Chapter II

Fragments and Combinatorics

Introduction. In the present chapter we shall elaborate proofs of various combinatorial principles in suitable fragments of arithmetic. In general, infinite principles deal with graphs, functions etc. on infinite sets, finite principles relate similarly to finite sets. We prove both some infinite and some finite principles; furthermore, we show some infinite principles to be equivalent to certain collection principles and some finite principles to be equivalent to certain consistency statements. Sections 1 and 2 deal with strengthenings of the infinite and finite Ramsey theorem (they will be formulated at the beginning of Sect. 1), in particular with various forms and instances of Paris-Harrington principle. This principle is very famous since it has been the first example of an arithmetical statement that has a clear combinatorial meaning, is true (in N) and is unprovable in PA.

Instances of Paris-Harrington principle will form a hierarchy of formulas, n-th of them will be proved in $I\Sigma_n$ ($n \geq 1$). As said above, in this chapter we deal with concrete proofs, not with unprovability; but unprovability results immediately follow from the results of this chapter using Gödel's incompleteness theorems (elaborated in Chap. III). We shall mention this on corresponding places in this chapter: $(n+1)$-th instance is unprovable in $I\Sigma_n$.

In Sect. 3 we shall deal with ordinals in fragments, introduce the notion of α-large sets (α an ordinal) and investigate another hierarchy of combinatorial statements, related to the first one. Results of this section will be used in Chap. IV for a characterization of functions provably recursive in $I\Sigma_n$ ($n \geq 1$).

1. Ramsey's Theorems and Fragments

(a) Statement of Results

1.1 First we shall recall Ramsey's theorems in an informal formulation. If X is a set of natural numbers, then $[X]^u$ is the set of all u-element subsets of X (or, equivalently, all increasing u-element sequences of elements of X). $F : [X]^u \to a$ (where a is a natural number) means that F is a mapping whose domain is $[X]^u$ and whose range is included in $\{0, 1, \ldots, a-1\}$. It is customary to call u the *arity* of F and a the *number of colours*. $Y \subseteq X$ is *homogeneous* for F if F restricted to $[Y]^u$ is constant. The *infinite Ramsey theorem* says that for each natural u and a and each $F : [X]^u \to a$ where X is unbounded (i.e. infinite) there is an unbounded $Y \subseteq X$ which is homogeneous for F. It is customary to denote this by

$$(\forall u, a)(\omega \to (\omega)^u_a) \quad \text{or} \quad (\forall u)(\omega \to (\omega)^u_{<\omega});$$

the first ω symbolizing the unboundedness of X and the second unboundedness of the homogeneous set.

1.2 For natural numbers, x, y, u, z, q, the symbol $[x, y] \to (q)^u_z$ means that if X is the closed interval $[x, y]$ of natural numbers, then each $f : [X]^u \to z$ has a homogeneous set of cardinality q. The *finite Ramsey theorem* is

$$(\forall x, q, u, z)(\exists y)([x, y] \to (q)^u_z).$$

The symbol $[x, y] \underset{*}{\to} (q)^u_z$ means that each f as above has a homogeneous set Y of cardinality q which is *relatively large*, i.e. $\min(Y) \leq card(Y)$. *Paris-Harrington principle* is

$$(\forall x, q, u, z)(\exists y)([x, y] \underset{*}{\to} (q)^u_z).$$

Evidently, this is a strengthening of the finite Ramsey theorem; Paris-Harrington principle follows from the infinite Ramsey theorem using König's lemma. (We shall discuss the proof below.)

1.3 Clearly, Paris-Harrington principle as well as the finite Ramsey theorem is expressible by a formula of first order arithmetic; let us write $PH(u, z)$ for $(\forall x, q)(\exists y)([x, y] \underset{*}{\to} (q)^u_z)$ (recall that u is the arity and z the number of colours). Thus Paris-Harrington principle is $(\forall u, z)PH(u, z)$. Our sequence of formulas that are "harder and harder to prove" is $(\forall z)PH(\overline{n}, z)$ for $n = 1, 2, \ldots$

1.4 On the other hand, we cannot formalize the infinite Ramsey theorem in first-order arithmetic as it stands since we cannot quantify over arbitrary sets of natural numbers. But we can quantify over sets of restricted complexity, e.g. over Δ_m sets (in $I\Sigma_1$) or *low* Δ_m (in $B\Sigma_m$). Thus we may express several partition relations saying that for each Γ_1-definable and unbounded X and for each Γ_1-definable $F : [x]^u \to z$ there is a Γ_2-definable unbounded homogeneous set (where Γ_1, Γ_2 are Δ_m, *low* Δ_m or so). Denote such a formula by

$$\omega \to (\omega)^u_z(\Gamma_1, \Gamma_2).$$

Recursive analysis of Ramsey's theorem consists in establishing truth of assertions of this type. (Pioneering work was done by Jockusch). Our aim is still more ambitious: we want to establish provability of such assertions in suitable fragments of arithmetic. We are now ready to present the main results of this section.

1.5 Theorem. For $m, n \geq 1$, $B\Sigma_{m+n}$ proves

$$\omega \to (\omega)^n_{<\omega}(low\ \Delta_{m+1}, low\ \Delta_{m+1}),$$

i.e.: For each z, if X is low Δ_{m+1} and unbounded and F is a low Δ_{m+1} mapping of $[X]^n$ into z (i.e. into $\{0, \dots, z-1\}$) then F has a low Δ_{m+1} homogeneous unbounded set.

For a proof (using low basis theorem) see below. Note that the assertion is *meaningful* in $B\Sigma_{m+1}$ and is expressible as a single formula using the coding of *low* Δ_{m+1} sets in $B\Sigma_{m+1}$ (see Chap. I, Sect. 2). Due to some obvious inclusions, we have e.g. the following corollary: for $m, n \geq 1$, $B\Sigma_{m+n}$ proves

$$\omega \to (\omega)^n_{<\omega}(\Delta_m, \Delta_{m+n}).$$

This assertion is weaker but is meaningful already in $I\Sigma_1$ and is equivalent over $I\Sigma_1$ to $B\Sigma_{m+n}$.

1.6 Theorem. For $m, n \geq 1$, $I\Sigma_1$ proves the following:

$$B\Sigma_{m+n} \equiv \omega \to (\omega)^n_{<\omega}(\Delta_m, \Delta_{m+1}).$$

(Here $B\Sigma_{m+n}$ is formulated as a single formula).
By Theorem 1.5, $B\Sigma_{m+n}$ proves an infinite Ramsey type theorem on mappings of arity n and complexity *low* Δ_{m+1}. We shall see that this theory also proves a Ramsey type theorem on mappings of arity $(n+1)$ and complexity *low* Δ_{m+1}; but it guarantees only a finite relatively large homogeneous set and the number of colours must be standard.

1.7 Theorem. Let $m, n \geq 1$. (1) $I\Sigma_{m+n-1}$ proves the following: If X is LL_m and unbounded and if $F : [X]^n \to z$ is LL_m, then for each q there is a relatively large homogeneous (finite) set having at least q elements. This can be expressed by

$$(\forall z)(\forall q)(\omega \underset{*}{\to} (q)^n_z(LL_m)).$$

(2) For each k, $I\Sigma_{m+n-1}$ proves

$$(\forall q)(\omega \underset{*}{\rightarrow} (q)_k^{n+1}(LL_m)\,.$$

$I\Sigma_{m+n}$ also proves that infinite homogeneous sets of some complexity need not exist. As an example we prove the following.

1.8 Theorem. For $m \geq 1$, $I\Sigma_{m+1}$ proves

$$\neg\omega \rightarrow (\omega)_2^2(\Delta_m, \Sigma_{m+1})$$

(Thus there is a Δ_m $F : [X]^2 \rightarrow 2$, X Δ_m, unbounded, with no Σ_{m+1} homogeneous unbounded set).

Note that a stronger result will be obtained in 1.28.

Let us summarize the above results for arity 2, 2 colours and Δ_1 mappings. We have the following:

$$\begin{aligned} I\Sigma_1 &\vdash (\forall q)\omega \underset{*}{\rightarrow} (q)_2^2(\Delta_1)\,, \\ I\Sigma_2 &\vdash \neg\omega \rightarrow (\omega)_2^2(\Delta_1, \Sigma_2)\,, \\ B\Sigma_3 &\vdash \omega \rightarrow (\omega)_2^2(\Delta_1, \mathit{low}\ \Delta_3)\,. \end{aligned}$$

Now let us consider finite Ramsey type theorems.

1.9 Theorem. For $n \geq 1$,

(a) $$I\Sigma_n \vdash (\forall x, z, q)(\exists y)([x, y] \underset{*}{\rightarrow} (q)_z^{\overline{n}})$$
(i.e. $I\Sigma_n \vdash (\forall z)PH(\overline{n}, z))$,

(b) for each k,
$$I\Sigma_n \vdash (\forall x, q)(\exists y)([x, y] \underset{*}{\rightarrow} (q)^{\overline{n}+1})$$
(i.e. for each k, $I\Sigma_n \vdash PH(\overline{n+1}, \overline{k})$.

Remark. Results of Chaps. III and IV enable us to add the following:

First, all formulas in question are Π_2.

Second, in Sect. 2 of the present chapter we prove a theorem implying that $I\Sigma_1$ proves $(\forall z)PH(\overline{n+1}, z) \rightarrow Con(I\Sigma_n^{\bullet})$ (consistency), thus we may apply Gödel's second incompleteness theorem (proved in III.2.21) to deduce the unprovability of $(\forall z)PH(\overline{n+1}, z)$ in $I\Sigma_n$. Thus we have a strictly increasing hierarchy of Π_2 formulas.

1.10 Theorem. $I\Sigma_1$ proves $(\forall x, z, q, u)(\exists y)([x, y] \rightarrow (q)_z^u)$.

This completes our list of results. In what follows we shall elaborate proofs.

(b) Proofs (of 1.5, 1.7, 1.9)

1.11 Definitions ($I\Sigma_1$). Let $F : [X]^2 \to a$ be Δ_1, X Δ_1 unbounded. An increasing sequence s of elements of X is *prehomogeneous* if for each $i < j < lh(s)$ the value $F((s)_i, (s)_j)$ depends only on the first argument, i.e. for each $i < j < k < lh(s)$ we have $F((s)_i, (s)_j) = F((s)_i, (s)_k)$. If $i < lh(s) - 1$ then the *colour* of $(s)_i$ (in s) is the common value $F((s)_i(s)_j)$ for $i < j < lh(s)$. If $lh(s) > 0$, $u > \max(s)$ and $s \frown \langle u \rangle$ is prehomogeneous, then the maximal element of s has a colour in $s \frown \langle u \rangle$, namely $F(\max(s), u)$. Let $lh(s) \geq 1$; $s \frown \langle u \rangle$ is a *minimal prehomogeneous* extension of s if $s \frown \langle u \rangle$ is prehomogeneous and there is no v between $\max(s)$ and u such that $s \frown \langle v \rangle$ is prehomogeneous and $\max(s)$ has the same colour in $s \frown \langle u \rangle$ as in $s \frown \langle v \rangle$. s is *hereditarily minimal prehomogenous* (or h.m.p.h.) if s is prehomogeneous, $s = \emptyset$ or $(s)_0 = \min X$ and for each i between 1 and $lh(s) - 1$, $s \restriction (i+1)$ is a minimal prehomogeneous extension of $s \restriction i$ is the initial segment of s of length i).

1.12 Definition. ($I\Sigma_1$). A tree T is *narrowly branching* if there is a number c such that each $s \in T$ has at most c immediate successors.

1.13 Lemma. ($I\Sigma_1$) Let $F : [X]^2 \to a$ be as in 1.11 and let T be the set of all h.m.p.h. sequences. Then T is Δ_1, and is an unbounded narrowly branching tree.

Proof. Evidently, T is a Δ_1 tree. It is narrowly branching since each $s \in T$ has at most a immediate successors. To prove that T is unbounded one easily shows that for each $x \in X$ there is an $s \in T$ such that $\max(s) = x$. Indeed, let $s_0 = \min X$; then $\langle s_0, x \rangle$ is trivially prehomogeneous and $\langle s_0 \rangle$ is h.m.p.h. Assume we have a h.m.p.h. sequence s such that $s \frown \langle x \rangle$ is prehomogeneous. If $s \frown \langle x \rangle$ is not h.m.p.h., then there is a $y < x$ such that $s \frown \langle y \rangle$ is h.m.p.h. and the colour of $\max(s)$ in $s \frown \langle y \rangle$ is the same as the colour of $\max(s)$ in $s \frown \langle x \rangle$; then $s \frown \langle y \rangle \frown \langle x \rangle$ is prehomogeneous. □

1.14 Lemma (1) ($I\Sigma_1$). If T is a Δ_1 narrowly branching unbounded tree, then for each level x, the set $\{s \in T \mid lh(s) = x\}$ is bounded.

(2) ($B\Sigma_2$). If T is as above then for each level x there is an $s \in T$ such that $lh(s) = x$ and the Δ_1 set $T_s = \{t \in T \mid t \supseteq s\}$ is unbounded. Thus $\{t \mid s \frown t \in T\}$ is an unbounded Δ_1 narrowly branching tree.

Proof. (1) Associate with each $s \in T$ a sequence $H(s) = s'$ of the same length defined as follows: $\emptyset'$ is $\emptyset$ and if $s \frown \langle u \rangle \in T$ then $(s \frown \langle u \rangle)' = s' \frown \langle i \rangle$ such that $s \frown \langle u \rangle$ is the i-th immediate successor of s in T. Note that H is Σ_1 and one-one on T. Now the set of all sequences t such that $[lh(t) = x$ and each member of t is $< a]$ is bounded by some b; by $S\Sigma_1$

$$(\exists d)(\forall t < b)[(\exists s \in T)(H(s) = t) \to (\exists s < d)(s \in T \,\&\, H(s) = t)].$$

Evidently, d is the desired bound.

(2) Given x, let c be a bound for all $s \in T$ of length x. Assume that all T_s $(s \in T, lh(s) = x)$ are bounded. Then

$$(\forall s < c)(\exists q)[s \in T \,\&\, lh(s) = x \to (\forall t \in T)(t \supseteq s \to t < q)].$$

By $B\Sigma_2$ we obtain

$$(\exists q)(\forall s < c)[s \in T \,\&\, lh(s) = x \to (\forall t \in T)(t \supseteq s \to t < q)];$$

thus

$$(\exists q)(\forall t \in T)(t < q),$$

a contradiction. □

1.15 Definitions ($I\Sigma_1$). Now let $F : [X]^u \to a$, $u \geq 3$, X Δ_1 and unbounded, $F \Delta_1$. An increasing sequence s of elements of X is *prehomogeneous* if for each $i_1 < \cdots < i_u < i_{u+1} < lh(s)$ we have $F((s)_{i_1}, \ldots, (s)_{i_{u-1}}, (s)_{i_u}) = F((s)_{i_1}, \ldots, (s)_{i_{u-1}}, (s)_{i_{u+1}})$, i.e. the value does not depend on the last argument. If $lh(s) \geq u-1$ and $s \frown \langle q \rangle$ is prehomogeneous, then the *colour* of $\max(s)$ in $s \frown \langle q \rangle$ is the finite mapping associating to each increasing sequence $i_1 < i_2 < \cdots < i_{u-2} < lh(s) - 1$ the value $F((s)_{i_1}, \ldots, (s)_{i_{u-2}}, \max(s), q)$.

