65] and Z. Ratajczyk's [75],[76] theories of iterations of the Paris-Harrington Principle and R. Sommer's related result on transfinite induction [86], independence results on pointwise induction by T. Arai [1], J. Avigad's statement with update procedures [4], P. Clote's PhD thesis and his anti-basis approach to independence results [25], Cichon's treatment of Goodstein sequences [24], Bigorajska-Kotlarski series of results on partitioning α-large sets (see e.g. [8], [9]), the Friedman-McAloon-Simpson early fundamental article [30] and the subsequent paper by S. Shelah [84], J.-P. Ressayre's statements unprovable in $I\Sigma_1$, PA and ZF that state existence of finite approximations of models of these theories [77], the treatment of infinitary strong statements by means of J. Paris-style density assertions in [14] and [15] by the author and A. Weiermann, the Kochen-Kripke proof with ultrapowers [57], the Buchholz Hydra [19], L. Carlucci's recent uniform treatment of hydras, worms and sequences with gap condition [23] and H. Friedman's many results on combinatorial statements unprovable in ZFC+ large cardinals [31], [34]. (Many new unprovable statements are announced regularly on H. Friedman's webpage, so we shall not even attempt to give an up-to-date account of these rapid developments here.)

The subject of first-order arithmetical unprovability has a big number of results, some of them masterpieces, produced by different authors, and is rapidly developing these days. Although it is one of the most central subjects in logic and is already seriously connected with many mathematical disciplines, there are currently no comprehensive monographs, no textbooks, not even surveys. This current sketchy and incomplete introduction is a mere glimpse at the subject and is intended to be a modest patch in this state of affairs, without any serious discussion of the subject's goals, its big fundamental questions, its history, its implications for philosophy of mathematics or its visions of the future and without any attempt to analyse and systematise its methods and ideas.

Methods of proving unprovability

Unprovability proofs so far usually fall into one of two broad categories. The first category consists of model-theoretic constructions that demonstrate how, assuming a statement, a model of our strong theory can be built directly, 'by hands'. Proofs of the first category work very well for Ramsey-style statements and for strong theories. The second category consists of combinatorial proofs springing from proof theory and the Ketonen-Solovay article [50] (which shows combinatorially that the function arising from PH eventually dominates every function of the Grzegorczyk-Schwichtenberg-Wainer hierarchy (since all

PA-provably recursive functions occur in this hierarchy, the result follows)) or from the study of well-quasi-ordered sets [79]. Most of the proofs in the volume [80] are of this category, as well as the articles [61] and [62] by M. Loebl and J. Nešetřil. Proofs in this category work very well (often better than model-theoretic proofs) at the lower end of consistency strength and for well-quasi-orders, like trees, whose ordinal treatment is well-developed but the model-theoretic treatment is currently absent.

Apart from the original articles we mentioned above, other good sources reporting on proofs of the first category would be the book [43] by P. Hájek and P. Pudlák and the papers [2], [3] by J. Avigad and R. Sommer (proof-theoretic aspects). A recent book manuscript [58] by H. Kotlarski is an exposition of both approaches as well as of many different proofs of Gödel's Theorem. The second category is well explained in an early article by W. Buchholz and S. Wainer [18], in the Friedman-Sheard article [33], in the Fairtlough-Wainer paper [29] and in the recent article by A. Weiermann [91].

Apparently, there is also a third category of unprovability proofs, proofs that interpret directly independent statements as reduction strategies for proof systems. We have little to say about it and refer the reader to a recent article by L. Carlucci [21], an old paper by H. Jervell [46] and the articles [44] and [45] by M. Hamano and M. Okada.

Once we have one unprovable statement, it is usually possible to find a series of mathematical statements in very different contexts that imply our statement, sometimes in a very non-trivial way. It often happens that non-logical mathematical obstacles in proving such implications are difficult and deep. In this way, the original Ramsey-theoretic and well-quasi-order-theoretic reasons for unprovability may end up very well hidden, and new mathematical statements appear natural and still carry a big amount of unprovability and logical strength.

Some people somehow view research in unprovability in terms of *big numbers and fast-growing functions.* Indeed, big numbers and fast-growing functions often occur in our discourse. (For example, the Paris-Harrington Principle implies every Π_2 theorem of Peano Arithmetic, including totality of every PA-provably recursive function.) However, this simple perception of the subject eventually misleads when one reaches the higher strata of consistency strength where the philosophical status of unprovable statements is very different. It is better, at least in the context of this article (and especially in the context of statements unprovable in higher theories) to view unprovable statements as possessing a large "amount of logical strength" that allows us to build approximations of models of strong theories.

2 Unprovability of PH

Nowadays unprovability of the Paris-Harrington Principle in Peano Arithmetic is presented via unprovability proof of KM followed by a proof that PH implies KM or by the combinatorial Ketonen-Solovay method using α-large sets or by a sequence of lemmas (gluing different colourings) as in the original article [69]. A model-theoretic proof was promised in the original article [69] but never appeared. This was a folklore understanding in the 1970s and 1980s that a model-theoretic proof can be written down. The closest it got is in the treatment of n-extendible initial segments by J. Paris on pages 324-328 in [71] and in the book by P. Hájek and P. Pudlák [43] but still far from being instantly accessible to a wide audience. We shall give a single simple proof (inspired by the Kanamori-McAloon proof [47]) that will be useful in logic courses. There will be one colouring which will yield a desired set of indiscernibles among its large homogeneous set. To our knowledge, the proof below is the first time a simple direct model-theoretic unprovability proof for PH appears in print.

Finally, let me mention that all statements below, unprovable in Peano Arithmetic, are easy consequences of the Infinite Ramsey Theorem, so will usually be labeled as "true".

Theorem 1. (Paris-Harrington)
The statement "for all a, k, m, c, there is $b > a$ such that for every $f\colon [a,b]^k \to c$, there is an f-homogeneous $H \subset N$ of size m such that $|H| > \min H$" is not provable in Peano Arithmetic.

The readers who need a reminder of basic notions from logic (e.g., models of arithmetic, overspill and the usual way in which sets, functions and formulas are encoded and treated within a model of arithmetic) are referred to Kaye's textbook [48].

Proof. Let $M \models I\Sigma_1$ be nonstandard, $a, e \in M$ be nonstandard and $\varphi_1(z, x_1, x_2, \ldots, x_e)$, $\ldots, \varphi_e(z, x_1, x_2, \ldots, x_e)$ be the first e Δ_0-formulas in at most the free variables shown. In particular, this list contains all Δ_0-formulas of standard size. Denote $R(2e+1, e+1, 5e+1)$ by $r(e)$. Let $b \in M$ be minimal such that for every $g\colon [a,b]^{4e+1} \to 3e+1$, there is a g-homogeneous $H \subset [a,b]$ of size $r(e)$ such that $|H| > \min H$. We shall build a model of PA between a and b.

