ELSEVIER

Theoretical Computer Science 201 (1998) 63–84

Theoretical Computer Science

Ordinal recursive bounds for Higman's theorem[1]

E.A. Cichon[a,*], E. Tahhan Bittar[b,2]

[a] CNRS-CRIN-INRIA-Lorraine, Campus Scientifique, BP 101, 54602 Villers-lès-Nancy, Cedex, France

[b] Laboratoire de Mathematiques Discrètes et Informatique, Institut de Mathématiques-Informatique, Université Claude-Bernard, 43 Bd du 11 Novembre 1918, 69622 Villeurbanne, Cedex, France

Received October 1994; revised October 1996
Communicated by M. Nivat

Abstract

The present paper concerns Higman's theorem for strings generated over a finite alphabet. We give a constructive proof of this theorem and we construct and characterise functions which bound the lengths of bad sequences. These bounding functions are described by ordinal-recursive definitions and their characterisation is achieved with reference to Hardy hierarchies of number-theoretic functions. © 1998 — Elsevier Science B.V. All rights reserved

1. Introduction

Over a number of years now a considerable amount of interest has been shown in both mathematical logic and computer science in combinatorial theorems concerning, in particular, the well-foundedness of certain structures. To those practising mathematical logic the question of what kind of system is needed to be able to express and prove such a theorem is of immense interest, especially if the system turns out to be fairly strong. For the computer scientist, well-founded structures provide a convenient and useful abstraction of the notion of termination of programs. The theorems of Higman and Kruskal have made a significant mark in the theory of term-rewriting, giving rise to methods for proving termination of rewrite systems which have been machine implemented.

The present paper concerns Higman's theorem for strings generated over a finite alphabet. This states that such a set of strings together with the embedding order

* Corresponding author. E-mail: cichon@loria.fr.

[1] The work presented here was effected during the second author's stay in Nancy with the Eureca group, a joint project of the CRIN and INRIA-Lorraine, during the period November 1993–November 1994.

[2] Research supported by a grant from the Universidad de Los Andes, ULA, Venezuela and the Concejo Nacional de Investigaciones Cientificas y Tecnologicas, CONICIT, Venezuela.

PII S0304-3975(97)00009-1

is a well-quasi order. Constructive proofs of this version of Higman's theorem have appeared in [3, 14, 11].

For more details concerning the Higman and Kruskal theorems the reader is referred to the article [6].

We present another constructive proof of Higman's theorem. We construct and characterise functions which bound the lengths of "bad" (or "counterexample") sequences. These bounding functions are described by ordinal-recursive definitions and their characterisation is achieved with reference to Hardy hierarchies of number-theoretic functions.

2. Ordinals and Hardy hierarchies

The *Hardy hierarchy* is defined over ordinal indices by

(i) $H_0(x) = x$,

(ii) $H_{\alpha+1}(x) = H_\alpha(x+1)$,

(iii) $H_\lambda(x) = H_{\lambda_x}(x)$, when λ is a limit ordinal.

The form of the above definition is typical of definitions of number theoretic classes of functions indexed by ordinals (or, more precisely, notations for ordinals) from some set Ω. The clause (iii) above introduces the problem of how to define a function indexed by a limit ordinal, λ. This is overcome by the association of a diagonalisation schema. For this one specifies an increasing sequence of ordinals $\{\lambda_n\}_{n\in\mathbb{N}}$, known as a *fundamental sequence*, which converges to λ. Thus, the limit ordinal λ is identified with a function $\lambda : \mathbb{N} \mapsto \Omega$. Since such a fundamental sequence is by no means unique, there arises the question as to what a *natural* choice of fundamental sequence might be. Furthermore, the nature of the definition of a fundamental sequences can have important consequences on the character of the class of functions obtained.

Note the schematic nature of the above definition. The calculation of values of H_λ will depend on the diagonalisation procedure induced by the definition of a fundamental sequence for λ.

The approach we describe here has evolved from the work of [4] (see also [5]).

2.1. Tree ordinals

Definition 2.1. 1. The set Ω of *countable tree-ordinals*, $\alpha, \beta, \gamma, \ldots$ is defined to be the smallest set X satisfying:

$$\frac{}{0 \in X} \qquad \frac{\alpha \in X}{\alpha + 1 \in X} \qquad \frac{\alpha : \mathbb{N} \mapsto X}{\alpha \in X}$$

2. The *tree-ordering*, $\prec$, on Ω, is defined as the transitive closure of the smallest relation satisfying

$$\frac{}{0 \preccurlyeq \alpha} \qquad \frac{}{\alpha \prec \alpha + 1} \qquad \frac{\alpha : \mathbb{N} \mapsto X}{\forall n \,.\, \alpha(n) \prec \alpha}$$

3. For $x \in \mathbb{N}$, the *pointwise-at-x ordering*, $\prec_x$, on Ω is defined as the transitive closure of the smallest relation satisfying

$$\frac{}{0 \preccurlyeq_x \alpha} \qquad \frac{}{\alpha \prec_x \alpha + 1} \qquad \frac{\alpha : \mathbb{N} \mapsto X}{\alpha(x) \prec_x \alpha}$$

It is clear that $\mathbb{N}$ can be viewed as a subset of Ω, and that the restrictions of $\prec$ and $\prec_x$ to $\mathbb{N}$ correspond to the usual ordering on $\mathbb{N}$. If α is a function from $\mathbb{N}$ into Ω we denote its value at x by α_x, rather than $\alpha(x)$, and we use $\langle \alpha_x \rangle$ as an alternative notation for α.