The definition of a minimal prehomogeneous extension is as above; s is h.m.p.h. if *either* $lh(s) < u - 1$ and s consists of the first $lh(s)$ elements of X *or* $lh(s) \geq u$, s begins by the first $(u-1)$ elements of X and for each i between $u-1$ and $lh(s)-1$, $s \restriction (i+1)$ is a minimal prehomogeneous extension of $s \restriction i$.

1.16 Lemma ($I\Sigma_1$). Let $F : [X]^u \to a$ be as in 1.15 and let T be the set of all h.m.p.h. sequences. Then T is Δ_1 and is an unbounded finitely branching tree.

Proof. Generalize the proof of 1.13 (but drop narrow branching; finite branching is evident). □

1.17 Lemma ($B\Sigma_{m+1}$). Let $u \geq 2$ and let $F : [X]^u \to a$ be low Δ_{m+1}, X low Δ_{m+1} and unbounded. Then the set of all h.m.p.h. sequences is a low Δ_{m+1} finitely branching unbounded tree.

Hint. Relativize the above.

1.18 Lemma. $B\Sigma_2 \vdash \omega \to (\omega)^1_{<\omega}(low\, \Delta_2, low\, \Delta_2)$, i.e. if X is low Δ_2 unbounded and $F : [X] \to a$ then there is an $i < a$ such that $F^{-1}(i)$ is unbounded.

Proof. Assume the contrary, i.e. $(\forall i < a)(\exists y)[(\forall u)(F(u) = i \rightarrow u < y)]$. Since F is low Δ_2, the formula in $[\ldots]$ is Δ_2 and by $B\Sigma_2$ we get $(\exists y)(\forall i < a)[(\forall u)F(u) = i \rightarrow u < y]$, i.e. $(\exists y)(\forall u)(u \in X \rightarrow u < y)$, which is contradiction. □

1.19 *Proof of Theorem* 1.5. By induction on n. For $n = 1$ see 1.18 (and the obvious relativization). Let $F : [X]^n \rightarrow a$ be as assumed. By 1.17 take the tree T of all h.m.p.h. sequences; it is low Δ_{m+1} and, by I.3.10 (5) has a low Δ_{m+2} unbounded branch. The branch defines a low Δ_{m+2} unbounded prehomogeneous set Y and F defines on $[Y]^{n-1}$ a function $G : [Y]^{n-1} \rightarrow a$; G is low Δ_{m+2}. By the induction hypothesis, G has an unbounded homogeneous set Z which is low $\Delta_{m+1+n-1}$, i.e. low Δ_{m+n}. Z is homogeneous also for F. □

1.20 *Proof of Theorem* 1.7. (1) The initial case for $n = 1$ is

$$I\Sigma_m \vdash (\forall q)(\forall z)(\omega \underset{*}{\rightarrow} (q)^1_z(LL_m));$$

due to relativization, LL_m may be replaced by Δ_m. We also take $m = 1$, i.e. we prove the following: if $X \in \Delta_1$ is unbounded and $F \in \Delta_1$ maps X into $(< z)$ then there is a relatively large set a of cardinality at least q such that F is constant on a. But this is easy: By $S\Sigma_1$ find a $b \in X$ such that for each colour $i < z$, $(\exists x \in X)(F(x) = i)$ implies $(\exists x < b)(x \in X \,\&\, F(x) = i)$. Then let X_0 be the set of first $(b+1) * z$ elements of X and let j be a colour such that $a = \{x \in X_0 \mid x \text{ has the colour } j\}$ has more than b elements. Since $\min(a) \leq b$, a is the desired relatively large homogeneous set.

The induction step is now analogous to the induction step in 1.19 but instead of *low* Δ_{m+1} and *low* Δ_{m+2} one works with LL_m and LL_{m+1}.

(2) First let us prove $I\Sigma_1 \vdash \omega \underset{*}{\rightarrow} (q)^2_k(LL_1)$. It is enough to replace LL_1 by Δ_1 and then relativize.

Assume k standard to be given; we proceed in $I\Sigma_1$. Let $X \in \Delta_1$ be unbounded, $F : [X]^2 \rightarrow k$, $F \in \Delta_1$, q arbitrary. Let T be the tree of all h.m.p.h. sequences; it is unbounded, k-branching and Δ_1. (See 1.13.) By 1.14 (1), for each x there is an upper bound b for all elements $s \in T$ of length $lh(s) \leq x$. By $L\Pi_1$, there is a least such b; call it $H(x)$ and observe that H is $\Sigma_0(\Sigma_1)$.

Take our q and put $G(x) = H(x) * (k+1)$ and $r = G^{k+1}(q)$. (Here we use the fact that k is standard; we may iterate G $(k+1)$-many times.) Clearly, T has elements of arbitrary length; fix an $s \in T$ such that $lh(s) = r$. For $i = 0, \ldots, k+1$, let s_i be the initial segment of s having the length $G^i(q)$. Thus if $r_i = lh(s_i)$ we get $r_{i+1} = G(r_i) = H(r_i) * (k+1)$. Assign colours to elements of s in the usal way and let $col(s_i)$ be the set of colours of elements of s_i. Pick an i such that $col(s_i) = col(s_{i+1})$ and let Z be the set of elements of s_{i+1}. The cardinality of X is $H(r_i) * (k+1)$, therefore for some colour j, the set $a = \{x \in Z \mid x \text{ has the colour } j\}$ has a cardinality bigger than $H(r_i)$. Now $H(r_i)$ is the maximum of elements on the level r_i, thus $H(r_i) \geq s_i \geq \max(s_i)$ (maximum of elements of s_i). And since $col(s_i) = col(s_{i+1})$ we get $\min(a) \leq \max(s_i) \leq H(r_i) < card(a)$. This proves our claim. The induction step is as above. $\square$

1.21 *Proof of Theorem* 1.9. (a) Assume that x, z, q are such that for no y we have $[x, y] \to (q)^n_z$. Thus for each y there is a *counter-example*-mapping $f : [x, y]^n \to z$ with no homogeneous relatively large set having at least q elements. Assume we have fixed a Δ_1 enumeration of $[x, \infty]^n = \bigcup_{y>x}[x, y]^n$ by all numbers such that for each y, $[x, y]^n$ forms an initial segment of length d_y. Then each counterexample is coded by a sequence s of length d_y such for each $i < d_y$, $(s)_i < z$. The set of all counterexamples determines naturally a Δ_1 tree T which is Δ_1-estimated and unbounded; by the low basis theorem (in $I\Sigma_1$) it has a LL_1 unbounded branch. This branch naturally determines a LL_1 mapping $F : [V]^n \to z$ (where V is the set of all numbers) with no relatively large homogeneous set having $\geq q$ elements. But this contradicts 1.7 (1).

(b) Replace n by $(n + 1)$, take k standard and apply 1.7. $\square$

(c) Proofs (of 1.6, 1.8, 1.10)

It remains to prove theorems 1.6, 1.8 and 1.10. The proofs depend neither on the above proofs nor on each other.

1.22 Remark. Observe that for $m, k > 1$, the following are equivalent over $I\Sigma_1$:

$$\omega \to (\omega)^1_{<\omega}(\Delta_m, \Delta_m)$$
$$\omega \to (\omega)^1_{<\omega}(\Delta_m, \Delta_{m+k})$$

(since for a $F : X \to a$ maximal homogeneous sets are just sets $F^{-1}(i)$, $i < a$).

1.23 Lemma. For $m \geq 1$,

$$(*) \qquad I\Sigma_1 + \omega \to (\omega)^1_{<\omega}(\Delta_m, \Delta_m) \vdash B\Sigma_{m+1}\,.$$

Proof. By I.2.23, $B\Sigma_{m+1}$ may be replaced by $R\Pi_{m-1}$. The proof is by induction on m. Let $m = 1$. Let θ be Π_0 and assume $(Cx)(\exists y < a)\theta(x, y)$. Let $\theta'(x, y) \equiv \theta(x, y) \mathbin{\&} (\forall y' < y)\neg\theta(x, y')$ (minimal selector); then θ' is Π_0, and defines a function $F(x) = i \equiv \theta'(x, y) \mathbin{\&} y < a$. F is Π_0, $dom(F)$ is Π_0 and unbounded; by $\omega \to (\omega)^1_a(\Delta_1, \Delta_1)$ we get an $i < a$ such that $F^{-1}(i)$ is unbounded, thus $(Cx)\theta(x, i)$. This proves $R\Pi_0$.

Now assume (*) for m and prove it for $m+1$. Thus assume $\omega \to (\omega)^1_{<\omega}(\Delta_{m+1}, \Delta_{m+1})$. Then $\omega \to (\omega)^1_{<\omega}(\Delta_m, \Delta_m)$, therefore by the induction hypothesis we have $B\Sigma_{m+1}$. We want to prove $R\Pi_m$. Let θ be Π_m and assume $(Cx)(\exists y < a)\theta$. Define θ' as above; by $B\Sigma_{m+1}$ θ' is Δ_{m+1} and the rest is as above. This proves the lemma. □

1.24. Till now we have investigated the combinatorial relation $\omega \to (\omega)^n_{<\omega}$ (Δ_i, Δ_j) (defined in 1.4, cf. 1.5). Denote this relation briefly by $Arrow(n,i,j)$. Let us now consider an apparently weaker partition relation, denoted by

$$\omega \to^0 (\omega)^n_{<\omega}(\Delta_i, \Delta_j) \text{ or briefly } Arrow^0(n,i,j)$$

(thus $\to$ replaced by $\to^0$) whose definition results from the definition of $Arrow(n,i,j)$ by assuming X to be just the whole universe V, thus:

For each Δ_1 function $F : [V]^n \to z$ (where z is any number) there is a Δ_j unbounded homogeneous set.

First consider the case $n = 1$. Evidently, in $I\Sigma_m$ $(m \geq 1)$ we have $Arrow(1,m,m) \equiv Arrow^0(1,m,m)$, since each Δ_m unbounded set is isomorphic with V by a Δ_m mapping (cf. I.2.65). We prove even more:

1.25 Lemma. $I\Sigma_1$ proves the equivalence of $Arrow(1,m,m)$ and $Arrow^0(1,m,m)$.

Proof. By induction on m, show $I\Sigma_1 + Arrow^0(1,m,m) \vdash I\Sigma_m$. Assume this for m and work in $(I\Sigma_1 + Arrow^0(1,m+1,m+1))$. By the induction hypothesis we have $I\Sigma_m$, thus $Arrow(1,m,m)$ and by 1.22, $B\Sigma_{m+1}$. Given a non-empty Σ_{m+1} set X such that $x \in X \equiv (\exists y)\theta(x,y)$ for some Π_m-formula θ, and an $a \in X$, define $F(s) = \min\{x \leq a \mid (\exists y < s)\theta(x,y)\}$ if this set is non empty, $= a+1$ otherwise.

By $B\Sigma_{m+1}$, $i = F(s)$ is Δ_{m+1} (and total) and by $Arrow^0(1,m+1,m+1)$, there is an $i < a+2$ such that $F^{-1}(i)$ is unbounded. This i is $\min X$. This proves $L\Sigma_{m+1}$ and hence $I\Sigma_{m+1}$. □

1.26 Lemma. For $n \geq 2$, $m \geq 1$, $I\Sigma_m$ proves the following:

$$Arrow^0(n,m,n+m) \text{ implies } Arrow^0(n-1,m+1,n+m).$$

Hint: Let $F : [V]^{n-1} \to a$ be Δ_{m+1}; by the limit theorem I.3.2 let $F(x) = \lim_s G(x,s)$ for a Δ_m function G. We may assume $G : [V]^n \to z$. An unbounded homogeneous set for G is homogeneous for F as well.

1.27 Theorem. Over $I\Sigma_1$, the following are equivalent $(n,m \geq 1)$

(i) $B\Sigma_{n+m}$

(ii) $Arrow(n, m, m+n)$, i.e.
$\omega \to (\omega)^n_{<\omega}(\Delta_m, \Delta_{m+n})$
(iii) $Arrow^0(n, m, m+n)$, i.e.
$\omega \to^0 (w)^n_{<\omega}(\Delta_m, \Delta_{m+n})$.

Proof. The only implication to be proved is (iii) $\to$ (i); but for $n = 1$ it follows by 1.25 and 1.23 and for $n > 1$ it follows using 1.26: Indeed assume (iii) for some $n \geq 1$ (and all m – induction hypothesis) and let $Arrow^0(n+1, m, m+n+1)$. Then, in particular, $Arrow^0(1, m, m+n+1)$, thus $Arrow^0(1, m, m)$ and $B\Sigma_{m+1}$; hence we may apply 1.26 and get $Arrow^o(n, m+1, n+m+1)$ hence $B\Sigma_{n+m+1}$ by the induction hypothesis. □

Clearly, Theorem 1.6 follows. Proofs of 1.8 and 1.10 will be sketchy; the reader may elaborate details as an exercise.

1.28 Theorem. For $m \geq 1$, $B\Sigma_{m+1}$ proves that there is a Δ_m mapping $F : [V]^2 \to 2$ (where V is the universe of all numbers) having no Σ_{m+1} o.t.u. set.

Hint: The proof in [Jockusch, 1972-JSL] (Theorem 3.1) formalizes easily and gives a Δ_m mapping $F : V^2 \to 2$ with no o.t.u. Δ_{m+1} homogeneous set. By I.3.24 F has no o.t.u. Σ_{m+1} homogeneous set.

1.29 Corollary. (1) For $m \geq 1$, $I\Sigma_{m+1}$ proves

$$\neg[\omega \to^0 (\omega)^2_2(\Delta_m, \Sigma_{m+1})]\,.$$

(2) Theorem 1.8 follows.

Hint: $I\Sigma_{m+1}$ proves that a Σ_{m+1} set is o.t.u. iff it is unbounded, see I.3.23.

1.30. We sketch a proof of the finite Ramsey theorem. It uses the following lemma:

(*) For each $u, a, q \geq 1$ there is a y such that for each x of cardinality y and each $f : [x]^u \to a$ there is a *prehomogenous* sequence of length q.