First define a function $f\colon [a,b]^{2e+1} \to e+2$ as follows[1]: for $c < \overline{d_1} < \overline{d_2}$ in $[a,b]$, put

$$f(c, \overline{d_1}, \overline{d_2}) = \min i \le e \ \exists p < c \ \left(\varphi_i(p, \overline{d_1}) \not\leftrightarrow \varphi_i(p, \overline{d_2})\right)$$

if such i exists and $e+1$ otherwise. (In order to write down this formula we should use a formula that represents the satisfaction relation for Δ_0-formulas.) Informally, $f(c, \overline{d_1}, \overline{d_2})$ is the first formula (with a parameter smaller than c allowed) that distinguishes the tuples $\overline{d_1}$ and $\overline{d_2}$.

Define another function h, defined on $[a,b]^{2e+1}$, as follows:

$$h(c, \overline{d_1}, \overline{d_2}) = \min p < c \ \left(\varphi_{f(c,\overline{d_1},\overline{d_2})}(p, \overline{d_1}) \ \not\leftrightarrow \ \varphi_{f(c,\overline{d_1},\overline{d_2})}(p, \overline{d_2})\right)$$

if $f(c, \overline{d_1}, \overline{d_2}) \ne e+1$ and c otherwise. I.e., the value $h(c, \overline{d_1}, \overline{d_2})$ is the first parameter p with which $\varphi_{f(c,\overline{d_1},\overline{d_2})}$ distinguishes $\overline{d_1}$ and $\overline{d_2}$.

Now let us introduce our colouring $g\colon [a,b]^{4e+1} \to 3e+1$ as follows:

$$g(c, \overline{d_1}, \overline{d_2}, \overline{d_3}, \overline{d_4}) = \begin{cases} 0 \quad \text{if} & h(c, \overline{d_1}, \overline{d_2}) = h(c, \overline{d_3}, \overline{d_4}) \\ \\ j \quad \text{if} & h(c, \overline{d_1}, \overline{d_2}) \ne h(c, \overline{d_3}, \overline{d_4}) \text{ and} \\ & 0 < j \le 3e \text{ and } h(c, \overline{d_1}, \overline{d_2}) \equiv j (\operatorname{mod} 3e) \end{cases}$$

Choose a g-homogeneous set $H \subseteq [a,b]$ of size at least $r(e)$ and such that $|H| > \min H$. Let us show that the value of g on $[H]^{4e+1}$ is 0. Suppose that the value of g on $[H]^{4e+1}$ is $j \ne 0$. Notice that then there are not more than $\left[\frac{\min H}{3e}\right] < \frac{|H|}{2e}$ points below $\min H$ that can be values of $h(\min H, \overline{d_1}, \overline{d_2})$ for $\overline{d_1} < \overline{d_2}$ in H. Hence, by the pigeonhole principle, there are $\overline{d_1} < \overline{d_2} < \overline{d_3} < \overline{d_4}$ such that $h(\min H, \overline{d_1}, \overline{d_2}) = h(\min H, \overline{d_3}, \overline{d_4})$, which is a contradiction. Hence the value of g on $[H]^{4e+1}$ is 0.

Since $|H| \ge r(e)$, there is a set $C \subseteq H$ of $5e+1$ points $c_0 < \cdots < c_{5e}$ in H such that for any two $(2e+1)$-tuples $c < \overline{d_1} < \overline{d_2}$ and $c' < \overline{d_1'} < \overline{d_2'}$ in C, we have $f(c, \overline{d_1}, \overline{d_2}) = f(c', \overline{d_1'}, \overline{d_2'})$. If the value of f on $[C]^{2e+1}$ is $i \ne e+1$ then define $p = h(c_0, c_{3e+1}, \ldots, c_{5e})$. Then for all $\overline{d_1} < \overline{d_2}$ among $\{c_1, \ldots, c_{3e}\}$, we have $h(c_0, \overline{d_1}, \overline{d_2}) = p$ because $g \equiv 0$ on $[H]^{4e+1}$. Hence

$$\varphi_i(p, c_1, \ldots, c_e) \not\leftrightarrow \varphi_i(p, c_{e+1}, \ldots, c_{2e}) \not\leftrightarrow \varphi_i(p, c_{2e+1}, \ldots, c_{3e}) \not\leftrightarrow \varphi_i(p, c_1, \ldots, c_e),$$

[1] we shall often shorten $c < a_1 < \cdots < a_e < b_1 < \cdots < b_e$ as $c < \overline{a} < \overline{b}$

which is impossible since we have only two truth values. Hence $i = e + 1$ and we can conclude that the set $C' = \{c_0 < c_1 < \cdots < c_{2e}\}$ is our desired set of indiscernible elements (=indiscernibles): for any standard Δ_0-formula $\varphi(z, x_1, \ldots, x_n)$ any $c \in C'$, any $d_1 < \cdots < d_n$ and $e_1 < \cdots < e_n$ above c and any $p < c$, we have $\varphi(p, d_1, \ldots, d_n) \leftrightarrow \varphi(p, e_1, \ldots, e_n)$. Now we shall repeat the usual argument that the initial segment $I = \sup_{k \in \mathbb{N}} c_k$ is a model of Peano Arithmetic. Once we prove that, observe that $I \vDash \mathrm{PA} + \neg\mathrm{PH}$ since $b > I$ was the minimal point for which a large homogeneous set exists for every colouring $g\colon [a, b]^{4e+1} \to 3e + 1$.