Definition 2.2. For $\alpha \in \Omega$, the *ordinal*, $\mathrm{ord}(\alpha)$, is defined by: $\mathrm{ord}(0) := 0$, $\mathrm{ord}(\alpha+1) := \mathrm{ord}(\alpha) + 1$, $\mathrm{ord}(\langle \alpha_n \rangle) := \sup_{n \in \mathbb{N}} \{\mathrm{ord}(\alpha_n) + 1\}$.

Lemma 2.3. *If* $\alpha, \beta \in \Omega$ *and* $\beta \prec \alpha$ *then* $\mathrm{ord}(\beta) \prec \mathrm{ord}(\alpha)$.

It follows easily from Lemma 2.3 that $\prec$ is a well-founded partial order on Ω. This is also true of $\prec_x$, for if $\beta \prec_x \alpha$, for some $x \in \mathbb{N}$, then $\beta \prec \alpha$.

3. Arithmetic functions on Ω

3.1. *Addition, multiplication, exponentiation*

Definition 3.1.

Addition	Multiplication	Exponentiation
$\alpha + 0 = \alpha$	$\alpha \cdot 0 = 0$	$\alpha^0 = 1,$
$\alpha + (\beta + 1) = (\alpha + \beta) + 1$	$\alpha \cdot (\beta + 1) = (\alpha \cdot \beta) + \alpha$	$\alpha^{\beta+1} = \alpha^\beta \cdot \alpha,$
$\alpha + \langle \lambda_n \rangle = \langle \alpha + \lambda_n \rangle$	$\alpha \cdot \langle \beta_x \rangle = \langle \alpha \cdot \beta_x \rangle$	$\alpha^{\langle \beta_x \rangle} = \langle \alpha^{\beta_x} \rangle$

Definition 3.2.

$$\omega := \langle x \rangle.$$

3.2. *Cantor normal forms below ε_0*

One first defines ε_0 as follows:

Definition 3.3. For $n \in \mathbb{N}$ define e_n by

$$e_0 := \omega, \qquad e_{n+1} := \omega^{e_n}.$$

Then,

$$\varepsilon_0 := \langle e_n \rangle_{n \in \mathbb{N}}.$$

Definition 3.4. The Cantor normal form for ordinals below ε_0, $\mathcal{CNF}$. First define, for $n \in \mathbb{N}$, the sets $\mathcal{CNF}^n$:

1. If $n \prec \omega$ then $n \in \mathscr{CNF}^0$.
2. If $\alpha \in \mathscr{CNF}^n$ then $\alpha \in \mathscr{CNF}^{n+1}$.
3. If $n_0, \dots, n_k \prec \omega$ and $\alpha_0, \dots, \alpha_k \in \mathscr{CNF}^n$ with $\alpha_0 \prec \cdots \prec \alpha_k$ then $\omega^{\alpha_k}.n_k + \cdots + \omega^{\alpha_0}.n_0 \in \mathscr{CNF}^{n+1}$.

And then define $\mathscr{CNF} := \bigcup_{n\in\mathbb{N}} \mathscr{CNF}^n$.

Theorem 1.

$$\{\alpha : \exists x.\alpha \prec_x \varepsilon_0\} = \mathscr{CNF}.$$

For ordinals in $\mathscr{CNF}$, taking fundamental sequences $\{\alpha_x \mid \alpha \text{ a limit ordinal}, x \in \mathbb{N}\}$ induced by the definitions of the ordinal functions of addition, multiplication and exponentiation, that is $(\alpha + \lambda)_x := \alpha + \lambda_x$, $(\alpha \cdot \lambda)_x := \alpha \cdot \lambda_x$ and $(\alpha^\lambda)_x := \alpha^{\lambda_x}$, one obtains the [17, 18] results $H_{\alpha+\beta}(x) = H_\alpha(H_\beta(x))$, and, taking $\omega_x := x$, as we do throughout this article, we obtain $H_{\omega^\alpha}(x) = F_\alpha(x)$, where $F_0(x) = x + 1$; $F_{\alpha+1}(x) = F_\alpha^x(x)$; $F_\lambda(x) = F_{\lambda_x}(x)$. These functions F_α are often referred to as *fast growing hierarchies*.

Fast growing hierarchies are hierarchies of number theoretic functions generated according to methods developed by [8] and extended in [13, 15–18]. These hierarchies arose out of interest in classifying recursive functions according to their computational complexity. The importance of such hierarchies for logic lies in their use for characterising the provably recursive functions of various formal systems of arithmetic. We briefly enumerate some of these characterisations:

1. For each $n \in \mathbb{N}$, F_n is a primitive recursive function. Its totality can be proved in the fragment $\Sigma_1^0 - IR$ of Peano arithmetic. Here $\Sigma_k^0 - IR$ denotes the subsystem of first order Peano Arithmetic where induction takes the form of a rule which can only be applied to Σ_k^0 formulas.

2. The function F_ω is a version of the *Ackermann Function* and is not primitive recursive. Its totality is not provable in $\Sigma_1^0 - IR$.

3. For each $\alpha \prec \omega^\omega$, F_α is a *multiply recursive function* (as defined in [12]). Its totality is provable in the fragment $\Sigma_2^0 - IR$, but not that of F_{ω^ω}.

4. For each $\alpha \prec \varepsilon_0$, F_α is a function whose totality is provable in Peano arithmetic, but the totality of F_{ε_0} is not provable in Peano arithmetic.