Suppose that for given u, a, q (*) does not hold and consider the tree of counterexamples like in 1.21. It is unbounded and Δ_1 estimated; an infinite LL_1 branch determines an infinite LL_1 mapping $F : [V]^u \to a$ with no prehomogeneous sequence of length q. But this contradicts 1.15, 1.16.

Now let us prove the finite Ramsey theorem.

Let $ph(u, a, q)$ be the minimal y satisfying (*); ph is Δ_1 and total. Define

$$rms(1, a, q) = aq$$
$$rms(u+1, a, q) = ph(u+1, a, rms(u, a, q))\,.$$

The function rms is Δ_1, total, and it is easy to show by induction on u that for each x of cardinality $rms(u,a,q)$ and each $f : [x]^u \to a$ there is a homogeneous sequence s (of elements of x) such that $lh(s) = q$. □

2. Instances of the Paris-Harrington Principle and Consistency Statements

(a) Introduction and Statement of Results

2.1 Introduction. In Sect. 1 we introduced the notion of a relatively large finite set (X is relatively large if $\min X < |X|$) and the "arrow" notation $[x,y] \underset{*}{\to} (q)^u_z$ (for each $f : [x,y]^u \to z$ there is a relatively large homogeneous set having at least q elements). We put

$$PH(u,z) \equiv (\forall x,q)(\exists y)([x,y] \underset{*}{\to} ((q)^u_z);$$

the *Paris-Harrington principle* was defined as the statement $(\forall u,z)PH(u+1,z)$. Write (PH) for the last statement and $(PH)_u$ for the formula $(\forall z)$ $PH(u+1,z)$. Paris and Harrington showed that PA proves

$$(PH) \equiv Con^\bullet(PA^\bullet + Tr(\Pi_1^\bullet))$$

where in $Con^\bullet(\ldots)$, $PA^\bullet$ stands for the natural Δ_1 definition of PA and $Tr(\Pi_1^\bullet)$ means the set of all true $\Pi_1^\bullet$-sentences. It follows by Gödel's second incompleteness theorem that (PH) is unprovable in PA. As it was shown above (1.9) for each $n \geq 1$,

$$I\Sigma_n \vdash (PH)_{n-1} \quad \text{(i.e. } I\Sigma_n \vdash (\forall z)PH(\overline{n},z))$$

and $I\Sigma_n$ proves all numerical instances of $(PH)_n$, i.e. for each k, $I\Sigma_n \vdash PH(\overline{n}+1,\overline{k})$.

The question whether the formulas $(PH)_n$ are related to statements assuring the consistency of fragments of $PA(+Tr(\Pi_1))$ is answered as follows by Paris's beautiful refinement of the result of Paris and Harrington:

2.2 Theorem. $I\Sigma_1$ proves that, for each $u \geq 1$,

$$(PH)_u \equiv Con^\bullet(I\Sigma_u^\bullet + Tr(\Pi_1^\bullet)).$$

The proof of this result is the main content of the present section.

2.3 Corollary. (1) For each $n \geq 1$, $(PH)_n$ is provable in $I\Sigma_{n+1}$ but not in $I\Sigma_n$.

(2) $I\Sigma_1$ proves that $(PH) \equiv Con(PA + Tr(\Pi_1))$.

(1) follows by Gödel's second incompleteness theorem, (2) is immediate from 2.2 (and the compactness theorem).

2.4 Discussion. Both $(PH)_n$ and $Con(I\Sigma_n + Tr(\Pi_n))$ are Π_2-statements; thus we have a hierarchy of sentences (1) forming an increasing hierarchy (the n-th of them is provable in $I\Sigma_{n+1}$ but not in $I\Sigma_n$), (2) being syntactically simple (Π_2) and (3) having a well understood double meaning: (a) combinatorial (mathematical), an instance of the Paris-Harrington principle, and (b) logical (metamathematical) – claiming the consistency of $I\Sigma_n^\bullet + Tr^\bullet(\Pi_1)$, which is a certain reflection principle for $I\Sigma_n^\bullet$ (as we shall see later).

Non-provabilities are negative results; but they follow immediately from the positive result 2.1 via Gödel's second incompleteness theorem so it is natural to mention them here.

Let us now present our general plan of the proof. In subsection (b) we prove some combinatorial facts related to $(PH)_u$ and as a by-product we find a simplified formulation of $(PH)_u$. In (c) we prove the implication $Con^\bullet(I\Sigma_u^\bullet + Tr^\bullet(\Pi_1)) \rightarrow (PH)_u$. We shall follow the corresponding proof of $Con(PA + Tr(\Pi_1) \rightarrow (PH)$ due to Paris and Harrington. Paris's proof of the former implication uses properties of α-large sets (α an ordinal) elaborated in the next section. The subsections (d)–(e) contain a proof of $(PH)_u \rightarrow Con(I\Sigma_u^\bullet + Tr(\Pi_1))$, together with various auxiliary things possibly useful elsewhere.

(b) Some Combinatorics

Recall that $PH(u, z)$ means

$$(\forall x, q)(\exists y)([x, y] \underset{*}{\rightarrow} (q)_z^u).$$

Note the obvious monotonicities:

2.5 Lemma ($I\Sigma_1$). If $[x, y] \underset{*}{\rightarrow} (q)_z^u$ and $x' \leq x$, $q' \leq q$, $z' \leq z$ and $y' \geq y$ then $[x', y'] \underset{*}{\rightarrow} (q')_{z'}^u$.

We are going to prove two results:

2.6 Theorem. $I\Sigma_1$ proves that, for each $u \geq 1$,

$$(PH)_u \equiv (\forall z) PH(u + 1, z) \equiv (\forall z)(\exists y)([0, y] \underset{*}{\rightarrow} (u + 2)_z^{u+1}).$$

2.7 Theorem. $I\Sigma_1$ proves that for each $u \geq 1$, $(\forall z)(\exists y)([0, y] \underset{*}{\rightarrow} (u + 2)_z^{u+1})$ implies that for each z there is a y such that for each $f : [0, y]^{u+1} \rightarrow z$ there is an H homogeneous for f and satisfying the following:

$$z \le \min H \le 2^{\min(H)} \le |H| .$$

2.8 Remark. We are going to prove 2.7; our proof follows an analogous proof from [Paris-Harrington]. Then we indicate how to prove 2.6 by the same methods.

The following lemmas are proved in $I\Sigma_1$:

2.9 Lemma. Let $f : [0,b]^e \to c$. $H \subseteq [0,b]$ is homogeneous for f iff each $(e+1)$-element subset of H is.

2.10 Lemma. Let $f_i : [0,b]^e \to c_i$, $i = 1,\dots,k$ and let $f(\mathbf{x}) = \langle f_1(\mathbf{x}),\dots, f_k(\mathbf{x})\rangle$. Then $f : [0,b]^e \to \Pi c_i$ and $H \subseteq [0,b]$ is homogeneous for f iff it is homogeneous for each f_i.

2.11 Lemma. Let $f : [0,b]^e \to c$. Then there is an $f' : [0,b]^{e+1} \to c+1$ such that a set $H \subseteq [0,b]$, $|H| > e+1$, is homogeneous for f iff it is homogeneous for f'.

Proof. For $\mathbf{x} \in [0,b]^{e+1}$ put $f'(\mathbf{x}) = 0$ iff $\mathbf{x}$ is homogeneous for f, $f'(\mathbf{x}) = f(x_0,\dots,x_{e-1})+1$ otherwise. If H is homogeneous for f then clearly $f'(\mathbf{x}) = 0$ for each $\mathbf{x} \in [H]^{e+1}$. Conversely, let H be homogeneous for f'; we prove that the value of f' on $[H]^{e+1}$ is 0, which implies that H is homogeneous for f. Let $\mathbf{x}$ be the least $(e+1)$-tuple in H and $f'(\mathbf{x}) = i = 1 + f(x_0,\dots,x_{e-1})$. Let $y > x_{e+1}$ be another element of H; for each $\mathbf{u} \in [\mathbf{x}]^e$,

$$f'(\mathbf{u},y) = f'(x_0,\dots,x_e) = 1 + f(x_0,\dots,x_{e-1}) = 1 + f(u),$$

thus $\mathbf{x}$ is homogeneous for f, contrary to our assumption. □

2.12 Remark. One can construct an $f' : [0,b]^{e+1} \to 1 + 2\sqrt{c}$ by refining the construction.

2.13 Lemma. For each b, there is an $f : [0,b]^2 \to 8$ such that, for each H relatively large and homogeneous for f,

$$x, y \in H \text{ and } x < y \text{ implies } 2^x < y .$$

Proof. Let $f_0(x,y) = 0$ if $2x < y$, $= 1$ o.w.
$f_1(x,y) = 0$ if $x^2 < y$, $= 1$ o.w.
$f_2(x,y) = 0$ if $2^x < y$, $= 1$ o.w.

Let f combine all these in accordance with Lemma 2.10 and let H be homogeneous and relatively large. Let $a = \min H < |H|$, $e = \max H$, thus $2a \le e$, $f_0(a,e) = 0$ and therefore $f_0(x,y) = 0$ for each $(x,y) \in [H]^2$. Thus $a^2 < e$, $f_1(a,b) = 0$ and therefore $f_1(x,y) = 0$ for each $(x,y) \in [H]^2$. Similarly, $2^a < e$ and $2^x < y$ for each $(x,y) \in [H]^2$. □

2.14 Lemma. For each b, e there is an $f : [0,b]^e \to e+6$ such that, for each H relatively large, homogeneous for f and and each $(x,y) \in [H]^2$ such that $x < y$, we have $2^x < y$.

By Lemmas 2.13 and 2.11.

2.15 Lemma. For each b, e, c there is an $f : [0,b]^e \to c+1$ such that for each H homogeneous for f and such that $|H| \ge e+1$ we have $\min H \ge c$.

Proof. Let $f(x_1, \ldots, x_e) = \min(x_1, c)$. □

2.16 Lemma. For each $f : [a,b]^e \to c$ there is an $\hat{f} : [0,b]^e \to c(c+1)(e+6)$ such that if there is an H homogeneous for $\hat{f}$ and relatively large then there is an H' homogeneous for f such that

$$c \le \min H' \le 2^{\min H'} < |H'|.$$

Proof. Using Lemma 2.15 and Lemma 2.10, replace f by $f_0[0,b]^e \to c(c+1)$ such that each H homogeneous for f_0 is homogeneous for f and satisfies $\min H \ge c$. Let $\log x$ be the maximal u such that $2^u \le x$, let $\log(x_1, \ldots, x_n)$ be $(\log(x_1), \ldots, \log(x_n))$. Define

$$f_1(\mathbf{x}) = f_0(\log(\mathbf{x})) \text{ for } \mathbf{x} \in [0,b]^e.$$
$$p : [0,b]^e \to e+6 \text{ from Lemma 2.14}$$

Let $\hat{f}(\mathbf{x})$ combine f_1, p and let H be relatively large, homogeneous for $\hat{f}$. Then H is homogeneous for f_1, $\min H \ge c$ and we have $2^x < y$ for $(x,y) \in [H]^2$.

Let $H' = \{\log x \mid x \in H\}$; we have $|H'| = |H|$, H' is homogeneous for f_0, $\min H' = \log(\min H)$, thus $2^{\min H'} < |H'|$ as desired. □

2.17 Remark. Theorem 2.7 follows directly. To prove 2.6 assume $(\forall z)(\exists y)([0,y] \underset{*}{\to} (u+2)^{u+1}_z)$ and let x, q, z be given; let $z' = z.\max(x,q)$ and let y be such that $[0,y] \underset{*}{\to} (u+2)^{u+1}_{z'}$. Assume $f : [x,y]^{u+1} \to z$; extend f arbitrarily to an $f_0 : [0,y]^{u+1} \to z$ and combine it with $f_1 : [0,y]^{u+1} \to \max(x,q)$ such that each H homogeneous for f_1 satisfies $\min(H) \ge x, q$. Let $\hat{f}$ be the resulting function and let H be homogeneous for $\hat{f}$, H relatively large, $|H| > u+1$. Then $\min H \ge x, q$ und H is homogeneous for f.

(c) Proof of $Con^{\bullet}(I\Sigma_u^{\bullet} + Tr(\Pi_1^{\bullet})) \rightarrow (PH)_u$ (for $u \geq 1$)

2.18 *Proof.* Recall 1.9: there we proved that for each $n \geq 1$ and for each $x \geq 1$, $I\Sigma_n$ proves $PH(n+1, k)$ (i.e. proves $(\forall x, q)(\exists y)([x, y] \underset{*}{\rightarrow} (q)_k^{n+1})$. The proof of this fact formalizes in $I\Sigma_1$, as an easy inspection shows, so that we have the following:

$$I\Sigma_1 \vdash (\forall z \geq 1)(\forall u \geq 1) Pr^{\bullet}_{I\Sigma_u}(PH(u+1, z)) \qquad (')$$

Now work in $I\Sigma_1 + Con^{\bullet}(I\Sigma_u^{\bullet} + Tr(\Pi_1))$. The added axiom can be evidently reformulated as saying that each Σ_1-sentence provable in $I\Sigma_u$ is true (otherwise its negation would be a true Π_1-sentence inconsistent with $I\Sigma_u$). Now take any x, z, q and observe that, by $(')$, $I\Sigma_u^{\bullet}$ proves $(\exists y)([\dot{x}, y] \underset{*}{\rightarrow} (\dot{q})_{\dot{z}}^{\dot{u}+1})$, which is a $\Sigma_1^{\bullet}$-sentence. Thus this sentence is true (in the sense of satisfaction of $\Sigma_1^{\bullet}$-sentences). But then, by the "it's snowing"-it's snowing lemma, we get $(\exists y)([x, y] \underset{*}{\rightarrow} (q)_z^u)$; we have proved $(\forall z) PH(u+1, z)$.

(d) Strong Indiscernibles

Recall the Σ'_n-formulas introduced in Chap. I, Sect. 2 (e).

2.19 Definition. For each Σ'_n-formula φ having the form

$$(\exists y_1)(\forall y_2)\ldots\varphi_0(\mathbf{x}, \mathbf{y})$$

let $\varphi \restriction$ be the Σ'_n-formula

$$(\exists y_1 \leq v_1)(\forall y_2 \leq v_2)\ldots\varphi_0(\mathbf{x}, \mathbf{y})$$

where $v_1, \ldots, v_n$ are variables not occurring in φ; they are called the *designated variables* of $\varphi \restriction$. Let $(\Sigma'_n) \restriction$ denote the set of all $\varphi \restriction$ for $\varphi \in \Sigma'_n$.

2.20 Observation. Definition 2.19 is meaningful in $I\Sigma_1$; thus in $I\Sigma_1$ we have, for each u, the Δ_1-set of all $(\Sigma'_u \restriction^{\bullet})$-formulas. Moreover, since $(\Sigma'_u \restriction^{\bullet})$-formulas are particular $\Sigma'^{\bullet}_0$-formulas and therefore we have a Δ_1 satisfaction for all $(\Sigma'_u \restriction^{\bullet})$-formulas (with arbitrary u); we denote it occasionally by $\models$.