To prove that I satisfies PA, we have to show that I is closed under addition and multiplication and satisfies the induction scheme. All other axioms are inherited from M. To show that I is closed under addition, it suffices to show that $c_1 + c_2 < c_3$. If it was the case that $c_1 + c_2 \geq c_3$ then for some $p < c_1$, $p + c_2 = c_3$. Then, by indiscernibility (using the formula $z + x_2 = x_3$ as φ), for any $c > c_2$ in H, we have $p + c_2 = c$, which is a contradiction. Hence $c_1 + c_2 < c_3$ and I is closed under addition. The same argument for multiplication: if $c_1 \cdot c_2 > c_3$ then there is $p < c_1$ such that $p \cdot c_2 < c_3 < p \cdot c_2 + c_2$. By indiscernibility this would now mean that $c_4 < p \cdot c_2 + c_2 < c_3 + c_2$, which is impossible since we already proved that $c_4 > c_2 + c_3$. Hence I is a structure in the language of arithmetic. Let us now show that the scheme of induction holds in I. Let $p \in I$ and $\varphi(p, x_1, x_2, \ldots, x_n, x)$ be such that $I \vDash \exists x\ \forall x_1 \exists x_2 \forall x_3 \ldots\ \varphi(p, x_1, x_2, \ldots, x_n, x)$, i.e., we take an arbitrary formula with a parameter $p \in I$ and one free variable x and assume that there is $x \in I$ which satisfies it. We want to find a minimal such x in I. For some $k \in \mathbb{N}$ and some $\ell \in I$, we have $\langle p, \ell \rangle < c_k$ and $I \vDash \forall x_1 \exists x_2 \forall x_3 \ldots \varphi(p, x_1, \ldots, x_n, \ell)$. By indiscernibility, for every $d \leq \ell$, $I \vDash \forall x_1 \exists x_2 \forall x_3 \ldots \varphi(p, x_1, \ldots, x_n, d)$ if and only if $M \vDash \forall x_1 < c_{k+1} \exists x_2 < c_{k+2} \forall x_3 < c_{k+3} \ldots \varphi(p, x_1, \ldots x_n, d)$. However, in M there is a minimal d such that $M \vDash \forall x_1 < c_{k+1} \exists x_2 < c_{k+2} \forall x_3 < c_{k+3} \ldots \varphi(p, x_1, \ldots x_n, d)$, hence this d is minimal such that $I \vDash \forall x_1 \exists x_2 \forall x_3 \ldots \varphi(p, x_1, \ldots, x_n, d)$. We have shown that $I \vDash \mathrm{PA} + \neg\mathrm{PH}$ and our proof is complete. □

In this proof, we actually needed much less from our nonstandard ground model than that it satisfies $I\Sigma_1$, namely Δ_0-induction and existence of Ramsey numbers (i.e., closedness under the tower-function) but the assumption of existence of Ramsey numbers can be eliminated (by the overspill argument below), so Δ_0-induction plus totality of exponentiation is enough.

Also notice that the model of PA + ¬PH was constructed from just one nonstandard instance of PH (the existence of $b \in M$ such that for any colouring $[a, b]^{4e+1} \to 3e + 1$ there exists a large homogeneous subset of size $r(e)$). Why would such an instance exist? We could of course start with a ground model $M \vDash I\Sigma_1 + \mathrm{PH}$ and not worry any more (this should be the pedagogically preferred option). However notice that for every standard $a, n \in M$ there is b_n such that for any colouring $[a, b_n]^{4n+1} \to 3n + 1$, there exists a homogeneous H of size $r(n)$ such that $|H| > \min H$ (because PH holds in the "standard model" part of M). Now do a Δ_0-overspill argument to find the nonstandard numbers a, e and b as needed in the beginning of the proof. (Notice also that this overspill argument immediately re-proves the famous fact that every nonstandard model of $I\Sigma_1$ has many initial segments satisfying PA. Probably this is a well-known observation.)

A good exercise for the reader at this stage would be to modify the proof of Theorem 1 to show that the number of colours in PH can be fixed as 2 and the statement stays unprovable. A stronger result (unprovability of the Paris-Harrington principle in fixed dimension and two colours) belongs to J. Paris and is Lemma 29 in [71]. Another exercise for the reader would be to modify the above proof and get unprovability with fixed $a = 0$.

3 Adapted version of PH

The unprovability proof for PH is now easy (as easy as for KM) but we can formulate unprovable statements with yet simpler unprovability proofs if we sacrifice one exponent. I suggest that logic courses can use the proofs below instead of proofs for PH or KM.

First, we present an adapted version of PH, whose unprovability is established easily by one straightforward step, and the proof is rid of all unnecessary combinatorics.

Theorem 2. The statement "for all m, n and c, there exists N such that for every $f\colon [N]^n \to c$, there is an f-homogeneous $H \subset N$, of size at least m and such that $|H| > n \cdot (2^{n \cdot \min H} + 1)$" is not a theorem of Peano Arithmetic.

Proof.
We shall prove unprovability already with $c = 2$. Let $M \vDash I\Sigma_1$ be nonstandard, $e \in M \smallsetminus \mathbb{N}$, and $\varphi_1(x_1, \ldots, x_e, y), \ldots, \varphi_e(x_1, \ldots, x_e, y)$ be the enumeration of the first e Δ_0-formulas in at most the free variables shown. Suppose $N \in M$ is the minimal point such that for every $f\colon [N]^{2e+1} \to 2$, there is an f-monochromatic $H \subseteq N$ such that $|H| > e \cdot (2^{e \cdot \min H} + 1)$. Define our colouring $f\colon [N]^{2e+1} \to \{0, 1\}$ as follows: if $\{a < b_1 < b_2 < \cdots < b_e < c_1 < \cdots < c_e\}$ is a $(2e+1)$-subset of N, put $f(a, b_1, \ldots, b_e, c_1, \ldots, c_e) = 0$ if for all $x < a$,

$$\{i \leq e \mid \varphi_i(b_1, \ldots, b_e, x)\} = \{i \leq e \mid \varphi_i(c_1, \ldots, c_e, x)\}$$

and 1 otherwise. Using the definition of N, extract an f-homogeneous set $H \subset N$ of size greater than $e \cdot (2^{e \cdot \min H} + 1)$.

For every e-tuple $b_1 < b_2 < \cdots < b_e$ in $H \smallsetminus \{\min H\}$, define the following sequence of e-many subsets of $[0, \min H)$:

$$\langle \{x < \min H \mid \varphi_1(b_1, \ldots, b_e, x)\}, \ldots, \{x < \min H \mid \varphi_e(b_1, \ldots, b_e, x)\} \rangle.$$

There can be no more than $2^{e \cdot \min H}$ such sequences, hence, by the pigeonhole principle, there are $b_1 < b_2 < \cdots < b_e < c_1 < \cdots < c_e$ in $H \smallsetminus \{\min H\}$ such that

$$f(\min H, b_1, \ldots, b_e, c_1, \ldots, c_e) = 0.$$

Hence, by homogeneity, f is constant 0 on $[H]^{2e+1}$.

Let $d_1 < \cdots < d_e$ be the last e elements of H. Then for any $a < b_1 < \cdots < b_e$ and $a < c_1 < \cdots < c_e$ in $H \smallsetminus \{d_1, \ldots, d_e\}$, we have: for every $x < a$,

$$\{i \leq e \mid \varphi_i(b_1, \ldots, b_e, x)\} = \{i \leq e \mid \varphi_i(d_1, \ldots, d_e, x)\} = \{i \leq e \mid \varphi_i(c_1, \ldots, c_e, x)\},$$

which is the indiscernibility condition we need. Again define an initial segment I as the supremum of the first $\mathbb{N}$ points of H and notice that $I \vDash \mathrm{PA}$ and that I satisfies the negation of our statement. □

4 Adapted version of KM

For KM, there is also a short adapted version which we first presented in [10].

Theorem 3. The statement "for all m, a and n with $n < m$, there is $b > a$ such that for every $f\colon [a, b]^n \to 2^{nb}$ such that $f(x_1, \ldots, x_n) < 2^{n \cdot x_1}$, there is an f-min-homogeneous subset of $[a, b]$ of size m" is not provable in Peano Arithmetic.