3.3. Generalising Hardy hierarchies

While the set $\mathscr{CNF}$ provides a particularly well behaved system of notations for ordinals, it does not suffice for our needs here because of the severe restriction on the form of term allowed. So, we sacrifice the unicity of notations and the totality of the ordering on terms for other advantages in considering the set $\mathscr{T}$ of terms which we now define.

Definition 3.5. $\mathscr{T}$ is defined as the set of closed terms generated over the signature $\{0, \omega, \mathit{successor}, +, \times, \mathit{exponentiation}\}$.

The set $\mathscr{CNF}$ of terms is a subset of $\mathscr{T}$. From now on we consider our indices as coming from $\mathscr{T}$.

Returning now to the Hardy hierarchies, it turns out that we can easily obtain further useful simplification properties for the H_α's. First we introduce a generalisation of the scheme for successor ordinals:

$$H_0(x)=x; \qquad H_{\alpha+1}(x)=H_\alpha(g(x)); \qquad H_\lambda(x)=H_{\lambda_x}(x).$$

The correspondence mentioned above with the F hierarchy remains true if one modifies the definition of F_0 by $F_0(x) := g(x)$. Now, writing g^α for H_α, we see that the Hardy hierarchy is nothing more than a definition of iteration of a function g extended into the *transfinite*. Thus

$$g^0(x)=x; \qquad g^{\alpha+1}(x)=g^\alpha(g(x)); \qquad g^\lambda(x)=g^{\lambda_x}(x). \tag{1}$$

We shall call g the *control function* for this Hardy hierarchy. The result $H_{\alpha+\beta}(x)=H_\alpha(H_\beta(x))$ now looks obvious, it translates into

$$g^{\alpha+\beta}(x)=g^\alpha(g^\beta(x)). \tag{2}$$

From now on we shall indicate the αth member of a Hardy hierarchy, controlled by g, by g^α.

The above result for ordinal addition naturally suggests that we should be able to extract useful identities with respect to ordinal multiplication and exponentiation. The following uses intuition concerning laws of iteration, but can easily be established directly by transfinite induction.

3.3.1. *Multiplication*

$$g^{\alpha\cdot\beta}=(g^\alpha)^\beta \tag{3}$$

The left-hand side is the $(\alpha\cdot\beta)$th member of a Hardy hierarchy controlled by the function g, the right hand side is the βth member of a Hardy hierarchy controlled by the function g^α. Thus one obtains relationships between different Hardy hierarchies.

3.3.2. *Exponentiation*

Here one gets a generalisation of the result above relating fast growing hierarchies to Hardy hierarchies. The definition of a fast growing hierarchy is appropriately generalised by fixing an ordinal β in advance and putting

$$F_0(x)=g(x); \qquad F_{\alpha+1}(x)=F_\alpha^\beta(x); \qquad F_\lambda(x)=F_{\lambda_x}(x),$$

where $F_\alpha^\beta(x)$ denotes the βth iterate of F_α. The relationship is now

$$g^{\beta^\alpha}=F_\alpha.$$

Putting $\beta=\omega$ with $\omega_x := x$, we extract the result mentioned in the first paragraph. Note that in this case the F hierarchy with $\beta=\omega$ behaves as follows – $F_{\alpha+1}(x)=F_\alpha^\omega(x)=$

$F_\alpha^{\omega_x}(x) = F_\alpha^x(x)$, coinciding with the definition in the first paragraph. In this definition of the hierarchy $\{F_\alpha\}$, the successor clause $F_{\alpha+1} := F_\alpha^\beta$ says that $F_{\alpha+1}$ is obtained by computing the βth member of the Hardy hierarchy based on the control function F_α. This is a generalisation of the traditional schemes.

3.3.3. *Length hierarchies*

We now introduce hierarchies which is closely related to the Hardy hierarchies – the *length* hierarchies. These hierarchies are based on the analysis of Goodstein sequences (see [1, 2]) and can be used to measure precisely the length of such sequences.

Definition 3.6. For a given control function g, define the hierarchy $\{g_\alpha\}$ by

$$g_0(x) = 0; \qquad g_{\alpha+1}(x) = g_\alpha(g(x)) + 1; \qquad g_\lambda(x) = g_{\lambda_x}(x).$$

Then we have

$$g^\alpha(x) = g^{g_\alpha(x)}(x) \tag{4}$$

and, in particular,

$$g_{\alpha+\beta}(x) = g_\alpha(g^\beta(x)) + g_\beta(x). \tag{5}$$

3.4. *The "direct limit" operator, $\Delta_{X\to\omega}[\ ]$*

We saw earlier how notations for ordinals can be induced by defining functions on Ω to obtain $\mathscr{T}$. We now introduce an innovation which consists in the ability to internalise the definition of limits for sequences of ordinal terms generated by a recursive enumeration function of these terms.

If we refer back to the definition of Ω we see that it contains tree-ordinals for *arbitrary* ω-sequences. Since this is excessively general, it is necessary to impose some restrictions on the generation of such sequences. The restriction we introduce here consists in the explicit definition of the sequencing operator $\Delta_{X\to\omega}[\ ]$. $\Delta_{X\to\omega}[\ ]$ is intended to be applicable when a sequence of ordinal terms is given by a recursive definition. The result is an ordinal term which dominates the given sequence.

We describe how this operator is used in the present work. Suppose we have defined a function $F : \mathbb{N} \mapsto \mathscr{T}$ and that, for each $n \in \mathbb{N}$, we have $F(n) = T[{}^X/_n]$ where $T[{}^X/_n]$ denotes the result of taking the ordinal term $T[X]$ with possibly several occurrences of the free variable X and simultaneously substituting them by n. Then $\Delta_{X\to\omega}[F(X)]$ denotes $T[{}^X/_\omega]$. At this point, nothing seems to have been gained as we might just as well have written $F(\omega)$, with an obvious meaning and lightening the notation.