2.21 Definition (in $I\Sigma_1$). A finite set $B = \{b_i \mid i < \lambda\}$ (increasing enumeration) is a set of *strong indiscernibles* for $(\Sigma'_u \restriction^{\bullet})$-formulas$^{\bullet}$ if for each $i < \lambda$ we have the following:

For each $(\Sigma'_u \restriction^{\bullet})$-formula$^{\bullet}$ $\varphi(\mathbf{x}, \mathbf{v})$ ($\mathbf{v}$ designated), such that $\varphi(\mathbf{x}, \mathbf{v}) < i$, each tuple $\mathbf{p}$ of possible meanings of $\mathbf{x}$, all $\leq b_i$, and each pair $\mathbf{b}, \mathbf{b}' \in (B \setminus [0, b_i])^u$ of increasing u-tuples of elements of B bigger than b_i,

$$(*) \qquad \models \varphi(\mathbf{p}, \mathbf{b}) \equiv \varphi(\mathbf{p}, \mathbf{b}').$$

(Remember that in the last equivalence, φ must be sufficiently small ($\le i$), the parameters $\mathbf{p}$ must be sufficiently small ($\le b_i$) and increasing u-tuples $\mathbf{b}, \mathbf{b}'$ of elements B sufficiently large (all elements $> b_i$).

2.22 Example. Let $u = 3$, let φ be

$$(\exists y_1 \le v_1)(\forall y_2 \le v_2)(\exists y_3 \le v_3)\psi(x_1, x_2, y)\,.$$

Assume $\varphi \le i$; $p_1, p_2 \le b_i$; $i < j < k < q$; $i < j' < k < q$. Then $(*)$ implies

$$\models (\exists y_1 \le b_j)(\forall y_2 \le b_k)(\exists y_3 \le b_q)\psi(p_1, p_2, y)$$

iff

$$\models (\exists y_1 \le b_{j'})(\forall y_2 \le b_{k'})(\exists y_3 \le b_{q'})\psi(p_1, p_2, y)$$

2.23 Theorem. $I\Sigma_1$ proves that, for each u, $(PH)_u$ implies the following: For each ν there is a set B of strong indiscernibles for $(\Sigma'_u \restriction^\bullet)$-formulas such that $|B| = \nu$.

We prove this theorem in the present section. The next section is devoted to a proof of the fact that, for each u, the conclusion of 2.23 (existence of arbitrarily large sets of strong indiscernibles for $(\Sigma'_u \restriction^\bullet)$-formulas$^\bullet$ implies $Con^\bullet(I\Sigma^\bullet_u + Tr(\Pi^\bullet_1))$.

For simplicity, we shall assume $u = 3$. But the method is perfectly general.

2.24 Conventions (only for this section). Define in $I\Sigma_1$ as follows: let $(\exists x \le v)\varphi(x, \mathbf{z})$ be a Σ'_0-formula$^\bullet$ such that v does not occur in φ; let $\mathbf{p}$ be a tuple of possible meanings of $\mathbf{z}$. Then the element *defined* in $[0, d]$ by this formula with parameters $\mathbf{p}$ is the minimal $a \le d$ such that $\models \varphi(a, \mathbf{p})$ (if there is such an a). Dually, the element *defined* in $[0, d]$ by $(\forall x \le v)\varphi(x, \mathbf{z})$ is the minimal $a \le d$ such that $\models \neg\varphi(a, \mathbf{p})$.

If d is a number and $b \subseteq [0, d]$ then $def_q(d, b)$ denotes the set of all elements of $[0, d]$ defined by formulas$^\bullet$ ψ of the above two forms such that $\psi < q$, with parameters from b. (In particular, you may use for ψ any $(\Sigma'_u \restriction^\bullet)$-formula$^\bullet$ or $(\Pi'_u \restriction^\bullet)$-formula$^\bullet$ ($u \ge 1$) w.r.t. its first designated variable.)

Let $\beta < \gamma < \delta < d$ be given. An increasing sequence $(a_q \mid q < \nu)$ of elements less than β is a ***Paris sequence*** (for β, γ, δ, d) if, for each $q < \nu - 1$,
(1) $[a_{q+1}, \beta] \cap def_q(d, [0, a_q] \cup \{\gamma, \delta\}) = \emptyset$,
(2) $[a_{q+1}, \gamma] \cap def_q(d, [0, a_q] \cup \{\delta\}) = \emptyset$,
(3) $[a_{q+1}, \delta] \cap def_q(d, [0, a_q]) = \emptyset$.

2.25 Lemma ($I\Sigma_1$). A Paris sequence is a set of strong indiscernibles for $(\Sigma'_3 \restriction^\bullet)$-formulas$^\bullet$.

Proof. Let $\psi(\mathbf{w}) \equiv (\exists x)(\forall y)(\exists z)\varphi(x,y,z,\mathbf{w})$ where φ is Σ'_0; assume $(\psi \restriction) < i < j < k < q < \nu$, $\mathbf{p} \leq a_i$. Then the following is true (in the sense of $\models$):

$$(\psi \restriction)(\beta,\gamma,\delta,\mathbf{p}) \equiv (\exists x \leq \beta)(\forall y \leq \gamma)(\exists z \leq \delta)\varphi(x,y,z,\mathbf{p}) \equiv$$
$$\equiv (\exists x \leq a_j)(\forall y \leq \gamma)(\exists z \leq \delta)\varphi(x,y,z,\mathbf{p})$$

(since otherwise the smallest x such that $(\forall y \leq \gamma)(\exists z \leq \delta)\varphi(x,y,z,\mathbf{p})$ would be in $[a_j,\beta]$, i.e. in $\mathit{def}_i(d,[0,a_i] \cup \{\gamma,\delta\}) \cap [a_{i+1},\beta]$, which contradicts (1)),

$$\equiv (\exists x \leq a_j)(\forall y \leq a_k)(\exists z \leq \delta)\varphi(x,y,z,\mathbf{p})$$

(since if we let x_0 be such that $x_0 \leq a_j$ and $(\forall y \leq a_k)(\exists z \leq \delta)\varphi(x_0,y,z,\mathbf{p})$ but not $(\forall y \leq \gamma)(\exists z \leq \delta)(\ldots)$, then the minimal y such that $\neg(\exists z \leq \delta)$ $\varphi(x_0,y,z,\mathbf{p})$ would lie in $[a_k,\gamma]$, hence in $\mathit{def}_j(d,[0,a_j] \cup \{\delta\}) \cap [a_{j+1},\gamma]$, which contradicts (2))

$$\equiv (\exists x \leq a_j)(\forall y \leq a_k)(\exists z \leq a_q)\varphi(x,y,z,\mathbf{p})$$

(otherwise let $x_0 \leq a_j$ be minimal such that

$$(\forall y \leq a_k)(\exists z \leq \delta)\varphi(x,y,z,\mathbf{p})$$

and take a $y_0 \leq a_k$ such that

$$\neg(\exists z \leq a_q)\varphi(x_0,y_0,z,\mathbf{p});$$

then the minimal z such that $\varphi(x_0,y_0,z,\mathbf{p})$ would lie in $[a_q,\delta]$, thus in $\mathit{def}_k(d,[0,a_k]) \cap [a_{k+1},\delta]$, which contradicts (3)).

Thus

$$\models (\psi \restriction)(\beta,\gamma,\delta,\mathbf{p}) \equiv (\psi \restriction)(a_i,a_j,a_k,\mathbf{p})$$

for all $i < j < k$ satisfying our condition, which shows that the Paris sequence $(a_q \mid q < \nu)$ is a sequence of strong indiscernibles for $(\Sigma'_3 \restriction)$-formulas. $\square$

To complete the proof of Theorem 2.23, it remains to prove the following

2.26 Lemma $(I\Sigma_1)$. For each u, $(PH)_u$ implies the existence of a Paris sequence of an arbitrary length.

Proof (for $u = 3$). Let the desired length $\nu \geq 5$ of a Paris sequence be given. We assume $(PH)_3$, and use 2.7. Take a sufficiently large c (w.r.t. ν; it turns out that $c = 2^\nu$ is sufficient) and let d be such that for each $f : [0,d]^4 \to c$ there is a homogeneous H such that $c \leq \min H \leq 2^{\min H} \leq |H|$.

For each $\alpha < \beta < \gamma < \delta < d$ define (cf. (1) above)

$$a_0 = 0$$
$$a_{q+1} = \max[\mathit{def}_q(d,[0,a_q] \cup \{\gamma,\delta\}) \cap [0,\beta]] + 1$$

$(q = 0,1,\ldots,\nu-1)$.

Thus $[a_{q+1}, \beta] \cap def_q(d_1, [0, a_q] \cup \{\gamma, \delta\}) = \emptyset$. Observe that if $a_q < \beta$ then $a_q < a_{q+1}$. We want to find $\alpha < \beta < \gamma < \delta < d$ such that the corresponding sequence of a_q's satisfies the following for each $q < \nu - 1$.

(1′) $\quad a_{q+1} \leq \alpha$, i.e. $[\alpha, \beta] \cap def_q(d, [0, a_q] \cup \{\gamma, \delta\}) = \emptyset$,

(2′) $\quad$ (2), i.e. $[\beta, \gamma] \cap def_q(d, [0, a_q] \cup \{\gamma\}) = \emptyset$,

(3′) $\quad$ (3), i.e. $[\gamma, \delta] \cap def_q(d, [0, a_q]) = \emptyset$.

Define the function $F : [0, d]^4 \to c$ as follows: for each $\langle \alpha, \beta, \gamma, \delta \rangle \in [0, d]^4$ let $(a_q \mid q < \nu)$ be the sequence defined above and put

$$\begin{aligned} F(\alpha, \beta, \gamma, \delta) &= (\min q < \nu)([\alpha, \beta] \cap def_q(d, [0, a_q] \cup \{\gamma, \delta\}) \neq \emptyset) \\ &\qquad \text{if there is such a } q, \textit{ else} \\ &= c/4 + (\min q < \nu)([\beta, \gamma] \cap def_q(d, [0, a_q] \cup \{\gamma\}) \neq \emptyset) \\ &\qquad \text{if there is such a } q, \textit{ else} \\ &= c/2 + (\min q < \nu)([\gamma, \delta] \cap def_q(d, [0, a_q]) \neq \emptyset)) \\ &\qquad \text{if there is such a } q, \textit{ else} \\ &= 3c/4 + 1 \end{aligned}$$

Evidently, $F : [0, d]^4 \to c$; if we prove that there is a homogeneous H such that the common value of F on $[H]^4$ is $3c/4 + 1$, then each quadruple $\langle \alpha, \beta, \gamma, \delta \rangle \in H$ determines a Paris sequence of length ν.

Now let H be homogeneous for F and such that $c \leq \min H \leq 2^{\min H} < |H|$; let $\{h_i \mid i < e\}$ be its increasing enumeration. ***First*** assume that the common value of F on H^4 is $q < c/4$. Then for $h_i < h_j < h_k < h_m$ from H we have

$$[h_i, h_j] \cap def_q(d, [0, a_q] \cup \{h_k, h_m\}) \neq \emptyset$$

and since $[h_i, h_j] \cap def_{q-1}(d, [0, a_{q-1}] \cup \{h_k, h_m\}) = \emptyset$ and $a_q = \max[def_{q-1}(d, [0, a_{q-1}] \cup \{h_k, h_m\}) \cap [0, h_j]] + 1$, we get $a_q \leq h_i$. Note that a_q depends on h_j, h_k, h_m but not on h_i; we have $F(h_0, h_j, h_k, h_m) = F(h_0, h_j, h_k, h_m) = F(h_i, h_j, h_k, h_m) = q$ and we get $a_q \leq h_0$. Hence for all i we have

$$[h_i, h_{i+1}] \cap def_q(d, [0, h_0] \cup \{h_{e-2}, h_{e-1}\}) \neq \emptyset.$$

But we have $\leq q$ formulas each with $\leq q$ free variables and $h_0 + 3$ parameters, thus

$$|def_q(d, [0, h_0] \cup \{h_{e-2}, h_{e-1}\})| \leq (q+1).(h_0+3)^q \leq (h_0+3)^{(q+1)}$$

(since $q < \nu < c < h_0$). The last set must intersect each of $(|H| - 3)/2$ disjoint intervals $[h_{2i-1}, h_{2i})$ $(i = 1, \ldots, e-2)$. From $2^{h_0} < |H|$ we get

$2^{h-2} < (2^{h_0} - 3)/2 < (|H| - 3)/2$; thus if we can prove $(h_0 + 3)^{(q+1)} \leq 2^{h_0-2}$, we have a contradiction. Now remember that we took $c = 2^\nu$. Thus $(h_0+3)^{(q+1)} \leq (h_0+3)^\nu$ and $(h_0+3) > h_0 \geq c = 2^\nu$. Hence it suffices to prove that $x \geq 2^\nu$ implies $x^\nu \leq 2^{x-5}$. But evidently this is true for $x = 2^\nu$ if $\nu \geq 5$ and therefore true for each $x \geq 2^\nu$ (if $\nu \geq 5$). Thus we get a contradiction and have excluded the first possibility in the definition of F.

Similarly we eliminate the second and third possibility. Take the second. We already know that $a_q < h_0$ for each q (since the first case does not occur). Assume $F(h_i, h_j, h_k, h_m) = c/4 + q$ for all respective h's. Thus

$$[h_j, h_k] \cap def_q(d, [0, a_q] \cup \{h_m\}) = \emptyset, \text{ thus}$$
$$[h_j, h_{j+1}] \cap def_q(d, [0_0] \cup \{h_{e-1}\}) = \emptyset$$

for all j (as above), which leads to a contradiction. The third case is analogous.

Thus the common value of F on $[H]^4$ is $c/4 + 1$ and therefore for each $\langle \alpha, \beta, \gamma, \delta \rangle \in [H]^4$ the corresponding sequence of a's is a Paris sequence. This completes the proof of Lemma 2.26 and of the Theorem 2.23. □

(e) Final Considerations

2.27 Recall Theorem I.4.37; it will be used to complete the proof of $(\forall m \geq 2)$ $((PH)_{m-1} \rightarrow Con^\bullet(L\Pi'_{m-1}{}^\bullet \cup Tr(\Pi'_1{}^\bullet)))$ in $I\Sigma_1$. (We use $m-1$ instead of u and $L\Pi'_{m-1}$ instead of $I\Sigma'_{m-1}$ to simplify our considerations.) Let S_0 be a finite set of closed instances of $Sk_0(L\Pi'_{m-1}{}^\bullet \cup Tr(\Pi'_1{}^\bullet))$; assuming $(PH)_{m-1}$ we shall construct a Δ_1-satisfaction $\models'$ for S_0 such that $\models'$ extends $\models$ (the usual Δ_1 satisfaction for Σ_0-formulas) and $\models' S_0$. By I.4.37, this implies the desired consistency. The satisfaction $\models'$ is constructed using a sufficently long set of strong $(\Sigma'_{m-1} \restriction^\bullet)$-indiscernibles, guaranteed by 2.26. The definition follows.