Proof.
Let $M \vDash I\Sigma_1$ be a nonstandard ground model, $e, a, b \in M \smallsetminus \mathbb{N}$, $e < a < b$ and $b \in M$ be minimal such that for every function f defined on $[a,b]^{e+1}$ such that $f(x_0, x_1, \ldots, x_e) < 2^{e \cdot x_0}$, there is a min-homogeneous subset of $[a,b]$ of size $2e$.

Let $\varphi_1(x_0, x_1, \ldots, x_e), \ldots, \varphi_e(x_0, x_1, \ldots, x_e)$ be the first e Δ_0-formulas in not more than the $e+1$ free variables shown. Define f as follows: for every $x_0 < x_1 < \cdots < x_e$ in $[a,b]$, put

$$f(x_0, x_1, \ldots, x_e) = \left\langle \begin{matrix} \{p < x_0 \mid \varphi_1(p, x_1, \ldots, x_e)\} \\ \vdots \\ \{p < x_0 \mid \varphi_e(p, x_1, \ldots, x_e)\} \end{matrix} \right\rangle .$$

i.e., $f(x_0, x_1, \ldots, x_e)$ is a code of a collection of e subsets of $[0, x_0 - 1)$, hence $f(x_0, x_1, \ldots, x_e) < 2^{e \cdot x_0}$. Extract an f-min-homogeneous subset $H = \{c_i\}_{i=1}^{2e} \subseteq [a,b]$ and notice that for every Δ_0-formula $\varphi(x_0, x_1, \ldots, x_n)$, any $i_0 < i_1 < \cdots < i_n$ and $i_0 < j_1 < \cdots < j_n$ and all $p < c_{i_0}$, we have $M \vDash \varphi(p, c_{i_1}, \ldots, c_{i_n}) \leftrightarrow \varphi(p, c_{j_1}, \ldots, c_{j_n})$, which is exactly our indiscernibility condition. Again define an initial segment I as the supremum of the first $\mathbb{N}$ points of H and notice that $I \vDash \mathrm{PA}$ and that I satisfies the negation of our statement. □

5 Model-theoretic proof of a threshold result for PH

The following theorem was first proved by A. Weiermann in [89] using ordinals.

Theorem 4. For every $n \in \omega$, Peano Arithemtic does not prove $\mathrm{PH}_{\log^{(n)}}$.

The combinatorial treatment of KM was done by G. Lee in [59], also using ordinals.

Theorem 5. For every $n \in \omega$, Peano Arithmetic does not prove $\mathrm{KM}_{\log^{(n)}}$.

We shall give one model-theoretic proof of both of these theorems. It is also possible to show by a combinatorial argument how each of these statements implies PH, by a usual trade-off where increased dimension allows for weaker largeness restrictions on a homogeneous set. The reader can think of these possible combinatorial arguments as being already incorporated into the complete model-theoretic proof below.

For every number a, let $\mathbf{X}(a)$ be the set coded by a. We fix a coding method such that every subset of $[0,a]$ has a code before $2^{a+1}+1$. For example, let every $X \subseteq [0,a]$ be coded by $\sum_{i=0}^{a} \chi(i) \cdot 2^i$, where χ is the characteristic function of X. Let us first write down the proof for $n = 1$.

Proof.
As usual, take a nonstandard ground model $M \vDash I\Sigma_1$ and elements $a > e > \mathbb{N}$. Let $\varphi_1(z, x_1, x_2, \ldots, x_e), \ldots, \varphi_e(z, x_1, x_2, \ldots, x_e)$ be the first e Δ_0-formulas in at most the free variables shown. In particular, this list will contain all Δ_0-formulas of standard size. Denote $R(2e+1, e+2, 10e+1)$ by $r(e)$. Let $b \in M$ be minimal such that for every $g\colon [a,b]^{8e+1} \to 5e+1$ there is $H \subset [a,b]$ of size at least $r(e)$ such that $|H| > \log(\min H)$. Define $f\colon [a,b]^{2e+1} \to e+2$, the minimal formula of disagreement, as before: for $c < \overline{d_1} < \overline{d_2}$ in $[a,b]$, put $f(c, \overline{d_1}, \overline{d_2}) = \min i \leq e \; \exists p < c \; (\varphi_i(p, \overline{d_1}) \not\leftrightarrow \varphi_i(p, \overline{d_2}))$ if such i exists and $e+1$ otherwise. Define a regressive function h, the minimal parameter of disagreement, as before: $h(c, \overline{d_1}, \overline{d_2}) = \min p < c \; (\varphi_{f(c,\overline{d_1},\overline{d_2})}(p, \overline{d_1}) \not\leftrightarrow \varphi_{f(c,\overline{d_1},\overline{d_2})}(p, \overline{d_2}))$ if $f(c, \overline{d_1}, \overline{d_2}) \neq e+1$ and c otherwise. Now, define a log-regressive function ℓ defined on $[a,b]^{4e+1}$. Let $c < \overline{d_1} < \overline{d_2} < \overline{d_3} < \overline{d_4}$ be a $(4e+1)$-subset in $[a,b]$. Let $p_1 = h(c, \overline{d_1}, \overline{d_2})$, $p_2 = h(c, \overline{d_3}, \overline{d_4})$. Put

$$\ell(c, \overline{d_1}, \overline{d_2}, \overline{d_3}, \overline{d_4}) = \min(\mathbf{X}(p_1) \Delta \mathbf{X}(p_2)) \text{ if } \; p_1 \neq p_2$$

and $\log(c)$ otherwise. Here Δ denotes the symmetric difference relation: $A\Delta B = (A \smallsetminus B) \cup (B \smallsetminus A)$. Clearly, $\ell(c, \overline{d_1}, \overline{d_2}, \overline{d_3}, \overline{d_4}) \leq \log(c)$.