The real effect of introducing $\Delta_{X\to\omega}[\ ]$ comes to light when we define hierarchies of functions over $\mathscr{T}$ which are indexed by elements of $\mathscr{T}$ (as we do later in this paper). Consider the following definition:

$$F_0(\delta) := \delta, \qquad F_{n+1}(\delta) := F_n({}^\delta/_{\omega\cdot\delta}), \qquad F_\omega(\delta) := \Delta_{X\to\omega}[F_X({}^\delta/_{X\cdot\delta})].$$

We easily see that $F_n(\delta)=\omega^n\cdot\delta$. However, $F_\omega(\delta)\neq\omega^\omega\cdot\delta$. Indeed, what we obtain is the following: $F_\omega(\delta) := \Delta_{X\to\omega}[F_X(^{\delta}/_{X\cdot\delta})] = \Delta_{X\to\omega}[(\omega^X\cdot\delta)\{^{\delta}/_{X\cdot\delta}\}] = \Delta_{X\to\omega}[\omega^X\cdot X\cdot\delta] = \omega^\omega\cdot\omega\cdot\delta = \omega^{\omega+1}\cdot\delta$. $\Delta_{X\to\omega}[\ ]$ enables us to extend definitions over transfinite indices in a "discontinuous" way. For this reason we think of $\Delta_{X\to\omega}[\ ]$ as a *direct limit* operator.

The above discussion suffices for our present needs, Some preliminary work developing on the above notions has been reported in [2].

In the sequel we shall derive equations which define functions $\{\alpha_{\omega.\gamma}\mid 1\preccurlyeq\gamma\prec\omega^\omega\}$. These functions $\{\alpha_\Gamma\}$ will be specifiable as functions $\alpha:\{\delta\mid\delta\prec\omega^\omega\}\times\mathscr{T}\mapsto\mathscr{T}$. The fact that each index Γ can be restricted to $\mathscr{CNF}$ corresponds to an evaluation strategy which concentrates on the co-ordinate of smallest complexity in the analysis of products described later in Lemmas 7.3 and 7.4.

We now state some majorisation properties of the hierarchy $\{g_\gamma\}_{\gamma\in\Omega}$. Proving results in the general context of Ω means that these results apply also to $\mathscr{T}$.

Theorem 2. *Suppose that the total function $g:\mathbb{N}\mapsto\mathbb{N}$ satisfies:*
1. *for each $x\in\mathbb{N}$, $x\leqslant g(x)$,*
2. *if $x,y\in\mathbb{N}$ and $x\leqslant y$, then $g(x)\leqslant g(y)$.*

Suppose also that $\alpha\in\Omega\backslash\mathbb{N}$, then, for every $\gamma\in\Omega\backslash\mathbb{N}$ with $\gamma\preccurlyeq\alpha$, we have:
1. *If $x<y$, then $g_\gamma(x)<g_\gamma(y)$.*
2. *If $\delta\prec_x\gamma$, then $g_\delta(x)\leqslant g_\gamma(x)$.*

The majorisation properties of a hierarchy $\{g^\gamma\}_{\gamma\preccurlyeq\alpha\in\Omega}$ can be derived from those of the hierarchy $\{g_\gamma\}_{\gamma\preccurlyeq\alpha\in\Omega}$ using Eq. (4).

4. Higman's theorem – Basic definitions and notation

For $n\in\mathbb{N}$, Σ_n denotes the n-letter alphabet $\{0,\ldots,n-1\}$. No order is presupposed on the letters of Σ_n. Σ_n^* denotes the set of finite words (finite ordered strings) over Σ_n. The number of letters occurring in a word $\mathbf{w}$ is denoted by $|\mathbf{w}|$. The *Higman ordering* $\preccurlyeq_{\Sigma_n^*}$ on Σ_n^* is *embedding*, that is, for $\mathbf{a}$ and $\mathbf{b}$ in Σ_n^*, $\mathbf{a}=a_1\ldots a_p\preccurlyeq_{\Sigma_n^*}\mathbf{b}$ if there are, possibly empty, words $\mathbf{b}_1,\ldots\mathbf{b}_{p+1}$ in Σ_n^* such that $\mathbf{b}=\mathbf{b}_1a_1\mathbf{b}_2a_2\ldots\mathbf{b}_pa_p\mathbf{b}_{p+1}$. We shall always omit the subscript in $\preccurlyeq_{\Sigma_n^*}$ and simply write $\preccurlyeq$. If $\mathbf{a}\preccurlyeq\mathbf{b}$ then it is possible to specify a particular embedding of $\mathbf{a}$ into $\mathbf{b}$, which we call a *left-embedding*, so that $\mathbf{b}=\mathbf{b}_1a_1\mathbf{b}_2a_2\ldots\mathbf{b}_pa_p\mathbf{b}_{p+1}$ and, for each $i\in 1\ldots p$, a_i does not occur in $\mathbf{b}_i$. Clearly $\mathbf{a}$ embeds into $\mathbf{b}$ if, and only if, $\mathbf{a}$ left-embeds into $\mathbf{b}$.