2.28 Definition ($I\Sigma_1$). Let Φ be $(Q_1x_1)\ldots(Q_kx_k)\varphi(\mathbf{x}, \mathbf{y})$, $\varphi \in \Sigma_0$. Recall the meaning of $(\Phi \restriction \mathbf{z})(\mathbf{x}, \mathbf{y})$, namely

$$(Qx_1 \leq z_1)\ldots(Q_kx_k \leq z_k)\varphi(\mathbf{x}, \mathbf{y}).$$

Let $A = \{a_q \mid q < \nu\}$ be a finite set in its increasing enumeration. $(\Phi \restriction a_q \rightarrow)$ obviously means the result of substituting $a_q, a_{q+1} \ldots$ for $z_1, z_2 \ldots$ into $(\Phi \restriction \mathbf{z})$, i.e.

$$(Q_1x_1 \leq a_q)\ldots(Q_kx_k \leq a_{q+k-1})\varphi(\mathbf{x}, \mathbf{y}).$$

Given $c_i, \mathbf{d}$, assume that q is the least number such that

(1) $(\Phi \restriction) \leq q$,

(2) $\leftarrow c_{i-1}, \mathbf{d} \leq a_q$,

(3) $q + k \leq \nu$.
If $Q_i = \exists$ put

$$f_i^{\Phi}(\leftarrow c_{i-1}, \mathbf{d}) = (\min c_i \leq a_{q+1})[\models (\Phi^{(i)} \restriction a_{q+2} \rightarrow)(\leftarrow c_i, \mathbf{d})].$$

In all other cases put $f_i^{\Phi}(\leftarrow c_i, \mathbf{d}) = 0$.
Thus we have interpreted all the function symbols of $Sk_0(\Phi)$ by Δ_1 functions (dependent on A). Similarly for a finite set of formulas instead just one. This determines, in the usual way, a Δ_1 satisfaction $\models'$ for any finite set of formulas of the form $Sk_0(\Phi)$ (cf. I.4.14).

2.29 Theorem $(I\Sigma_1)$. Let $m \geq 2$ and $(PH)_{m-1}$. For each finite set S_0 of closed instances of $Sk_0(L\Pi'_{m-1}{}^{\bullet} \cup Tr(\Pi_1'^{\bullet}))$ there is a ν such that if A is a set of strong indiscernibles for $(\Sigma'_{m-1} \restriction^{\bullet})$-formulas of the cardinality ν and $\models'$ is the satisfaction for S_0 given by definition 2.28, then $\models' S_0$.

2.30 Corollary. Theorem 2.2 follows.

Elaboration.

2.31 Lemma $(I\Sigma_1)$. If Φ is Σ'_{m-1} and A is a set of strong $(\Sigma'_{m-1} \restriction^{\bullet})$-indiscernibles, then for *any* q satisfying 2.28 (1)–(3) we have

$$f_i^{\Phi}(\leftarrow c_{i-1}, \mathbf{d}) = (\min c_i \leq a_{q+1})[\models (\Phi^{(i)} \restriction a_{q+2} \rightarrow)(\leftarrow c_i, \mathbf{d})].$$

(Obvious from the definition of strong indiscernibles.)

2.32 Lemma $(I\Sigma_1)$. Let Φ be Σ'_{m-1} or Π'_{m-1} and let $\varphi(\mathbf{s}, \mathbf{u})$ be a closed instance of $Sk_0(\Phi)$. Let A be a set of strong $(\Sigma'_{m-1} \restriction)$-indiscernibles, $\models'$ the corresponding satisfaction and let q satisfy 2.28 (1)–(3) for $\mathbf{c} = V(\mathbf{s})$, $\mathbf{d} = V(\mathbf{a})$. ($V$ is the interpretation of the term s).
Then

$$\models' (\Phi \restriction a_{q+1})(\mathbf{u})) \rightarrow \varphi(\mathbf{s}, \mathbf{u}).$$

Proof. As in the proof of I.4.37 prove

$$\models' (\Phi \restriction a_{q+1}(\mathbf{u}) \rightarrow (\Phi^{(i)} \restriction a_{q+1})(\leftarrow s_i, \mathbf{u}).$$

It suffices to show

$$(') \qquad \models' (\Phi^{(i)} \restriction a_{q+1})(\leftarrow s_i, \mathbf{u}) \rightarrow (\Phi^{(i+1)} \restriction a_{q+1})(\leftarrow s_{i+1}, \mathbf{u}).$$

But $\Phi^{(i)}(\leftarrow x_i, \mathbf{y})$ is $(Q_{i+1} x_{i+1})\Phi^{(i+1)}(\leftarrow x_{i+1}, \mathbf{y})$.

Consider the following two cases:

Case 1. $Q_{i+1} = \forall$. Since $\models' s_{i+1} \le a_q$ (by (2)), we have

$$\begin{aligned}\models' (\Phi^{(i)} \restriction a_{q+1})(\leftarrow s_i, \mathbf{u}) &\rightarrow (\forall x_{i+1} \le a_{q+1})(\Phi^{(i+1)} \restriction a_{q+2})(\leftarrow s_i, x_{i+1}, \mathbf{u}) \\ &\rightarrow (\Phi^{(i+1)} \restriction a_{q+2})(\leftarrow s_{i+1}, \mathbf{u}) \\ &\rightarrow (\Phi^{(i+1)} \restriction a_{q+1})(\leftarrow s_{i+1}, \mathbf{u})\end{aligned}$$

(by indiscernibility).

Case 2. $Q_{i+1} = \exists$. Thus $s_{i+1} = F^{\Phi}_{i+1}(\leftarrow s_i, \mathbf{u})$. Similarly as above, but using also the definition of f^{Φ}_{i+1}, we have $\models' (\Phi^{(i)} \restriction a_{q+1})(\leftarrow s_i, \mathbf{u}) \rightarrow (\exists x_{i+1} \le a_{q+1})(\Phi^{(i+1)} \restriction a_{q+2})(\leftarrow s_i, x_{i+1}, u) \rightarrow (\Phi^{(i+1)} \restriction a_{q+2})(\leftarrow s_{i+1}, \mathbf{u}) \rightarrow (\Phi^{(i+1)} \restriction a_{q+1})(\leftarrow s_{i+1}, \mathbf{u})$. This completes the proof. □

2.33 *Proof of* 2.29. Let a finite set $S_0 \subseteq inst(Sk_0(\Pi'_{m-1}{}^{\bullet} \cup Tr(\Pi'_1{}^{\bullet})))$ be given. Let ν_0 be such that for each instance $\psi(s_1, \dots, s_k) \in S_0$ of the Skolemization of an axiom $\Psi \in L\Pi'_{m-1}{}^{\bullet} \cup Tr(\Pi'_1{}^{\bullet})$ we have

$$(\Psi \restriction) < \nu_0 \text{ and } s_1, \dots, s_k < \nu_0$$

and let $A = \{a_q \mid q < \nu_0 + 3m\}$ be a set of strong $(\Sigma'_{m-1} \restriction^{\bullet})$-indiscernibles. (As we shall see, if $\Psi \in L\Pi'_{m-1}$ then its "prenex normal form with bounded kernel" has $\le 3m$ unbounded quantifiers.) Our aim is to show $\models' S_0$ for the Δ_1-satisfaction given by A.

(1) Let $\Psi \in Tr(\Pi'_1)$, $\Psi = (\forall x)\psi(x)$ where ψ is bounded$'$. Let $\psi(s) \in S_0$; then $\models \psi(V(s))$, i.e. $\models' \psi(s)$.

(2) Now let Ψ be $L\neg\Phi$, where $\Phi(x_1)$ is a Σ'_{m-1}-formula $(\exists x_2) \dots (Q_m x_m)$ $\varphi(x_1, \dots, x_m)$. Thus Ψ is the following:

$$(\forall x_1)[\Phi(x_1) \vee (\exists y_1 \le x_1)(\neg\Phi(y_1) \,\&\, (\forall z_1 < y_1)\Phi(z_1))] .$$

Hence an instance of $Sk_0(\psi)$ has the form

$$\varphi(s) \vee (t_1 \le s_1 \,\&\, \neg\varphi(\mathbf{t}) \,\&\, (r_1 < t_1 \rightarrow \varphi(\mathbf{r})), \text{ in short, } \hat{\psi}(\mathbf{s}, \mathbf{t}, \mathbf{r}) .$$

(Precise conditions on the form of the terms $\mathbf{s}, \mathbf{t}, \mathbf{r}$ will be considered later.)

Let q be minimal such that $(L\neg\Phi \restriction)) \le q$ and $V(\mathbf{s}), V(\mathbf{t}), V(\mathbf{r}) \le a_q$.

We have two cases

(a) $\models' \varphi(\mathbf{s})$; then $\hat{\psi}(\mathbf{s}, \mathbf{t}, \mathbf{r})$ and we are done.

(b) $\models' \neg\varphi(\mathbf{s})$; then, by Lemma 2.32, we have $\models' (\neg\Phi \restriction a_{q+1})(s_1)$.

Let e be the least number such that $\models (\neg\Phi \restriction a_{q+1})(e)$; assume $\mathbf{b} = V(\mathbf{s})$ (i.e. $b_1 = V(s_1)$ etc.).

(3) *Claim.* $V(t_1) = e$.
Indeed, for some $h \leq q$,

$$\begin{aligned} f^L_{m+1}(\mathbf{b}) &= (\min c_1 \leq a_h) \models \varphi(\mathbf{b}) \vee .c_1 \leq b_1 \,\&\, (\neg\Phi \restriction a_{h+1})(c_1) \\ &\qquad \&\, (\forall z_1 < c_1)(\Phi \restriction a_{h+m+1})(z_1) \\ &= (\min c_1 \leq a_h) \models (\neg\Phi \restriction a_{q+1})(c_1)) \,\&\, (\forall z_1 < c_1)(\Phi \restriction a_{q+1})(z_1) \\ &= (\min c_1 \leq a_h) \models (\neg\Phi \restriction a_{q+1})(y_1) \\ &= e\,; \end{aligned}$$

thus $V(t_1) = e$.

(4) Now by Lemma 2.32, $\models (\neg\Phi \restriction a_{q+1})(t_1)$ implies $\models \neg\varphi(t'_1, t'_2, \ldots, t'_m)$ where t'_i is t_i for $i = 1$ or i even and is given by $\neg\Phi$ for i odd, $i > 1$. Our t_i are given by $L_{\neg\Phi}$.

(5) *Claim.* For $i = 1, \ldots, m$, $\models t_i = t'_i$; thus $\models \neg\varphi(\mathbf{t})$.

For i even and $i = 1$ we trivially have $\models t_i = t_i$; for i odd, $i > 1$ we proceed by induction. Now $t'_i = F^{\neg\Phi}_i(\leftarrow t_{i-1})$ and $t_i = F^L_{m+1}(\mathbf{s}, t_{i-1})$. Assume $\mathbf{b} = V(\mathbf{s})$ and $\leftarrow c_{i-1} = V(\leftarrow t_{v-1}) = V(\leftarrow t'_{i-1})$ and compute:

$$\begin{aligned} f^L_{m+i}(\mathbf{b}, \leftarrow c_{i-1}) &= (\min c_i \leq a_p)(\varphi(\mathbf{b}) \vee c_1 \leq b_1 \,\&\, (\neg\Phi^{(i)} \restriction a_{p+1})(\leftarrow c_i) \\ &\qquad \&\, (\forall_1 < c_1)\Phi \restriction a_p(z_1) \\ &= (\min c_i \leq a_p)(\neg\Phi^{(i)} \restriction a_{p+1}(\leftarrow c_i) \\ &= (\min c_i \leq a_q)(\neg\Phi^{(i)} \restriction a_{q+1}(\leftarrow c_i) \\ &= f^{\neg\Phi}_i(\leftarrow c_{i-1})\,. \end{aligned}$$

Thus $V(t_i) = V(t'_i)$.

(6) Now take r_1; if $\models r_1 \geq t_1$ nothing need be proved. Thus assume $V(r_1) < V(t_1) = d_1$; then $\models \Phi \restriction a_{q+1}(t_1)$, therefore $\models \varphi((r'_1, \ldots, r'_m)$ where $r'_i = r_i$ for i odd and similarly to above we prove $\models r_i = r'_i$ for i even. We have proved

(7) *Claim.* Under the present assumption we have $\varphi(\mathbf{r})$. Thus we have proved $\hat{\psi}(\mathbf{s}, \mathbf{t}, \mathbf{r})$, which completes the proof. □

3. Schwichtenberg-Wainer Hierarchy and α-large Sets

In the preceding two sections we studied instances of Paris-Harrington principle and showed (1) that the k-th instance $(PH)_k$ is provable in $I\Sigma_{k+1}$ and (2) that, provably in $I\Sigma_1$, $(PH)_k$ is equivalent to $Con^\bullet(I\Sigma^\bullet_k + Tr^\bullet(\Pi_1))$. (Evidently, (2) implies (1) since $I\Sigma_{k+1}$ proves the above consistency, cf. I.4.33–34, but the explicit proof of $(PH)_k$ in $I\Sigma_{k+1}$ that we presented is of independent interest.)

Now we are going to study a different but related combinatorial principle and its instances. We shall call it (W) or *the principle of α-large invervals.* For background see bibliographical remarks; we shall present the principle and and relate its instances to instances of Paris-Harrington principle. This will be done in the following steps: (a) we introduce ordinals in $I\Sigma_1$ and derive their important properties, (b) we show which induction is sufficient to get enough induction for ordinals, (c) we introduce and study α-large sets, and (d) we define the principle (W) and relate instances of (W) to instances of Paris-Harrington principle. Note that results of this section may be used to get the characterization of functions provably total in $I\Sigma_k$ (and PA) using model theoretic means; this will be done in Chap. IV.

(a) Ordinals in $I\Sigma_1$

3.1. We are going to define in $I\Sigma_1$ a Δ_1 class ε linearly ordered by a Δ_1 ordering $\preccurlyeq$ with a least element 0 and with a Δ_1 operation Σ assigning to each finite non-empty decreasing sequence $\mu_1 \ldots \mu_x$ of elements of ε and each sequence of non-zero numbers $a_1, \ldots, a_x$ of the same length an element of ε denoted by $\sum_{i=1}^{n} \omega^{\mu_i} a_i$; the ordering is related to Σ as follows:

$$\sum_{i=1}^{x} \omega^{\mu_i} a_i \preccurlyeq \sum_{i=1}^{y} \omega^{\nu_i} b_i \text{ iff}$$

(1) there is an $i \leq x, y$ such that $\mu_i \neq \nu_i$ or $a_i \neq b_i$, and, for the least such i, $\mu_i < \nu_i$ or ($\mu_i = \nu_i$ and $a_i < b_i$) or

(2) for each $i \leq x$, $\mu_i = \nu_i$ and $a_i = b_i$.