Now, define the main colouring $g\colon [a,b]^{8e+1} \to 5e+1$ as follows:

$$g(c, \overline{d_1}, \dots, \overline{d_8}) = \begin{cases} 0 & \text{if} \qquad \ell(c, \overline{d_1}, \dots, \overline{d_4}) = \ell(c, \overline{d_5}, \dots, \overline{d_8}) \\ j & \text{if} \qquad \ell(c, \overline{d_1}, \dots, \overline{d_4}) \neq \ell(c, \overline{d_5}, \dots, \overline{d_8}) \text{ and} \\ & \quad 0 < j \leq 5e \text{ and } \ell(c, \overline{d_1}, \dots, \overline{d_4}) \equiv j (\text{mod } 5e) \end{cases}$$

Extract a g-homogeneous $H \subset [a,b]$ of size $r(e)$ such that $|H| > \log(\min H)$. Observe, as usual, that g is constant 0 on $[H]^{8e+1}$. Indeed, if $g|_{[H]^{8e+1}}$ is $j \neq 0$ then below $\log(\min H)$ there are at most $\left[\frac{\log(\min H)}{5e}\right]$ possible values of $\ell(\min H, \overline{d_1}, \overline{d_2}, \overline{d_3}, \overline{d_4})$, hence for $\frac{|H|}{4e}$ successive $4e$-subsets in $H \smallsetminus \{\min H\}$, there are not enough separate spaces, hence for some $\overline{d_1} < \dots < \overline{d_8}$, $\ell(\min H, \overline{d_1}, \overline{d_2}, \overline{d_3}, \overline{d_4}) = \ell(\min H, \overline{d_5}, \overline{d_6}, \overline{d_7}, \overline{d_8})$. Hence $g|_{[H]^{8e+1}} = 0$.

Choose $H_1 \subset H$ of size $10e+1$, homogeneous for f. This can be done by our definition of $r(e)$. Let H_2 be $H_1 \smallsetminus \{\text{the last } 4e \text{ elements of } H_1\}$.

Let us show that $f|_{[H_2]^{2e+1}} = e+1$. Suppose that $f|_{[H_2]^{2e+1}}$ has constant value $i \neq e+1$ and we shall get a contradiction. List H_2 as $c < \overline{d_1} < \dots < \overline{d_6}$. Denote the value of $\ell(c, \dots)$ on H_2 by p and let $h(c, \overline{d_1}, \overline{d_2}) = p_1$, $h(c, \overline{d_3}, \overline{d_4}) = p_2$, $h(c, \overline{d_5}, \overline{d_6}) = p_3$. Let us first notice that p_1, p_2 and p_3 are all different. If two of them coincided, say $p_1 = p_2$ then $\ell(c, \overline{d_1}, \overline{d_2}, \overline{d_3}, \overline{d_4}) = p = \log c$ and hence $p_1 = p_3$. So, let us assume $p_1 = p_2 = p_3 = q$. Since $i \neq e+1$, we know that $q < c$. Suppose without loss of generality that $\varphi_i(q, \overline{d_1})$, $\varphi_i(q, \overline{d_3})$ and $\varphi_i(q, \overline{d_5})$ are true (and $\varphi_i(q, \overline{d_2})$, $\varphi_i(q, \overline{d_4})$ and $\varphi_i(q, \overline{d_6})$ are false). Since $\ell(c, \overline{d_1}, \overline{d_2}, \overline{d_3}, \overline{d_5}) = \ell(c, \overline{d_1}, \overline{d_2}, \overline{d_3}, \overline{d_4})$, we conclude that $\overline{d_3}$ and $\overline{d_5}$ disagree on q, which contradicts our knowledge that $\varphi_i(q, \overline{d_3})$ and $\varphi_i(q, \overline{d_5})$ are both false. We arrived at a contradiction, hence $p_1 \neq p_2$, $p_1 \neq p_3$ and $p_2 \neq p_3$ and the sets $\mathbf{X}(p_1), \mathbf{X}(p_2)$ and $\mathbf{X}(p_3)$ are different.

Notice that by definition of ℓ, we have $p \in \mathbf{X}(p_1) \Leftrightarrow p \notin \mathbf{X}(p_2)$, $p \in \mathbf{X}(p_2) \Leftrightarrow p \notin \mathbf{X}(p_3)$ and $p \in \mathbf{X}(p_1) \Leftrightarrow p \notin \mathbf{X}(p_3)$, which are contradictory assertions. Hence $f|_{[H_2]^{2e+1}} = e+1$ and H_2 is our desired set of indiscernibles. □

On the way we also proved PA $\nvdash$ $\mathrm{KM}_{\log}$: indeed the set H_1 is obtained as min-homogeneous for the log-regressive function ℓ.

Now the proof for arbitrary n. The idea is the same as in case $n = 1$. We shall make n steps downstairs and focus the contradiction below $\log^{(n)}(c)$.

Proof.
As usual, take a nonstandard ground model $M \vDash I\Sigma_1$ and elements $a > e > \mathbb{N}$. Let $\varphi_1(z, x_1, x_2, \dots, x_e), \dots, \varphi_e(z, x_1, x_2, \dots, x_e)$ be the first e Δ_0-formulas in the free variables shown. In particular, this list will contain all Δ_0-formulas of standard size. Denote $R(2e+1, e+2, 5 \cdot 2^n \cdot e + 1)$ by $r(e)$. Let $b \in M$ be minimal such that for every $g\colon [a,b]^{e \cdot 2^{n+2}+1} \to 2^{n+2}e+1$ there is $H \subset [a,b]$ of size at least $r(e)$ and such that $|H| > \log^{(n)}(\min H)$.

Define $f\colon [a,b]^{2e+1} \to e+2$, the minimal formula of disagreement as before: for $c < \overline{d_1} < \overline{d_2}$ in $[a,b]$, put $f(c, \overline{d_1}, \overline{d_2}) = \min i \leq e \; \exists p < c \; \varphi_i(p, \overline{d_1}) \not\leftrightarrow \varphi_i(p, \overline{d_2})$ if such i exists and $e+1$ otherwise. Define a (regressive) function h, the minimal parameter of disagreement as before: $h(c, \overline{d_1}, \overline{d_2}) = \min p < c \; \left[\varphi_{f(c, \overline{d_1}, \overline{d_2})}(p, \overline{d_1}) \not\leftrightarrow \varphi_{f(c, \overline{d_1}, \overline{d_2})}(p, \overline{d_2})\right]$ if $f(c, \overline{d_1}, \overline{d_2}) \neq e+1$ and c otherwise.

Now, define a $\log^{(n)}$-regressive function ℓ defined on $[a,b]^{2^{n+1} \cdot e+1}$. Let $c < \overline{d_1} < \dots < \overline{d_{2^{n+1}}}$ be a $(2^{n+1} \cdot e+1)$-subset in $[a,b]$. First, define for every $i = 1, 2, \dots, n+1$ a sequence of points $\{p^i_k\}_{k=1}^{2^{i-1}}$ as follows. For all $k = 1, 2, \dots, 2^n$, put $p^{n+1}_k = h(c, \overline{d_{2k-1}}, \overline{d_{2k}})$. For every $1 \leq i < n+1$ and every $k = 1, 2, \dots, 2^{i-1}$, put $p^{i-1}_k = \min(\mathbf{X}(p^i_{2k-1}) \Delta \mathbf{X}(p^i_{2k}))$. Now, put $\ell(c, \overline{d_1}, \ldots, \overline{d_{2^{n+1}}}) = p_1^1$. Clearly, $\ell(c, \overline{d_1}, \ldots, \overline{d_{2^{n+1}}}) \leq \log^{(n)}(c)$. Now, define the main colouring $g\colon [a,b]^{2^{n+2} \cdot e+1} \to 2^{n+2}e+1$ as follows: $g(c, \overline{d_1}, \ldots, \overline{d_{2^{n+1}}}, \overline{d_{2^{n+1}+1}} \ldots, \overline{d_{2^{n+2}}}) =$