Higman's theorem for strings over a finite alphabet can now be stated as follows:

Higman's theorem. *For every infinite sequence $\{\mathbf{a}_i\}_{i\in\mathbb{N}}$ of words from Σ_n^*, there exist $i<j\in\mathbb{N}$ such that $\mathbf{a}_i\preccurlyeq\mathbf{a}_j$.*

The Cartesian product $\Sigma_{n_1}^*\times\cdots\times\Sigma_{n_k}^*$ is given by the product order:

$$(\mathbf{a}_1,\ldots,\mathbf{a}_k)\preccurlyeq(\mathbf{b}_1,\ldots,\mathbf{b}_k)\text{ if, and only if, }\mathbf{a}_1\preccurlyeq\mathbf{b}_1\wedge\cdots\wedge\mathbf{a}_k\preccurlyeq\mathbf{b}_k.$$

The notation $\boldsymbol{a}$ will be used to denote a sequence $\mathbf{a}_1, \mathbf{a}_2, \ldots$ of elements taken from some space.

We suppose throughout this paper that g is a unary function on $\mathbb{N}$ and that, for all $x \in \mathbb{N}$, $g(x) > x$. Let $x \in \mathbb{N}$ and suppose that $\boldsymbol{a} = \mathbf{a}_0, \mathbf{a}_1, \ldots$ is a sequence in Σ_n^*. We say that $\boldsymbol{a}$ is *controlled* by (g,x) if, for each i, $|\mathbf{a}_i| \leqslant g^i(x)$. This notion of control extends to product spaces as follows: if $\boldsymbol{a} = \mathbf{a}_0, \mathbf{a}_1, \ldots$ is a sequence in $\Sigma_{n_1}^* \times \cdots \times \Sigma_{n_k}^*$ then $\boldsymbol{a}$ is controlled by (g,x) if, for each $i = 1, \ldots, k$, the sequence obtained by projecting the ith co-ordinate of each member of $\boldsymbol{a}$ is controlled by (g,x). For $m_1, \ldots, m_k$ and $n_1, \ldots, n_k \in \mathbb{N}$, the Cartesian product

$$\underbrace{\Sigma_{n_1}^* \times \cdots \times \Sigma_{n_1}^*}_{m_1 \text{ times}} \times \cdots \times \underbrace{\Sigma_{n_k}^* \times \cdots \times \Sigma_{n_k}^*}_{m_k \text{ times}}$$

will be denoted by the ordinal term

$$\omega^{n_1} \cdot m_1 + \cdots + \omega^{n_k} \cdot m_k.$$

In the sequel it will suffice to consider only those product spaces $\Sigma_{n_1}^* \times \cdots \times \Sigma_{n_k}^*$ for which $n_1 \geqslant \cdots \geqslant n_k$. Consequently, an ordinal term σ which represents a space will be an ordinal term in Cantor Normal Form.

Definition 4.1 (*The Higman function*). Let σ be an ordinal term denoting some space, r be a positive integer, and (g,x) be control information as described above. We define $Hig(\sigma, r, g)(x)$ to be the least positive integer N such that every (g,x)-controlled sequence $\mathbf{a}_0, \ldots, \mathbf{a}_N$ from σ contains a $\preccurlyeq$-increasing subsequence of length at least r.

Our aim is to show that the functional Hig is total and to characterise its recursive complexity. Following terminology of Harvey Friedman, the proof of the totality of the functional Hig will be called a *miniaturisation* of Higman's theorem. Note that the totality of $\lambda r.\lambda g.\lambda x.Hig(\sigma, r, g)(x)$ is equivalent to Higman's theorem that σ is a well quasi-order.

Lemma 4.2 (Monotonicity properties for $\lambda r.\lambda g.\lambda x.Hig(\sigma, r, g)(x)$). *If $r_1 \leqslant r_2, g_1 \leqslant g_2$ (i.e. $\forall x. g_1(x) \leqslant g_2(x)$) and $x_1 \leqslant x_2$, then $Hig(\sigma, r_1, g_1)(x_1) \leqslant Hig(\sigma, r_2, g_2)(x_2)$.*

The proof of this lemma is evident. It is worth noting that any sequence controlled by (g_1, x_1) is controlled by (g_2, x_2).

Lemma 4.3. *Suppose that $Hig(\sigma, 2, g)(x)$ exists and that there is an ordinal term α such that $Hig(\sigma, 2, g)(x) \leqslant g_\alpha(x)$. Then, for all $r \geqslant 2$, $Hig(\sigma, r, g)(x)$ exists and, for all $r \geqslant 0$, $Hig(\sigma, r+2, g)(x) \leqslant g_{\alpha \cdot (\alpha+1)^r}(x)$.*

Proof. It is quite easy to describe a general construction for bounding $Hig(\sigma, r, g)(x)$, for $r > 2$, on the basis of the assumption that $Hig(\sigma, 2, g)(x)$ can be determined. The extra hypothesis, that there is an ordinal term α, such that $Hig(\sigma, 2, g)(x) \leqslant g_\alpha(x)$, is convenient in the present context for calculating control information. The construction, by induction on $r \geqslant 2$, is as follows:

Suppose that $\mathbf{a}_1, \mathbf{a}_2, \ldots$ is a sequence taken from the space σ. Define the function F as follows:

$$F(0)=0,$$

$$F(k+1)-F(k)=Hig(\sigma, r, g)(g^{F(k)}(x)).$$

This definition of F ensures that every subsequence $\mathbf{a}_F(k), \mathbf{a}_F(k)+1, \ldots, \mathbf{a}_F(k+1)$ contains an increasing subsequence of length r. Let t_k denote the rth element of the r-subsequence occurring in $\mathbf{a}_F(k), \mathbf{a}_F(k)+1, \ldots, \mathbf{a}_F(k+1)$ and consider the sequence $t_0, t_1, \ldots, t_k, \ldots$. This sequence is controlled by $(G, g^{Hig(\sigma,r,g)(x)}(x))$ where G is determined below. Now, taking $N := Hig(\sigma, 2, G)(g^{Hig(\sigma,r,g)(x)}(x))$ ensures that $t_0, t_1, \ldots, t_N$ contains a 2-subsequence, $t_i \preccurlyeq t_j$, say. Since t_i is the rth element of an r-subsequence, adjoining t_j gives a subsequence of length $r+1$.