Furthermore, ε is least Δ_1-class containing 0 and closed under Σ.

This is what we expect from ordinals $< \varepsilon$; we have to show that this can be achieved in $I\Sigma_1$. (We also expect well-order; but, as we shall see, this costs induction.) Thus let us make the following.

3.2 Definition. A *regular tree* is a set t of finite sequences of numbers such that

(i) t contains with each s each initial segment of s, and

(ii) for each i, j, s if $s \frown \langle j \rangle \in t$ and $i < j$ then $s \frown \langle i \rangle \in t$.

(Thus upper neighbours of s in t are $s \frown \langle 0 \rangle$, $s \frown \langle 1 \rangle, \ldots, s \frown \langle i \rangle$ for some i).

A *pre-ordinal* is each regular tree t together with a mapping e (evaluation) assigning to each non-empty $s \in t$ a non-zero number $e(s)$. The *height* of t is the maximum of lengths of elements of t. We define an *operation* Σ applicable to each pair (μ, a) where μ is a non-empty sequence of pre-ordinals and a is a sequence of positive numbers such that $lh(\mu) = lh(a)$.

The pre-ordinal $(t,e) = \sum_1^x \omega^{\mu_i} a_i$ is defined as follows: let $\mu_i = (t_i, e_i)$ and let $t = \bigcup_{i=1}^x \{\langle i \rangle \frown s \mid s \in t_i\} \cup \{\emptyset\}$, $e(\langle i \rangle) = a_i$, and for $\emptyset \neq s \in t_i$ let $e(\langle i \rangle \frown s) = e_i(s)$. (It is easily seen that this corresponds to joining the evaluated trees $\mu_1, \ldots, \mu_x$ over a new root and evaluating the old root of μ_i by a_i.)

Now define total Δ_1 functions $\mathcal{O}_{x,y}, \preccurlyeq_{x,y}$ as follows: $\mathcal{O}_{0,y} = \{\emptyset\}$, $\mathcal{O}_{x+1,y} = \{\sum_1^z \omega^{\mu_i} a_i \mid \mu_i \in \mathcal{O}_{x,y}, \mu_i$ descending $\preccurlyeq_{x,y}$, $a_i \leq y\}$; $\preccurlyeq_{x,y}$ using the obvious modifications of 3.1 (1), (2) above.

3.3 Fact. For each y, $\mathcal{O}_{x,y}$ is a total Δ_1 function of x; $\mathcal{O}_{x,y} \subseteq \mathcal{O}_{x+1,y}$, $\mathcal{O}_{x,y} \subseteq \mathcal{O}_{x,y,+1}$, analogously for $\preccurlyeq_{x,y}$. (Proofs in $I\Sigma_1$ evident).

3.4 Definition.

$$\mathcal{O}_y = \bigcup_x \mathcal{O}_{x,y}, \quad \preccurlyeq_y = \bigcup_x \preccurlyeq_{x,y}$$
$$\Omega_x = \bigcup_y \mathcal{O}_{x,y}, \quad \preccurlyeq'_x = \bigcup_y \preccurlyeq_{x,y}$$
$$\varepsilon = \bigcup_{x,y} \mathcal{O}_{x,y}, \quad \preccurlyeq \text{ is } \bigcup_{x,y} \preccurlyeq_{x,y}$$

3.5 Fact. $\mathcal{O}_y$, $\preccurlyeq_y$, Ω_x, $\preccurlyeq'_x$, ε, $\preccurlyeq$ are Δ_1.
(Evidently, they are Σ_1; but for $\mu = (t,e)$ we have

$$\mu \in \mathcal{O}_y \to \mu \in \mathcal{O}_{x,y} \text{ for } x = \mathit{height}(t)$$
$$\mu \in \Omega_x \to \mu \in \mathcal{O}_{x,y} \text{ for } y = \max(\mathit{range}(e))$$

3.6 Fact. (1) $\preccurlyeq$ linearly orders ε. (2) ε is the smallest Δ_1 class X containing $\emptyset$ and closed under sums $\sum \omega^{\mu_i} a_i$ ($\mu_i \in X$ descending, a_i positive). (3) For each x, each non-empty Δ_1 subset of $\mathcal{O}_x$ has a $\preccurlyeq$-least (i.e. $\preccurlyeq_x$-least) element.

(To prove (2) show that each $\mathcal{O}_{x,y} \subseteq X$; to prove (3) observe that each $\mathcal{O}_{x,y}$ is a finite set and $\mathcal{O}_{x+1,y}$ is an end-extension of $\mathcal{O}_{x,y}$, i.e. each old element precedes each new one.)

3.7 Definition. 0 is *isolated*; $\sum_1^x \omega^{\mu_i} a_i$ is *isolated* if $\mu_x = 0$; otherwise it is a *limit ordinal* (or simply a limit).

3.8 Fact. μ is a limit iff $\mu > 0$ and has no predecessor.

3.9 Fact. Each $\mu \in \varepsilon$ has a successor.

3.10 Definition. For $\alpha = \sum_1^x \omega^{\mu_i} a_i$, $\beta = \sum_1^y \omega^{\nu_i} b_i$ define $\alpha \gg \beta$ iff $\nu_1 \preccurlyeq \mu_x$ (μ_x is the least exponent in α, ν_1 the greatest in β). Further we put $\alpha \gg 0$

if $\alpha \neq 0$. For $\alpha \gg \beta$ define $\alpha + \beta$ as follows: $\alpha + 0 = \alpha$; furthermore, if $\mu_x > \nu_1$ then $\alpha + \beta$ is given by exponents $\mu_1, \ldots, \mu_x, \nu_1, \ldots, \nu_y$ and coefficients $a_1, \ldots, a_x, b_1, \ldots, b_y$; if $\mu_x = \nu_1$ then it is given by exponents $\mu_1, \ldots, \mu_x, \nu_2, \ldots, \nu_y$ and coefficients $a_1, \ldots, a_{x-1}, (a_x + b_1), b_z, \ldots, b_y$.

(Thus e.g. $(\omega^3.3 + \omega^2.4) + (\omega^2.4 + \omega^0.7) = \omega^3.3 + \omega^2.8 + \omega^0.7$.)

3.11 Lemma. Each limit ordinal $\alpha \in \varepsilon$ can be uniquely written in the form $HD + \omega^{\mu}.1$ where ($HD \in \varepsilon$ (head) and $HD \gg \omega^{\mu}.1$) or HD is empty and is disregarded. (Evident.)

3.12 Theorem. There is a Δ_1 function $\{\alpha\}(x)$ defined for each $\alpha \in \varepsilon$ and each x satisfying the following:

(i) If $\alpha = \beta + 1$ then $\{\alpha\}(x) = \beta$; $\{0\}(x) = 0$;
if α is limit, $\alpha = HD + \omega^{\mu}.1$ ($\mu \geq 1$) then

$$\{\alpha\}(x) = \begin{cases} HD + \omega^{\mu-1} & \text{if } \mu \text{ is isolated } (\mu - 1 \text{ is the predecessor}); \\ HD + \omega^{\{\mu\}(x)} & \text{if } \mu \text{ is limit.} \end{cases}$$

For $x = 0$ the member $\omega^{\mu-1}x$ is deleted, thus the result is HD.

(ii) $\alpha \neq 0$ implies $\{\alpha\}(x) < \alpha$;
$\alpha \in \mathcal{O}_{x,y}$ and $z \leq y$ implies $\{\alpha\}(z) \in \mathcal{O}_{x,y}$;
$\beta \gg \alpha > 0$ implies $\{\beta + \alpha\}(x) = \beta + \{\alpha\}(x)$ and $\beta \gg \{\alpha\}(x)$.

Proof. For each $y > 0$, define the function $D_y(\alpha, x) = \{\alpha\}_y(x)$ on $\mathcal{O}_y \times (\leq y)$ by the evident analogues of (i). Show that this is a Δ_1 function with domain $\mathcal{O}_y \times (\leq y)$ and that for $z > y$, D_y is a restriction of D_z. (To this end, show by induction on x that $D_y \restriction (\mathcal{O}_{x,y} \times (\leq y))$ is a well-defined function with range included in $\mathcal{O}_{x,y}$. D_y is Σ_1 and its domain is Δ_1; thus D_y is Δ_1.)

The last claim follows from the evident fact that, under given assumptions, $\beta + \alpha$ is isolated iff α is isolated and if they are limit then for $\alpha = HD + \omega^{\mu}$ we have $\beta + \alpha = (\beta + HD) + \omega^{\mu}$. □

3.13 Definition. $\alpha \underset{x}{\rightarrow} \beta$ means that there is a finite sequence $s = \langle \alpha_0, \ldots, \alpha_r \rangle$ such that $\alpha_0 = \alpha$, $\alpha_r = \beta$ and, for $i < r$, $\alpha_{i+1} = \{\alpha_i\}(x)$. The sequence s is called the *witness* of $\alpha \underset{x}{\rightarrow} \beta$.

3.14 Lemma. (1) $\alpha \underset{x}{\rightarrow} \beta$ is Δ_1. (2) If $X \subseteq \varepsilon$ is Σ_1 or Π_1 then so is $Y = \{\alpha \mid (\forall \beta)(\alpha \underset{x}{\rightarrow} \beta. \rightarrow .\beta \in X)\}$.

Proof. By induction on the length of the witness of $\alpha \underset{z}{\rightarrow} \beta$ show: if $\alpha \in \mathcal{O}_{x,y}$, $z \preccurlyeq x$ and $\alpha \underset{z}{\rightarrow} \beta$ then $\beta \in \mathcal{O}_{x,y}$. Let $F(\alpha, z) = \mathcal{O}_{height(\alpha),z}$ and $G(\alpha, z) =$ the set of all decreasing sequences of elements of $F(\alpha, z)$. F, G are Δ_1 and defined for all $\alpha \in \varepsilon$ and all z. The existential quantifier $(\exists s)$ in the above definition of $\alpha_z \rightarrow \beta$ may be replaced by $(\exists s \in G(\alpha, \max(e, z^*))$ where z^* is the maximal number from all the evaluations in α and z. This shows that $\alpha \underset{z}{\rightarrow} \beta$ is Δ_1. The proof of (2) is similar. □

3.15 Theorem (Properties of $\rightarrow$).
(1) $\alpha \underset{z}{\rightarrow} \beta \underset{z}{\rightarrow} \gamma$ implies $\alpha \underset{z}{\rightarrow} \gamma$
(2) If $\beta \gg \alpha > 0$ and $\alpha \underset{z}{\rightarrow} \gamma$ then $\beta + \alpha \underset{z}{\rightarrow} \beta + \gamma$
(3) $\alpha \underset{z}{\rightarrow} 0$; more generally, if $\alpha \geq \omega^0.k$ and $z \geq k$ then $\alpha \underset{z}{\rightarrow} \omega^0.k$
(4) $x < y$ implies $\omega^\alpha.y \underset{z}{\rightarrow} \omega^\alpha.x$
(5) $\omega^{\delta+1} \underset{z}{\rightarrow} \omega^\delta$ for each $z > 0$
(6) $\alpha \underset{z}{\rightarrow} \beta$ implies $\omega^\alpha \underset{z}{\rightarrow} \omega^\beta$
(7) $x < y$ implies $\{\alpha\}(y) \underset{1}{\rightarrow} \{\alpha\}(x)$
(8) $x < y$ and $\alpha \underset{x}{\rightarrow} \beta$ implies $\alpha \underset{y}{\rightarrow} \beta$
(9) $x \geq 1$, $\alpha \underset{x}{\rightarrow} \beta$ and $\alpha > \beta$ implies $\alpha \underset{x+1}{\rightarrow} \beta + 1$
(10) for $\alpha > 0$ and $x > 1$, $\omega^\alpha \underset{x}{\rightarrow} \omega^{\{\alpha\}(x-1)}.x$
(11) $\alpha \underset{x}{\rightarrow} \beta$ implies $\{\alpha\}(x) \underset{x}{\rightarrow} \{\beta\}(x)$;
if $\alpha > \beta$ and $\alpha \underset{x}{\rightarrow} \beta$ then $\{\alpha\}(x) \underset{x}{\rightarrow} \beta$.

Remark. Needless to say, ω^α stands for $\omega^\alpha.1$, i.e. for the corresponding one-element sum.

Proofs. (2) follows from the last claim in 3.12 (ii). (Inspect, by induction, each number of the witnessing sequence.)

(3) Let $k \leq z \leq x$. Let α be the smallest element of $\mathcal{O}_x$ such that $\alpha \geq \omega^0.k$ and not $\alpha \underset{z}{\rightarrow} \omega^0.k$, then $\alpha > \omega^0.k$, $\{\alpha\}(z) < \alpha$, $\{\alpha\}(z) \in \mathcal{O}_x$ and one shows by checking the defining properties in 3.12 that $\{\alpha\}(z) \geq \omega^0.k$; thus $\{\alpha\}(z) \underset{z}{\rightarrow} \omega^0.k$.

(4) By (3), $\omega^\alpha(y-x) \underset{z}{\rightarrow} 0$; furthermore, $\omega^\alpha.x \gg \omega^\alpha.(y-x)$. Thus, by (2), $\omega^\alpha.y \underset{z}{\rightarrow} \omega^\alpha.x$.

(5) $\omega^{\delta+1} \underset{z}{\rightarrow} \omega^\delta.z \underset{z}{\rightarrow} \omega^\delta$ (by (4)).

(6) It is enough to show: if $\beta = \{\alpha\}(x)$ then $\omega^\alpha \underset{z}{\rightarrow} \omega^\beta$. First assume α a limit: then $\{\omega^\alpha\}(z) = \omega^{\{\alpha\}(z)} = \omega^\beta$. Now let $\alpha = \gamma + 1$, then $\beta = \gamma$ and, by (5), $\omega^\alpha \underset{z}{\rightarrow} \omega^\beta$.