$$= \begin{cases} 0 \quad \text{if} & \ell(c, \overline{d_1}, \ldots, \overline{d_{2^{n+1}}}) = \ell(c, \overline{d_{2^{n+1}+1}}, \ldots, \overline{d_{2^{n+2}}}) \\ \\ j \quad \text{if} & \ell(c, \overline{d_1}, \ldots, \overline{d_{2^{n+1}}}) \neq \ell(c, \overline{d_{2^{n+1}+1}}, \ldots, \overline{d_{2^{n+2}}}) \text{ and} \\ & 0 < j \leq 2^{n+2}e \text{ and } \ell(c, \overline{d_1}, \ldots, \overline{d_{2^{n+1}}}) \equiv j (\operatorname{mod} 2^{n+2}e) \end{cases}$$

Choose a g-homogeneous set $H \subseteq [a,b]$ of size $r(e)$ such that $|H| > \log^{(n)}(\min H)$. Let us show that the value of g on $(2^{n+2} \cdot e+1)$-subsets of H is 0. Suppose that the value of g on $(2^{n+2} \cdot e+1)$-subsets of H is $j \neq 0$. Notice that then there are not more than $\left\lceil \frac{\log^{(n)}(\min H)}{2^{n+2}e} \right\rceil$ points below $\log^{(n)}(\min H)$ that can possibly be values of $\ell(\min H, \overline{d_1}, \ldots, \overline{d_{2^{n+1}}})$. Hence, by the pigeonhole principle (using the fact that $\frac{\log^{(n)}(\min H)}{2^{n+2}e} < \frac{|H|}{2^{n+1}e}$), there are $\overline{d_1} < \cdots < \overline{d_{2^{n+1}}} < \overline{d_{2^{n+1}+1}} < \cdots < \overline{d_{2^{n+2}}}$ such that $\ell(c, \overline{d_1}, \ldots, \overline{d_{2^{n+1}}}) = \ell(c, \overline{d_{2^{n+1}+1}}, \ldots, \overline{d_{2^{n+2}}})$, which is a contradiction. Hence the value of g on $(2^{n+2}e+1)$-subsets of H is 0.

Since $|H| \geq r(e)$, we can choose an f-homogeneous $H' \subseteq H$, of size $5 \cdot 2^n \cdot e + 1$.

Let us show that the value of f on $(2e+1)$-subsets of H' is $e+1$, thus producing our desired set of indiscernibles. Let $H'' = H' \setminus \{\text{the last } 2^{n+1} \cdot e \text{ points of } H'\}$. List the set H'' as $c < \overline{d_1} < \cdots < \overline{d_{3 \cdot 2^{n+1}}}$.

Notice that the function $\ell(c, \ldots)$ is constant on H''. Denote the value of ℓ on H'' by p. Again, show (by an argument similar to the argument above) that $p_1^n \neq p_2^n \neq p_3^n \neq p_1^n$ and notice that since $\ell(c, \overline{d_1}, \ldots, \overline{d_{2^n}}, \overline{d_{2^n+1}}, \ldots, \overline{d_{2^{n+1}}}) =$

$$= \ell(c, \overline{d_1}, \ldots, \overline{d_{2^n}}, \overline{d_{2^{n+1}+1}}, \ldots, \overline{d_{3 \cdot 2^n}}) = \ell(c, \overline{d_{2^n+1}}, \ldots, \overline{d_{2^{n+1}}}, \overline{d_{2^{n+1}+1}}, \ldots, \overline{d_{3 \cdot 2^n}}) = p,$$

we have $p \in \mathbf{X}(p_1^n) \Leftrightarrow p \notin \mathbf{X}(p_2^n) \Leftrightarrow p \in \mathbf{X}(p_3^n) \Leftrightarrow p \notin \mathbf{X}(p_1^n)$, which is a contradiction. Hence f is constant $e+1$ on H'' and H'' is our desired set of indiscernibles. □

It is clear that on the way we proved that Peano Arithmetic does not prove $\mathrm{KM}_{\log^{(n)}}$: notice that an ℓ-min-homogeneous set of size $r(e)$ produces the same set of indiscernibles as in our proof.

We conjecture that the proof above can now be converted into a proof of the full A. Weiermann's threshold result from [89]: Peano Arithmetic does not prove PH_f, where $f(n) = \log^{H_{\varepsilon_0}^{-1}(n)}(n)$, where H_{ε_0} is the ε_0th function in the Hardy hierarchy (that eventually dominates all PA-provably recursive functions).

Another approach to proving threshold results for PH would be to use Lemma 3.3. of [47] (which gives KM-thresholds) and then prove combinatorial implications between PH_f and KM_g for different functions f and g. In the case of PH-thresholds, this would require the same amount of work as we did above.

References

[1] Arai, T. (1991). A slow growing analogue to Buchholz' proof. *Ann. Pure Appl. Logic*, 54, no. 2, pp. 101-120.

[2] Avigad, J., Sommer R. (1997). A model-theoretic approach to ordinal analysis. *Bull. Symbolic Logic* 3, pp. 17-52.

[3] Avigad, J., Sommer, R. (1999). The model-theoretic ordinal analysis of theories of predicative strength. *Journal of Symbolic Logic*, 64, pp. 327-349.

[4] Avigad, J. (2002). Update procedures and 1-consistency of arithmetic. *Mathematical Logic Quarterly*, 48, pp. 3-13.

[5] Beklemishev, L. D. (2006). The worm principle. *Logic Colloquium 2002*, Lecture Notes in Logic, pp. 75-95.

[6] Beklemishev, L. D. (2007). Representing worms as a term rewriting system. *Oberwolfach Report* 52/2006, pp. 7-9.

[7] Berline, McAloon, Ressayre eds. (1981). *Model Theory and Arithmetic.* Lecture Notes in Mathematics 890, Springer-Verlag.

[8] Bigorajska, T., Kotlarski, H. (1999). A partition theorem for α-large sets. *Fund. Math.* 160, pp. 27-37.

[9] Bigorajska, T., Kotlarski, H. (2006). Partitioning α-large sets: some lower bounds. *Transactions of the AMS*, 358, pp. 4981-5001.

[10] Bovykin, A. (2005). Several proofs of PA-unprovability. *Contemporary Mathematics series of the American Mathematical Society*, 380, pp. 21-36.

[11] Bovykin, A. (2007). Unprovability of sharp versions of Friedman's sine-principle. *Proceedings of the American Mathematical Society*, 135, pp. 2967 - 2973.