A suitable function G is obtained as follows:

Suppose, by induction hypothesis, that $Hig(\sigma, r, g)(x) \leqslant g_\delta(x)$, for some δ. We show that $Hig(\sigma, r+1, g)(x) \leqslant g_{\delta(\alpha+1)}(x)$:

$$\begin{aligned}
g^{F(k+1)}(x) &= g^{Hig(\sigma,r,g)(g^{F(k)}(x))+F(k)}(x) \\
&= g^{Hig(\sigma,r,g)(g^{F(k)}(x))}(g^{F(k)}(x)) \\
&\leqslant_{\mathrm{IH}} g^{g_\delta(g^{F(k)}(x))}(g^{F(k)}(x)) \\
&= g^\delta(g^{F(k)}(x)).
\end{aligned}$$

Hence, $g^{F(k)}(x) = g^{\delta \cdot k}(g^\delta(x)) = (g^\delta)^k(g^\delta(x))$. Thus we can take G to be g^δ. We therefore have

$$\begin{aligned}
G^{Hig(\sigma,2,G)(g^{Hig(\sigma,r,g)(x)}(x))}(G(x)) &\leqslant (g^\delta)^{Hig(\sigma,2,g^\delta)(g^\delta(x))}(g^\delta(x)) \\
&\leqslant (g^\delta)^{(g^\delta)_\alpha(g^\delta(x))}(g^\delta(x)) \\
&= (g^\delta)^\alpha(g^\delta(x)) \\
&= g^{\delta\cdot(\alpha+1)}(x).
\end{aligned}$$

It now follows that, for $0 \leqslant r \in \mathbb{N}$, $Hig(\sigma, r+2, g)(x) \leqslant g_{\alpha\cdot(\alpha+1)^r}(x)$. $\square$

Remark 1. The lemma above is quite general. The problem with this construction in the present context is that it gives us bounds which are too coarse. To obtain improved bounds for $Hig(\sigma, r, g)(x)$ we describe constructions in which r occurs as a parameter.

5. Higman's theorem for Σ_1^*

$(\Sigma_1^*, \preccurlyeq)$ is a special case in our analysis. Clearly $(\Sigma_1^*, \preccurlyeq)$ can be identified with the standard ordering of the natural numbers $(\mathbb{N}, \leqslant)$. The ordinal term denoting Σ_1^*

is ω. Here we are able to determine $Hig(\omega, r, g)(x)$ precisely in terms of Hardy functions.

Lemma 5.1. *Suppose that $\boldsymbol{a} = \mathbf{a}_1, \ldots, \mathbf{a}_N$ is a sequence of natural numbers and that $\boldsymbol{a}$ is controlled by a constant (function), that is, for $i \in 1 \ldots N, \mathbf{a}_i < k$. Then $\boldsymbol{a}$ contains a constant subsequence, and hence an increasing subsequence, of length at least r provided that $N \geqslant k \cdot (r-1) + 1$.*

Proof. Optimality is shown by considering the sequence:

$$\overline{k-1}, \overline{k-2}, \ldots, \overline{0}.$$

Here $\bar{\imath}$ denotes $r-1$ copies of i. Any sequence of length $k \cdot (r-1)$ which does not contain a constant subsequence of length r will be a permutation of this sequence. Appending a natural number smaller than k we obtain a constant subsequence of length r. $\square$

Lemma 5.2. $Hig(\omega, 2, g)(x) = x$.

Proof. The longest possible descending sequence controlled by (g,x), that is, starting from $x-1$, is $x-1, \ldots, 0$, which is of length x, hence $Hig(\omega, 2, g)(x) = x$. It is clear that if g is totally defined then $Hig(\omega, 2, g)(x)$ is totally defined. $\square$

Lemma 5.3. *For $r \geqslant 2, Hig(\omega, r, g)(x)$ is totally defined.*

Proof. The proof will proceed by an induction on r, for $r \geqslant 2$, for which Lemma 5.2 establishes the base case.

So suppose that $Hig(\omega, r, g)(x)$ is totally defined, where $r \geqslant 2$. Let

$$\boldsymbol{a} = \mathbf{a}_0, \ldots, \mathbf{a}_{Hig(\omega,r,g)(x)}, \ldots, \mathbf{a}_{Hig(\omega,r,g)(x)+r \cdot g^{Hig(\omega,r,g)(x)}(x)+1}$$

be a sequence controlled by (g,x). Then $\boldsymbol{a}$ contains an increasing subsequence of length $r+1$. The argument for this is as follows: by the induction hypothesis, the initial part $\mathbf{a}_0, \ldots, \mathbf{a}_{Hig(\omega,r,g)(x)}$ of $\boldsymbol{a}$ contains an increasing subsequence of length r. Let t be the rth element of this increasing subsequence and note that $t < g^{Hig(\omega,r,g)(x)}(x)$ by virtue of the fact that $\boldsymbol{a}$ is controlled by (g,x). Now consider the part of $\boldsymbol{a}$ starting from $\mathbf{a}_{Hig(\omega,r,g)(x)+1}$, that is, the sequence

$$\mathbf{a}_{Hig(\omega,r,g)(x)+1}, \ldots, \mathbf{a}_{Hig(\omega,r,g)(x)+r \cdot g^{Hig(\omega,r,g)(x)}(x)+1}.$$