(7) Trivial for α isolated; assume α limit. Let $x, y \leq z$ and $\alpha \in \mathcal{O}_z$; we proceed by induction in $\mathcal{O}_z$. Let $\alpha = HD + \omega^\mu$; then $\omega^\mu \in \mathcal{O}_z$. If HD is non-empty, then the induction assumption gives $\{\omega^\mu\}(y) \underset{1}{\rightarrow} \{\omega^\mu\}(x)$ and the result follows by (2). Thus assume HD empty, $\alpha = \omega^\mu$. If μ is isolated

then see (4); if it is a limit then $\{\omega^\mu\}(y) = \omega^{\{\mu\}(y)} \underset{1}{\rightarrow} \omega^{\{\mu\}(x)} = \{\omega^\mu\}(x)$ by the induction assumption (since $\mu, \{\mu\}(y), \{\mu\}(x) \in \mathcal{O}_z$).

(8) Assume $\alpha \in \mathcal{O}_z$, $z > x, y$, $\beta = \{\alpha\}(x)$ and use induction on α: $\alpha \underset{y}{\rightarrow} \{\alpha\}(y) \underset{1}{\rightarrow} \{\alpha\}(x)$ by (7), thus $\{\alpha\}(y) \underset{y}{\rightarrow} \{\alpha\}(x) = \beta$ and $\alpha \underset{y}{\rightarrow} \beta$.

(9) It suffices to assume $\{\alpha\}(x) = \beta$ and to prove $\alpha \underset{x+1}{\rightarrow} \beta + 1$. To this end, it suffices to prove the following for each fixed z and $x + 1 \le z$, $x \ge 1$:

For each $\alpha \in \mathcal{O}_z$, if $\{\alpha\}(x) = \beta$ then $(\beta + 1) \in \mathcal{O}_z$ and $\alpha \underset{x+1}{\rightarrow} \beta + 1$. This is proved by induction on α. The case of α being isolated is trivial; assume $\alpha = HD + \omega^\delta$. First assume HD non-empty and put $\{\omega^\delta\}(x) = \beta_0$. Then $\{\alpha\}(x) = HD + \beta_0$ and, by the induction assumption, $\omega^\delta \underset{x+1}{\rightarrow} \beta_0 + 1$ and $\beta_0 + 1 \in \mathcal{O}_z$. Thus $\alpha = HD + \omega^\delta \underset{x+1}{\rightarrow} HD + \beta_0 + 1 = \{\alpha\}(x) + 1 = \beta + 1$ and one easily sees that $\beta + 1 \in \mathcal{O}_z$ ($\beta_0 + 1 \in \mathcal{O}_z$ and the last exponent in HD is strictly greater than the exponent of β_0).

Assume HD empty, thus $\alpha = \omega^\delta$. If δ is $\mu + 1$ then $\alpha \underset{x+1}{\rightarrow} \{\mu^\delta\}(x + 1) = \omega^\mu(x+1) = \omega^\mu.x + \omega^\mu \underset{x+1}{\rightarrow} \omega^\mu.x + 1$ (since, by (3), $\omega^\mu \underset{x+1}{\rightarrow} 1$), thus $\alpha \underset{x+1}{\rightarrow} \beta + 1$. If δ is a limit then $\omega^\delta \underset{x+1}{\rightarrow} \omega^{\{\delta\}(x+1)} \in \mathcal{O}_z$ (since $x + 1 \le z$); by (7) and (6) $\omega^{\{\delta\}(x+1)} \underset{1}{\rightarrow} \omega^{\{\delta\}(x)}$ and $\underset{1}{\rightarrow}$ may be replaced by $\underset{x}{\rightarrow}$ (by (8) since $1 \le x$). By the induction hypothesis we get $\omega^{\{\delta\}(x+1)} \underset{x+1}{\rightarrow} \omega^{\{\delta\}(x)} + 1$ which gives $\alpha = \omega^\delta \underset{x+1}{\rightarrow} \omega^{\{\delta\}(x)} + 1 = \beta + 1$ as desired.

(10) Assuming $\alpha > 0$ and $x > 1$ we prove $\omega^\alpha \underset{x}{\rightarrow} \omega^{\{\alpha\}(x-1)}.x$. Clearly, $\alpha \underset{x-1}{\rightarrow} \{\alpha\}(x-1)$ (and $x - 1 \ge 1$); by (9), $\alpha \underset{x}{\rightarrow} \{\alpha\}(x-1) + 1$. Thus, by (6), $\omega^\alpha \underset{x}{\rightarrow} \omega^{\{\alpha\}(x-1)+1} \underset{x}{\rightarrow} \omega^{\{\alpha\}(x-1)}.x$ and we are done.

(11) This is a triviality: $\alpha = \beta$ implies $\{\alpha\}(x) = \{\beta\}(x)$ and $\alpha > \beta$ and $\alpha \underset{x}{\rightarrow} \beta$ implies that the witnessing sequence is $\alpha, \{\alpha\}(x), \ldots, \beta$, thus $\{\alpha\}(x) \underset{x}{\rightarrow} \beta$. This gives $\{\alpha\}(x) \underset{x}{\rightarrow} \{\beta\}(x)$. □

3.16 Definition. For each $\mu \in \varepsilon$ we put $\omega_0^\mu = \mu$, $\omega_{x+1}^\mu = \omega^{\omega_x^\mu}$.

3.17 Remark.

(1) Evidently, if $y \ge 1$ and $\mu \in \mathcal{O}_y$ then $\omega_x^\mu \in \mathcal{O}_y$ for each x.

(2) If $\mu = \omega^0.n$ we shall write ω_x^n instead of ω_x^μ.

(3) Evidently,

$$\omega_x^y \underset{z}{\rightarrow} \omega_x^{y-1} \quad \text{for } y \ge 1 \text{ and } z > 0\,;$$

furthermore, for $x > 0$, and $z > 0$,

$$\omega_x^1 \underset{z}{\rightarrow} \omega_{x-1}^z$$

(note that $\omega_x^0 = \omega_{x-1}^1$ for $x > 0$). The first relation is by 3.15 (5,6), the second from $\omega^1.1 \underset{z}{\to} \omega^0.z$ by (6).

(4) Also evidently, for each $\alpha \in \varepsilon$, and each x, $\alpha \in \Omega_x$ iff $\alpha < \omega_x^1$. (Prove this Π_1-property of x by induction on x.)

(b) Transfinite Induction and Fragments

3.18 Theorem. For each $m, k, n \geq 1$, $I\Sigma_{m+k-1}$ proves the following: each non-empty Σ_m set of ordinals less than ω_k^n has a least element in the ordering $\preccurlyeq$.

Proof. By induction on k. First, for each n, $I\Sigma_m$ proves that each non-empty Σ_m subset of $(\preccurlyeq \omega_1^n)$ has a least element: this can be proved by induction on n. The case $n = 1$ is clear since $I\Sigma_1$ proves that $(< \omega_1^1)$ is Δ_1 isomorphic with the universe of all numbers with $\leq$. It also proves that ω_1^n is Δ_1 isomorphic to the cartesian power of n copies of the universe ordered lexicographically. Assume we have proved on $I\Sigma_m$ the claim for n and let $X \subseteq \underbrace{\omega \times \cdots \times \omega}_{(n+1) \text{ times}}$. Let a be the least number such that, for some n-tuple s, $(\langle a \rangle \frown s) \in X$. Such an a exists since the condition is Σ_m. Let Y be the set of all such sequences s; by the induction assumption, we can prove that Y has a least element s_0 in the lexicographic ordering. Thus $\langle a \rangle \frown s_0$ is least in X; this completes the proof in $I\Sigma_m$.

Now assume we can prove the theorem for k, n and $(m+1)$; we show that it holds for $(k+1)$, n and m. This will complete the whole proof. We proceed in $I\Sigma_{k+m}$. Let $X \neq \emptyset$ be Σ_m, $X \subseteq (\leq \omega_{k+1}^n)$. Define a function F as follows:

$F(0)$ is the minimal (μ_0, a_0) such that $\mu_0 \in (< \omega_k^n)$, $a_0 > 0$ and X contains an element $\omega^{\mu_0}a_0 + \ldots$. (Existence clear; minimality is understood lexicographically, using $\prec$ and $\leq$.) Let $F(x) = (\mu_x, a_x)$ be given; we define $F(x+1)$.

$F(x+1)$ is the minimal (μ_{x+1}, a_{x+1}) such that $\mu_{x+1} \in (\prec \omega_k^n)$, $a_{x+1} > 0$ and X contains an element $\omega^{\mu_0}a_0 + \cdots + \omega^{\mu_x}a_x + \omega^{\mu_{x+1}}a_{x+1} + \cdots$ if there is such an element; otherwise $F(x+1) = F(x)$.

Clearly, such F is well-defined in $I\Sigma_{k+m}$ (it is Δ_{m+1}) and the set $Y = \{\mu_x \mid x\}$ is Σ_{m+1} and non-empty; furthermore, $Y \subseteq (\preccurlyeq \omega_k^n)$. By the induction hypothesis Y has a least element μ; μ is μ_y for some y. But this means that the element $\sum_1^y \omega^{\mu_x} a_x$ is the least element of X. $\square$

Remark. Let $L(\omega_k^n, \Sigma_m)$ be the statement "each non-empty Σ_m subset of $(< \omega_k^n)$ has a least element". Thus we proved, for each $m, k, n \geq 1$,

$$I\Sigma_{m+k-1} \vdash L(\omega_k^n, \Sigma_m).$$

(c) α-large Sets in $I\Sigma_1$

The notion of an α-large set is technically very useful and is also appealing in its own right. In the next subsection (d) we shall show that it is naturally related to the Schwichtenberg-Wainer hierarchy of functions.

3.19 Definition. Let A be a finite set and let $(a_0, \ldots, a_q)$ be its increasing enumeration; we define $\{\alpha\}A$ for each $\alpha \in \varepsilon$. If $A = \emptyset$ then $\{\alpha\}A = \alpha$; otherwise we put

$$\{\alpha\}(a_0, \ldots, a_q) = \{\{\alpha\}(a_0)\}(a_1, \ldots, a_q).$$

(Clearly, this defines $\{\alpha\}A$ as a total Δ_1 function.) A is *α-large* if $\{\alpha\}A = 0$.

3.20 Remark. (1) A is x-large (i.e. $\omega^0.x$-large) if and only if $card(A) \geq x$.

(2) A is ω-large (i.e. $\omega^1.1$-large) iff $card(A) > \min A$.

(3) $A = (a_0, \ldots, a_q)$ is α-large iff $A - (a_0, \ldots, a_i)$ is $\{\alpha\}(a_0, \ldots, a_i)$-large. (Evident.)

3.21 Theorem (1) If A is α-large, $x \leq \min A$ and $\alpha \underset{x}{\rightarrow} \beta$ then A is β-large.

(2) If $A = (a_0, \ldots, a_q)$, $B = (b_0, \ldots, b_r)$, $q \leq r$, for $i \leq q$ we have $b_i \leq a_i$ and A is α-large then B is β-large; (in particular) if $A \subseteq B$ and A is α-large then B is α-large.

(3) Let $\alpha \gg \beta > 0$. Then A is $(\alpha + \beta)$-large iff there are B, C such that $A = B \cup C$, $\max B < \min C$, B is β-large and C is α-large.

(4) Let $\alpha \geq 1$, $A = (a_0, \ldots, a_q)$, $a_0 \geq 2$, $a_0 = x_0 < \ldots < x_{a_0} = a_q$ (i.e. $[a_0, a_q]$ is decomposed into a_0 intervals). If A is ω^α-large, then there is an $i < a_0$ such that $(x_i, x_{i+1}] \cap A$ is $\omega^{\{\alpha\}(a_0-1)}$-large.

(Here

$$(x_i x_{i+1}] = \{z \mid x_i < z \leq x_{i+1}\}.)$$

(5) If $A = (a_0, \ldots, a_q)$ is ω_x^y-large, and $a_0, y \geq 1$ then both $(a, \ldots, a_q)$ and $(a_0, \ldots, a_{q-1})$ are ω_x^{y-1}-large.

Proofs. (1) By 3.15 (8), we may assume $x = \min A$. Put $\alpha_i = \{\alpha\}(a_0, \ldots, a_{i-1})$, $\beta_i = \{\beta\}(a_0, \ldots, a_{i-1})$; thus $\alpha_0 = \alpha$, $\beta_0 = \beta$, $\alpha_{q+1} = 0$, $\alpha_0 \underset{a_0}{\rightarrow} \beta_0$. By 3.15 (11), $\alpha_1 \underset{a_0}{\rightarrow} \beta_1$ thus $\alpha_1 \underset{a_1}{\rightarrow} \beta_1$; similarly we get $\alpha_i \underset{a_i}{\rightarrow} \beta_i$, thus $\alpha_{q+1} \underset{a_q}{\rightarrow} \beta_{q+1}$, i.e. $0 = \alpha_{q+1} \geq \beta_{q+1} = 0$.

(2) Let $A = (a_0, \ldots, a_q)$, $B = (b_0, \ldots, b_r)$, $r \geq q$; we have $b_i \leq a_i$ for $I \leq q$. Let $\alpha_i = \{\alpha\}(a_0, \ldots, a_{i-1})$, $\beta_i = \{\alpha\}(b_0, \ldots, b_{i-1})$, thus $\alpha_{i+1} = \{\alpha_i\}(a_i)$ and similarly for β. First, $\alpha_1 = \{\alpha\}(a_0) \underset{1}{\rightarrow} \{\alpha\}(b_0) = \beta_1$ (for $a_0 \geq 1$ by 3.15 (7)), thus $\alpha_1 \underset{a_0}{\rightarrow} \beta_1$ by 3.15 (8). Now assume

$\alpha_{i+1} \underset{a_i}{\rightarrow} \beta_{i+1}$, $i+1 \leq q$. If $\beta_{i+1} = 0$ then $\{\alpha\}(B) = 0$; if $\beta_{i+1} > 0$ then $\alpha_{i+1} \underset{a_{i+1}}{\rightarrow} \beta_{i+1} \underset{a_{i+1}}{\rightarrow} \{\beta_{i+1}\}(a_{i+1}) \underset{a_{i+1}}{\rightarrow} \{\beta_{i+1}\}(b_{i+1}) = \beta_{i+2}$ and $\alpha_{i+1} > \beta_{i+2}$; by 3.15 (11) $\alpha_{i+2} = \{\alpha_{i+1}\}(a_{i+1}) \underset{a_{i+1}}{\rightarrow} \beta_{i+2}$. Thus in any case $\{\alpha\}B = 0$.

(3) Let $\alpha, \beta > 0$, $\alpha \gg \beta$. First assume $A = (a_0, \ldots, a_q)$ to be $(\alpha+\beta)$-large. Put $\lambda_i = \{\alpha+\beta\}(a_0, \ldots, a_{i-1})$, $\beta_i = \{\beta\}(a_0, \ldots, a_{i-1})$. We have $\lambda_{q+1} = 0$; thus, for some i, $\beta_i = 0$. Let m be minimal such i. Then $\lambda_m = \alpha$, $(a_0, \ldots, a_{m-1}) = B$ is β-large and $(a_m, \ldots, a_q) = C$ is α-large.