[12] Bovykin A. (2007). Exact unprovability results for compound well-quasi-ordered combinatorial classes. *Submitted.* Available online at `http://logic.pdmi.ras.ru/∼andrey/research.html`.

[13] Bovykin, A., Carlucci, L. (2006). Long games on braids. Preprint. Available online at `http://logic.pdmi.ras.ru/∼andrey/research.html`.

[14] Bovykin, A., Weiermann, A. (2005). The strength of infinitary ramseyan statements can be accessed by their densities. *Submitted.*

[15] Bovykin, A., Weiermann, A. (2007). The strength of infinitary statements can be approximated by their densities. Preprint.

[16] Bovykin, A. Weiermann, A. (2007). Unprovable statements based on diophantine approximation and distribution of values of zeta-functions. Preprint. `http://logic.pdmi.ras.ru/∼andrey/research.html`.

[17] Bovykin, A., Carlucci, L., Dehornoy, P., Weiermann, A. (2007). Unprovability results involving braids. *Preprint.*

[18] Buchholz, W., Wainer, S. (1987). Provably computable functions and the fast growing hierarchy, *Contemporary Mathematics series of the AMS*, 65, pp. 179-198.

[19] Buchholz, W. (1987). An independence result for Π^1_1-CA + BI. *Annals of Pure and Applied Logic*, 33, pp. 131-155.

[20] Buchholz, W. (2007). Another rewrite system for the standard hydra battle. *Oberwolfach Report* 52/2006, pp. 13-15.

[21] Carlucci, L. (2003). A new proof-theoretic proof of the independence of Kirby-Paris' Hydra Theorem. *Theoretical Computer Science.* 1-3 (300), pp. 365-378.

[22] Carlucci, L., Lee G., Weiermann, A. (2006). Classifying the phase transition threshold for regressive Ramsey functions. Submitted.

[23] Carlucci, L. (2005). Worms, gaps and hydras. *Mathematical Logic Quarterly*, 51 (4), pp. 342–350.

[24] Cichon, E. A. (1983). A short proof of two recently discovered independence results using recursion theoretic methods. *Proc. Amer. Math. Soc.* 87, pp. 704-706.

[25] Clote, P. (1980). Weak partition relations, finite games and independence results in Peano Arithmetics. Lecture Notes in Mathematics, 834, pp. 92-107.

[26] Clote, P., McAloon, K. (1983). Two further combinatorial theorems equivalent to the 1-consistency of Peano Arithmetic. *Journal of Symbolic Logic*, 48, pp. 1090-1104.

[27] Dehornoy, P., Dynnikov, I., Rolfsen, D., Wiest, B. (2002). Why are braids orderable? Panoramas et syntheses. Société Mathématique de France.

[28] Erdös, P., Rado, R. (1952). Combinatorial theorems on classifications of subsets of a given set. *Proc. London Math. Soc.* 2 (3), pp. 417-439.

[29] Fairtlough, M., Wainer, S. (1998). Hierarchies of provably recursive functions. *Handbook of proof theory, Studies in Logic*, 137, pp. 149–207.

[30] Friedman, H., McAloon, K., Simpson, S. (1982). A finite combinatorial principle which is equivalent to the 1-consistency of predicative analysis. *Logic Symposium I (Patras 1980)*, North-Holland, pp. 197-220.

[31] Friedman, H. (1981). On the necessary use of abstract set theory. *Advances in Mathematics* 41, pp. 209-280.

[32] Friedman, H., Robertson, N., Seymour, P. (1987). The metamathematics of the graph minor theorem. *Contemporary Mathematics*, vol. 65, pp. 229-261.

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

[34] Friedman, H. (1998). Finite functions and the necessary use of large cardinals. *Annals of Mathematics* 148, pp. 803-893.

[35] Friedman, H. (2002). Internal finite tree embeddings. *Reflections on the Foundations of Mathematics: Essays in honor of Solomon Feferman, Lecture Notes in Logic*, volume 15, pp. 62-93.

[36] Friedman, H. (2001). Long finite sequences. *Journal of Combinatorial Theory*, ser. A, vol. 95, pp. 102-144.

[37] Friedman, H. (2007). Interpretations of set theory in discrete mathematics and informal thinking. Preprint. Tarski Lectures. Available online.

[38] Friedman, H. (2007). Interpreting set theory in discrete mathematics. Preprint. Tarski Lectures. Available online.

[39] Gordeev, L. N. (1989). Generalizations of the one-dimensional version of the Kruskal-Friedman theorems. *Journal of Symbolic Logic* 54, no. 1, pp. 100–121.

[40] Gordeev, L. N. (1990). Generalizations of the Kruskal-Friedman theorems. *Journal of Symbolic Logic*, 55, no. 1, pp. 157–181.

[41] Gordeev, L. N. (1994). A modified sentence unprovable in PA. *Journal of Symbolic Logic*, 59, no. 4, pp. 1154–1157.

[42] Hájek P., Paris, J. (1986). Combinatorial principles concerning approximations of functions. *Archive for Mathematical Logic*, 26, pp. 13-28.

[43] Hájek, P., Pudlák, P. (1993). Metamathematics of first-order arithmetic. Springer-Verlag.

[44] Hamano, M., Okada, M. (1997). A Relationship among Gentzen's Proof-Reduction, Kirby-Paris' Hydra Game, and Buchholz's Hydra Game, *Math. Logic Quarterly*, 3, pp. 103-120.

[45] Hamano, M., Okada, M. (1998). A direct independence proof of Buchholz's Hydra game on finite labeled trees, *Archive for Mathematical Logic*, 37, pp. 67-89.

[46] Jervell, H. (1985). Gentzen games. *Z. Math. Logik Grundlag. Math.*, 31, no. 5, pp. 431–439.

[47] Kanamori, A., McAloon, K. (1987). On Gödel incompleteness and finite combinatorics. *Annals of Pure and Applied Logic*, 33, pp. 23-41.

[48] Kaye, R. (1991). Models of Peano Arithmetic. Oxford Logic Guides.

[49] Ketonen, J. (1979). Set theory for a small universe, I. The Paris-Harrington Axiom. Unpublished manuscript.

[50] Ketonen, J., Solovay, R. (1981). Rapidly growing Ramsey Functions. *Annals of Mathematics* (ser 2) 113, pp. 267-314.

[51] Ketonen, J. (1981). Some Remarks on Finite Combinatorics. Unpublished Manuscript.

[52] Kirby, L., Paris, J.B. (1977). Initial segments of models of Peano's axioms. *Set theory and hierarchy theory, V (Proc. Third Conf., Bierutowice, 1976)*, pp. 211–226. Lecture Notes in Mathematics 619.

[53] Kirby, L. (1977). Initial segments of models of arithmetic. PhD Thesis, Manchester University.