Its length is $r \cdot g^{Hig(\omega,r,g)(x)}(x) + 1$. If there is no element, $\mathbf{a}_j$ in this sequence such that $t \leqslant \mathbf{a}_j$ then this sequence is controlled by the constant $g^{Hig(\omega,r,g)(x)}(x)$. By Lemma 5.1, it must contain a constant sequence of length $r+1$. $\square$

This argument gives the flavour of those to come in later sections and could (but will not in this case) be used to calculate a bound on $Hig(\omega, r, g)(x)$. It is possible to be precise in characterising $Hig(\omega, r, g)(x)$. This is now undertaken.

Lemma 5.4. *For* $r \geqslant 2$,

$$Hig(\omega, r+1, g)(x) = Hig(\omega, r, g)(x) + Hig(\omega, 2, g)(g^{Hig(\omega, r, g)(x)}(x)).$$

Proof. For $r \geqslant 2$, consider a worst case situation for the given controls (g, x), that of a sequence $\boldsymbol{a} = \mathbf{a}_0, \ldots, \mathbf{a}_N$ of maximum length containing no increasing subsequence of length $r+1$. Within the context of this worst case situation, we present the following definitions and sub-lemmas:

Definition 5.4.1. An increasing subsequence $\mathbf{b}_1, \ldots, \mathbf{b}_k$ of $\boldsymbol{a}$ will be called *maximal* if there is no increasing subsequence of $\boldsymbol{a}$ which strictly contains it, that is, there is no increasing subsequence of $\boldsymbol{a}$ of the form $\boldsymbol{c}_1, \mathbf{b}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_k, \mathbf{b}_k, \boldsymbol{c}_{k+1}$, where at least one of the $\boldsymbol{c}_i$ is non-empty.

Definition 5.4.2. An increasing subsequence of $\boldsymbol{a}$ whose length is r will be called an *r-subsequence.*

Sublemma 5.4.3. *Every r-subsequence in* $\boldsymbol{a}$ *is maximal.*

Proof. This follows directly from the worst-case hypothesis. □

Sublemma 5.4.4. *Every element of* $\boldsymbol{a}$ *must belong to some r-subsequence.*

Proof. Suppose that some element $\mathbf{b}$ of $\boldsymbol{a}$ does not belong to any r-subsequence, then we can extend $\boldsymbol{a}$ to $\boldsymbol{a}'$ by inserting a copy of $\mathbf{b}$ into the position immediately to the right of $\mathbf{b}$'s position, that is,

$$\boldsymbol{a} = \mathbf{a}_0 \ldots b \ldots \mathbf{a}_N$$

$$\boldsymbol{a}' = \mathbf{a}_0 \ldots bb \ldots \mathbf{a}_N$$

To see that this does not create an $r+1$-subsequence, note that any increasing subsequence of $\boldsymbol{a}$ containing $\mathbf{b}$ is at most lengthened by 1 in $\boldsymbol{a}'$. Since we have assumed that any maximal subsequence of $\boldsymbol{a}$ containing $\mathbf{b}$ has length strictly less than r, this procedure would create, at worst, a new increasing subsequence of length r. But $\boldsymbol{a}'$ is longer than $\boldsymbol{a}$ and has the same controls. This contradicts the worst case hypothesis on $\boldsymbol{a}$. □

Definition 5.4.5. For each element $\mathbf{b}$ of $\boldsymbol{a}$ we say that the *status* of $\mathbf{b}$ is i if $\mathbf{b}$ is the ith member of some r-subsequence.

Sublemma 5.4.6. *Every element of* $\boldsymbol{a}$ *has a unique status.*

Proof. Suppose that $\mathbf{b}$ has status i and j with respect to the two r-subsequences

$$\begin{array}{lcl} \mathbf{a}_{p_1} \dots \mathbf{a}_{p_{i-1}} & \multirow{2}{*}{$\mathbf{b}$} & \mathbf{a}_{p_{i+1}} \dots\dots \mathbf{a}_{p_r} \\ \mathbf{a}_{q_1} \dots\dots \mathbf{a}_{q_{j-1}} & & \mathbf{a}_{q_{j+1}} \dots \mathbf{a}_{q_r} \end{array}$$

where $i<j$. Then the sequence

$$\mathbf{a}_{q_1} \dots\dots \mathbf{a}_{q_{j-1}} \ \mathbf{b} \ \mathbf{a}_{p_{i+1}} \dots\dots \mathbf{a}_{p_r}$$

is an increasing subsequence of length greater than r, which is a contradiction. □

Definition 5.4.7. A *terminal element* is an element of $\boldsymbol{a}$ which has status r.

Sublemma 5.4.8. *Terminal elements must occur in strictly decreasing order of size. Furthermore, the terminal elements in* $\boldsymbol{a}$ *form an end-segment of* $\boldsymbol{a}$, *that is, if* $\boldsymbol{a} = \dots \mathbf{a}_i, \mathbf{a}_{i+1}, \dots$ *and* $\mathbf{a}_i$ *is terminal, then so is* $\mathbf{a}_{i+1}$.