Conversely, let $A = B \cup C$, $\max B < \min C$, let B be β-large, C α-large; assume that B is the least possible. Then $\{\beta\}(B) = 0$, $\{\alpha+\beta\}(B) = \alpha$, $\{\alpha+\beta\}(B \cup C) = \{\alpha\}(C) = 0$.

(4) By 3.15 (10), $\omega^\alpha \underset{a_0}{\rightarrow} \omega^{\{\alpha\}(a_0-1)}.a_0$; thus by 3.15 (11), $\{\omega^\alpha\}(a_0) \underset{a_0}{\rightarrow} \omega^{\{\alpha\}(a_0-1)}.a_0$. If $A = (a_0, \ldots, a_q)$ is ω^α-large then $A - (a_0)$ is $\{\omega^\alpha\}(a_0)$-large and therefore $\omega^{\{\alpha\}(a_0-1)}a_0$-large. By (3) this means that $A - (a_0)$ may be decomposed into $B_1, \ldots, B_a$ that are mutually disjoint and such that $\max B_i < \min B_{i+1}$ and each B_i is $\omega^{\{\alpha\}(a_0-1)}$-large. Thus if we have the decomposition $a_0 = x_0 < \cdots < x_{a_0} = a_q$ of $[a_0, a_q]$, at least one half-closed interval $(x_i, x_{i+1}]$ must contain B_i and therefore is $\omega^{\{\alpha\}(a_0-1)}$-large.

(5) Evidently, $(a_1, \ldots, a_q)$ is $\{\omega^y_x\}(a_0)$-large; since $\omega^y_x \underset{a_0}{\rightarrow} \omega^{y-1}_x$ we have $\{\omega^y_x\}(a_0) \underset{a_0}{\rightarrow} \omega^{y-1}_x$ (3.17); the result follows by (1). For $(a_0, \ldots, a_q)$ use (2). □

(d) Schwichtenberg-Wainer Hierarchy

3.22 We shall investigate the hierarchy of number theoretic functions defined informally as follows:

$$f_0(x) = x+1$$
$$f_{\alpha+1}(x) = f^x_\alpha(x+1)$$
$$f_\lambda(x) = f_{\{\lambda\}(x)}(x+1) \text{ for } \lambda \text{ limit}.$$

Here α, λ vary over ordinals $< \varepsilon$; $f^x(y)$ means x-th iteration of f, i.e. $f^0(y) = y$, $f^{x+1}(y) = f(f^x(y))$. This is a variant of the hierarchy investigated by Schwichtenberg and Wainer and the second from two hierarchies investigated by Solovay and Ketonen. We shall show that this hierarchy is Δ_1 definable in $I\Sigma_1$ as a hierarchy of partial functions, show conditions sufficient to prove that a given function f_α is total and relate the hierarchy to α-large sets (cf. Theorem 3.30).

3.23 Remark. (1) It is easy to show in $I\Sigma_1$ that if F is a total one-argument Δ_1 function then there is a unique total two-argument function G such that $G(x,y) = F^x(y)$ for each x, y.

(2) Similarly, $I\Sigma_1$ proves that if q is a finite one-argument function then there is a finite two-argument function q' which is the maximal function such that for each $(x,y) \in dom(q')$, $q'(x,y) = q^x(y)$. Easy proofs are left to the reader.

(3) We shall work with finite two-argument functions defined for *some* pairs (α, x) where $\alpha \in \varepsilon$ and x is a number. If q is such a function then q_α will be the unique function such that for all x, $q_\alpha(x)$ is defined iff $q(\alpha, x)$ is defined and then $q_\alpha(x) = q(\alpha, x)$. (Needless to say, "$q_\alpha(x)$ is defined" means $x \in dom\, q_\alpha$.)

3.24 Definition. Define a predicate $WD(q, \alpha, x, y)$ (read "q is a derivation of $f(x) = y$" or, pedantically, "q is a derivation of the fact that the value of x in the α-th function in the Schwichtenberg-Wainer hierarchy is y") as follows:
(1) q is a finite function, $dom(q) \subseteq \varepsilon \times \omega$;
(2) $q_\alpha(x) = y$
(3) wherever $q_\beta(z)$ is defined then
 (i) if $z > 0$ then $q_\beta(z-1)$ is defined,
 (ii) if $\beta = \gamma + 1$ then $q_\gamma^z(z+1)$ is defined and equal to $q_\beta(z)$,
 (iii) if β is a limit then $q_{\{\beta\}(z)}(z+1)$ is defined and equal to $q_\beta(z)$.

3.25 Lemma. (1) WD is Δ_1. (2) If $WD(q, \alpha, x, y)$, $q_\beta(z)$ is defined and $\beta \underset{z}{\rightarrow} \gamma$ then $q_\gamma(z)$ is defined. (3) $WD(q, \alpha, x, y)$ and $WD(q', \alpha, x, y')$ implies $y = y'$.

Proof (in $I\Sigma_1$) is easy. (1) Note that ε is Δ_1; all "is defined"-quantifiers can be bounded by q. It remains to observe that "$u = f^x(y)$" is Δ_1 in f, u, x, y.

(2) Assume $\gamma = \{\beta\}(z)$. If β is limit then use (iii) and (i); if β is $\beta_0 + 1$ then $\{\beta\}(z) = \beta_0$ and use (ii) and (i). □

3.26 Lemma. Assume $WD(q, \alpha, x, y)$ and $q_\beta(z)$ defined. Then
(1) $q_\beta(z) \geq z + 1$
(2) $w < z$ implies $q_\beta(w) < q_\beta(z)$
(3) $z > 0$, $\beta \underset{z}{\rightarrow} \gamma$ and $\beta \neq \gamma$ implies $q_\beta(z) > q_\gamma(z)$.

Proof. Prove simultaneously (1) & (2) & (3) by induction on β running over all γ such that $q_\gamma \neq \emptyset$. □

3.27 Corollary. Let $WD(q, \alpha, x, y)$, $\alpha \in \mathcal{O}_{u,v}$ and $y \leq v$. Then the restriction q' of q to $\mathcal{O}_{u,v} \times (\leq y)$ satisfies $WD(q', \alpha, x, y)$.

Proof. Verify the conditions 3.23 (i), (ii), (iii) for q' by induction on $\beta \in \mathcal{O}_{u,v}$ using 3.26 (1),(2). □

3.28 Definition. $y = f_\alpha(x)$ iff $(\exists q)\, WD(q, \alpha, x, y)$.

$$[x, y)] = \{z \mid x \leq z < y\}; \quad [(x, y)] = \{z \mid x < z < y\}; \quad \text{similarly } [(x, y]\,.$$

3.29 Remark. (1) Clearly this defines a partial two-argument function. Observe that 3.27 implies that this function is Δ_1.

Note also that by 3.26, the functions satisfy the following whenever defined:

(i) $f_\alpha(z) \geq z+1$,

(ii) $w < z$ implies $f_\alpha(w) < f_\alpha(z)$,

(iii) $z > 0$, $\beta \underset{z}{\rightarrow} \gamma \neq \beta$ implies $f_\beta(z) > f_\gamma(z)$.

(2) Recall that, for $k \geq 1$ and any n, $I\Sigma_{k+1} \vdash L(\omega_k^n, \Sigma_2)$. Using this we may prove in $I\Sigma_{k+1}$ that for each limit $\gamma \prec \omega_k^n$, $\gamma = \sup_x\{\gamma\}(x)$. (This is a Π_2 condition on γ; for its proof, $L(\omega_k^n, \Sigma_2)$ is sufficient: consider the least γ not satisfying this.) Also observe that $I\Sigma_{k+1}$ proves that for each $\alpha \prec \beta \prec \omega_k^n$ there is a z such that $\beta \underset{z}{\rightarrow} \alpha$. (This again is Π_2 in β.)

3.30 Theorem. For $k \geq 0$ and any n, $I\Sigma_{k+1}$ proves the following:

(1) For each $\alpha \preccurlyeq \omega_k^n$, f_α is total.

(2) For each $\alpha \prec \beta \prec \omega_k^n$, there is a z such that

$$(\forall w > z)(f_\alpha(w) < f_\beta(w)).$$

(3) For each $\alpha \prec \omega_k^n$ and each x, $f_\alpha(x)$ is the least y such that the interval $[x, y)]$ is ω^α-large.

Proof. (1) For $k = 0$ this follows by applying n times Remark 3.23. For $k > 0$ observe that the statement in question is Π_2 so that $L(\omega_k^n, \Sigma_2)$ suffices to prove it for all $\alpha \prec \omega_k^n$ (using Remark 3.23).

(2) By 3.29 (2), take a z such that $\beta \underset{z}{\rightarrow} \alpha$; by 3.29 (1) (iii), this implies $f_\alpha(z) < f_\alpha(z)$. If $w > z$ then, by 3.15 (8), we have $\beta \underset{w}{\rightarrow} \alpha$ and therefore $f_\alpha(w) < f_\beta(w)$.

(3) We shall proceed by induction on $\alpha \prec \omega_k^n$; observe that assuming totality of all f_α, $\alpha \prec \omega_k^n$, the assertion in question is Π_1. The case $\alpha = 0$ is clear.

Claim. Assume the assertion of (3) for α; then for each $y \geq 1$ and each x, $[x, f_\alpha^y(x))]$ is the minimal $(\omega^\alpha.y)$-large interval beginning with x. Proof by induction on y (the present assertion is Δ_1 in y). For $y = 1$ this is our assumption; assume the present assumption for $y-1$. Then $[x, z)]$ is $\omega^\alpha.y$-large iff $[f_\alpha(x), z)]$ is $\omega^\alpha.(y-1)$-large (by 3.21 (3)) iff $z \geq f_\alpha^{y-1}(f_\alpha(x)) = f_\alpha^y(x)$. This proves the claim.

Continuing the proof of (3), consider $\alpha + 1$. For $x = 0$ we easily see that $f_{\alpha+1}(0) = 1$, and $\{\omega^{\alpha+1}\}(0) = 0$, thus the one-element set (0) is $\omega^{\alpha+1}$-large. Thus assume $x > 0$ and use the claim: $[x, z)]$ is $\omega^{\alpha+1}$-large iff $[x+1, z)]$ is $\omega^\alpha.x$-large, iff $z \geq f_\alpha^x(x+1) = f_{\alpha+1}(x)$.

It remains to consider α being limit. Then $[x, z)$ is ω^α-large iff $[x+1, z)]$ is $\omega^{\{\alpha\}(x)}$-large iff $z \geq f_{\{\alpha\}(x)}(x+1) = f_\alpha(x)$. This completes the proof. □

3.31 Remark. The reader may verify the following as an exercise:

$$f_1^y(x) = 2^y(x+1) - 1$$
$$f_2(x) = 2^x(x+2) - 1 .$$

3.32 Theorem. $I\Sigma_1$ proves the following: for each α, x, z, $z = f_\alpha(x)$ iff z is minimal such that $[x,z)]$ is ω^α-large.

(Observe that we do not claim that $f_\alpha(x)$ exists, but we claim that *if* it exists and equals z then $[x,z)]$ is ω^α-large and z is minimal with that property; and if there is a z such that $[x,z)]$ is ω^α-large and z is minimal with this property then $f_\alpha(x)$ exists and equals z. Our proof is an inspection of the proof of 3.30 (2).)

Proof. Let $\alpha \in \mathcal{O}_q$ and $x, z \le q \ge 1$. We prove by induction on $\alpha \in \mathcal{O}_q$ the following Δ_1 property of α:

$(*)$ $(\forall z \le q)(\forall x \le q)(z = f_\alpha(x)$ iff z is minimal such that $[x,z)]$ is ω^α-large$)$.

This is clear for $\alpha = 0$. Assume $(*)$ for α and let $\alpha + 1 \in \mathcal{O}_y$.

Claim. For all $1 \le y \le q$, $z \le q$, $z = f_\alpha^y(x)$ iff z is minimal such that $[x,z)]$ is $(\omega^\alpha.y)$-large. (See the proof of 3.30.)

We may assume $x \ge 1$. By the claim, $[x,z)]$ is $\omega^{\alpha+1}$-large iff $[x+1,z)]$ is $\omega^\alpha.x$-large iff $z \ge f_\alpha^x(x+1) = f_{\alpha+1}(x)$ – as in 3.30. Similarly for α being limit. □

3.33 Definition. Let $(W)_u$ be the formula

$$(\forall x, z)(\exists y)([x,y] \text{ is } \omega_u^z\text{-large})$$

(the principle of ordinal-large intervals).

3.34 Facts. (1) $(W)_u$ is a Π_2-formula.

(2) $I\Sigma_1 \vdash (\forall u)((W)_u \equiv (\forall z)(\forall \alpha \prec \omega_{u-1}^z)(f_\alpha \text{ is total}))$

(3) For each $k \ge 0$, $I\Sigma_{k+1} \vdash (W)_k$ and, for each n, $I\Sigma_{k+1} \vdash (\forall x)(\exists y)([x,y]$ is ω_{k+1}^n-large$)$.

((2) follows by Theorem 3.32.)

3.35 Corollary. $I\Sigma_1$ proves the following:

$$(\forall u)(Con(I\Sigma_u + Tr(\Pi_1)) \to (W)_u) .$$

This follows from 3.30 (formalized in $I\Sigma_1$) exactly as the analogous statement in 2.17.

3.36 Theorem. For each k, $I\Sigma_1$ proves $(PH)_k \equiv (W)_k$; thus it proves $(W)_k \equiv Con(I\Sigma_k + Tr(\Pi_1))$.

Comment. One implication (easy) is 3.35. For the converse, thanks to the main result of Sect. 2 is enough to show in $I\Sigma_1$ the following:

$$(*) \qquad (\forall u)((W)_u \rightarrow (PH)_u),$$

or, at least, to prove each instance of this.

Here we have two possibilities:

(a) Solovay and Ketonen proved that, for each $k \geq 1$, $c \geq 2$, $b > a \geq 3$, if $[a, b]$ is ω_k^{c+5}-large then $[a, b] \underset{*}{\rightarrow} (k+2)_c^{k+1}$. If one checks that their proof works in $I\Sigma_1$ (which we expect but have not checked) then the implication $(*)$ is proved.

(b) Paris has a model-theoretic proof of $(W)_k \rightarrow (PH)_k$ (for any standard k). We shall elaborate it in Chap. IV (see IV.3.37).

3.37 Problem. Find a reasonably simple proof of $I\Sigma_1 \vdash (W)_u \rightarrow (PH)_u$ or, at least, $I\Sigma_1 \vdash (\forall u)(W)_u \rightarrow (\forall u)(PH)_u)$ or, alternatively, $I\Sigma_1 \vdash (\forall u)(W)_u \rightarrow Con(PA + Tr(\Pi_1))$. Are details of Solovay-Ketonen's paper dispensable?