[54] Kirby, L. (1982). Flipping properties in arithmetic. *Journal of Symbolic Logic*, 47, pp. 416-422.

[55] Kojman, M., Shelah, S. (1999). Regressive Ramsey numbers are ackermannian. *Journal of Combinatorial Theory, Ser. A*, 86(1), pp. 177-181.

[56] Kojman, M., Lee, G., Omri, E., Weiermann, A. (2006). Sharp thresholds for the phase transition between primitive recursive and ackermannian Ramsey numbers. *Submitted*.

[57] Kochen, S., Kripke, S. (1982). Nonstandard models of Peano Arithmetic. *Logic and Algorithmic Monographie de l'Enseignment Mathematique*, 30, pp. 275-295.

[58] Kotlarski, H. (2003). A model-theoretic approach to proof theory for arithmetic. Unpublished book manuscript.

[59] Lee, G. (2005). Phase transitions in axiomatic thought. PhD Thesis. University of Münster.

[60] Loebl, M. and Matoušek, J. (1987). On undecidability of the weakened Kruskal Theorem. *Contemporary Mathematics*, vol. 65. pp. 275-280.

[61] Loebl, M. and Nešetřil, J. (1991). Fast and slow growing (a combinatorial study of unprovability). *Surveys in combinatorics 1991*, London Mathematical Society Lecture Notes Ser. 166, pp. 119-160.

[62] Loebl, M. and Nešetřil, J. (1992). An unprovable Ramsey-type theorem. *Proceedings of the American Mathematical Society*, 116 (3), pp. 819-824.

[63] Matiyasevich, Yu. (1993). Hilbert's Tenth Problem. The MIT Press.

[64] McAloon, K. (1980). Progressions transfinies de théories axiomatiques, formes combinatoires du théoreme d'incomplétude et fonctions récursives a croissance rapide (in French). *Modèles de l'Arithmetique, Asterisque* 73, pp. 41-58.

[65] McAloon, K. (1985). Paris-Harrington incompleteness and progressions of theories. *Proc. Symp. Pure Math.*, 42, pp. 447-460.

[66] Mills, G. (1980). A tree analysis of unprovable combinatorial statements. *Model Theory of Algebra and Arithmetic* Lecture Notes in Mathematics 834, pp. 312-337.

[67] Moser, G. (2006). The Hydra battle revisited. Unpublished manuscript.

[68] Pacholski et al. eds. (1980). *Model Theory of Algebra and Arithmetic*, Lecture Notes in Mathematics 834. Springer-Verlag.

[69] Paris, J., Harrington, L. (1977). A mathematical incompleteness in Peano arithmetic. *Handbook for Mathematical Logic*, ed. J. Barwise. North-Holland.

[70] Paris, J. (1978). Some independence results for Peano arithmetic. *Journal of Symbolic Logic*, 43, pp. 725-731.

[71] Paris, J. (1980). Hierarchies of cuts in models of arithmetic. *Model Theory of Algebra and Arithmetic*, Lecture Notes in Mathematics 834, pp. 312-337.

[72] Paris, J., Kirby L. (1982). Accessible independence results for Peano Arithmetic. *Bulletin of the London Mathematical Society*, 14, pp. 285-293.

[73] Pudlak, P. (1979). Another combinatorial principle independent of Peano's axioms. Unpublished.

[74] Pudlak, P. (2005). Consistency and games - in search of new combinatorial principles. *Logic Colloquium 2003*, pp. 244-281.

[75] Ratajczyk, Z. (1992). Arithmetical transfinite induction and hierarchies of functions. *Fundamenta Mathematicae*, 141, pp. 1-20.

[76] Ratajczyk, Z. (1993). Subsystems of true arithmetic and hierarchies of functions, *Annals of Pure and Applied Logic*, 64, pp. 95-152.

[77] Ressayre, J.-P. (1987). Non standard universes with strong embeddings, and their finite approximations. *Contemporary Mathematics*, vol. 65, pp. 333-358.

[78] Simpson, S., Schütte, K. (1985). Ein in der reinen Zahlentheorie unbeweisbarer Satz über endliche Folgen von natürlichen Zahlen. (in German) [A theorem on finite sequences of natural numbers that is unprovable in pure number theory] *Arch. Math. Logik Grundlag.* 25, no. 1-2, pp. 75–89.

[79] Simpson, S. (1985). Nonprovability of certain combinatorial properties of finite trees. In: *Harvey Friedman's research on the foundations of mathematics*, Elsevier Science Publishers, pp. 87-117.

[80] Simpson, S. ed. (1987). Logic and Combinatorics. Contemporary Mathematics series of the AMS, volume 65.

[81] Simpson, S. ed. (1985). *Harvey Friedman's research on the foundations of mathematics*, Elsevier Science Publishers.

[82] Simpson, S. (1988). Ordinal numbers and the Hilbert Basis Theorem, *Journal of Symbolic Logic*, 53, pp. 961-974.

[83] Simpson, S. (2007). Subsystems of Second-Order Arithmetic. Second Edition. Perspectives in Logic.

[84] Shelah, S. (1984). On logical sentences in PA. *Logic Colloquium 1982*, pp. 145-160.

[85] Smoryński, C. (1977). The incompleteness theorems. *Handbook for mathematical logic*, pp. 821–865.

[86] Sommer, R. (1995). Transfinite induction within Peano arithmetic. *Ann. Pure Appl. Logic*, 76, no. 3, pp. 231–289.

[87] Touzet, H. (1998). Encoding the hydra battle as a rewrite system. *Lecture Notes in Computer Science*, vol. 1450, pp. 267-276.

[88] Weiermann, A.(2003). An application of graphical enumeration to PA. *Journal of Symbolic Logic*, 68 (1), pp. 5-16.

[89] Weiermann, A. (2004). A classification of rapidly growing Ramsey functions. *Proceedings of the American Mathematical Society*, 132, pp. 553-561.

[90] Weiermann, A. (2005). Analytic combinatorics, proof-theoretic ordinals, and phase transitions for independence results. *Annals of Pure and Applied Logic*, 136, pp. 189-218.

[91] Weiermann, A. (2006). Classifying the provably total functions of PA. *Bull. Symbolic Logic* 12, no. 2, pp. 177–190.

[92] Weiermann, A. (2007). Phase transition thresholds for some Friedman-style independence results. *Mathematical Logic Quarterly* 53, no. 1, pp. 4-18.

[93] Woods, A. R. (1981). Some problems in logic and number theory and their connections. PhD Thesis. Manchester University.

St. Petersburg Department of Steklov Mathematical Institute,
Fontanka 27, St. Petersburg, 191 023, Russia.
andrey@logic.pdmi.ras.ru

Department of Computer Science
The University of Liverpool,
Liverpool, L69 7ZF, United Kingdom