Proof. That terminal elements occur in strictly decreasing order of size is obvious, otherwise $\boldsymbol{a}$ would contain an increasing subsequence of length $r+1$. Now, if a terminal element, t, has to its immediate right a non-terminal element, b, then these two can be permuted within $\boldsymbol{a}$ to obtain $\boldsymbol{a}'$

$$\boldsymbol{a} = \mathbf{a}_0 \dots tb \dots \mathbf{a}_N$$

$$\boldsymbol{a}' = \mathbf{a}_0 \dots bt \dots \mathbf{a}_N$$

and we say that *the terminal element has been filtered to the right in* $\boldsymbol{a}$. Note that both $\boldsymbol{a}$ and $\boldsymbol{a}'$ have the same length and controls.[3] Neither has an increasing subsequence of length $r+1$ but $\boldsymbol{a}'$ has more subsequences of length r than $\boldsymbol{a}$. This process can be iterated, by working over the set of terminal elements from right to left, until all the terminal elements have been filtered to the right of non-terminal elements.

But this situation cannot arise if $\boldsymbol{a}$ is indeed a worst case sequence for the given controls. Let t_0 denote the first (i.e. leftmost) terminal element of $\boldsymbol{a}$. If any filtering takes place, then t_0 has to be shifted at least one place to the right and cannot be of the maximum size allowed by the controls for its new position. This means that, after the filtering process has finished, the result $\boldsymbol{a}'$ can be lengthened by the insertion of a word (natural number) larger than t_0 immediately to the left of t_0. The new sequence is longer than, and has the same controls as, $\boldsymbol{a}$ and has no increasing subsequence of length $r+1$, which is a contradiction. Hence the result stated. □

Thus, it should now be clear that the way in which a worst case situation occurs is when t_0 is as large as possible which, in turn, is achieved when the length of the initial part of $\boldsymbol{a}$ up to t_0, $a_0 \dots t_0$, is of maxmum length.[4] In other words, t_0 is $\mathbf{a}_{Hig(\omega,r,g)(x)}$

[3] We emphasise here that it is the fact that $|\mathbf{b}|<|\mathbf{t}|$ which makes filtering possible. This argument breaks down if we try to apply it to spaces other than Σ_1^*.

[4] This tells us that the worst case situation for $(r+1)$-subsequences extends the worst case situation for r-subsequences.

and $|t_0| = t_0 = g^{Hig(\omega,r,g)(x)}(x) - 1$. Appending any element to the sequence $\boldsymbol{a}$ generates an $r+1$-subsequence and so, this element is $\mathbf{a}_{Hig(\omega,r+1,g)(x)}$. So we have

$$\boldsymbol{a} = \mathbf{a}_0, \ldots, \underbrace{\mathbf{a}_{Hig(\omega,r,g)(x)}}_{=t_0}, \underbrace{\mathbf{a}_{Hig(\omega,r,g)(x)+1}}_{=t_0-1}, \ldots, \mathbf{a}_{Hig(\omega,r,g)(x)+\underbrace{g^{Hig(\omega,r,g)(x)}(x)-1}_{=t_0}}$$

Hence

$$Hig(\omega,\, r+1,\, g)(x) = Hig(\omega,\, r,\, g)(x) + g^{Hig(\omega,r,g)(x)}(x)$$

This completes the proof of Lemma 5.4. □

Finally, we have

Lemma 5.5. *For* $r \geqslant 1$,

$$Hig(\omega,\, r+1,\, g)(x) = g_{\omega\cdot r}(x)$$

Proof. By induction on $r \geqslant 1$. For $r=1$ we have $Hig(\omega,\, r+1,\, g)(x) = x$ by Lemma 5.2. By a direct and straightforward calculation, $g_\omega(x) = x$. For the case $r+1$ the induction hypothesis is $Hig(\omega,\, r,\, g)(x) = g_{\omega.(r-1)}(x)$, and we have

$$\begin{aligned} Hig(\omega,r+1,g)(x) &= Hig(\omega,r,g)(x) + Hig(\omega,2,g)(g^{Hig(\omega,r,g)(x)}(x)) \\ &= g_{\omega.(r-1)}(x) + g_\omega(g^{g_{\omega.(r-1)}(x)}(x)) \\ &= g_{\omega.(r-1)}(x) + g_\omega(g^{\omega.(r-1)}(x)) \quad \text{by Eq. (4)} \\ &= g_{\omega+\omega\cdot(r-1)}(x) \quad \text{by Eq. (5)} \\ &= g_{\omega.\,r}(x). \qquad \square \end{aligned}$$

We now define (see the discussion at the beginning of Section 7).

Definition 5.6. For $r \geqslant 1$, $\alpha_\omega(r) := \omega \cdot r$.

By way of example, we have

Example 5.7. Suppose that we have sequences of natural numbers which are controlled by (s,x), where s denotes the successor function, $s(x) = x+1$. Thus, from Lemma 5.1, $Hig(\omega, r+1, s)(x) = s_{\omega\cdot r}(x)$. Now, it is easily established by induction on α that $s^\alpha(x) = s_\alpha(x) + x$. Furthermore, a straightforward calculation gives $s^\omega(x) = 2x$, and so, $s^{\omega\cdot r}(x) = (s^\omega)^r(x) = 2^r.x$ Thus

$$Hig(\omega, r+1, s)(x) = 2^r.x - x = (2^r - 1).x.$$

6. Decomposing a word into a product

We now begin our analysis of $Hig(\Sigma_n^*, r, g)(x)$. The decomposition procedure described in this section is a crucial part of this analysis.

Following an idea of Jullien [9], we describe the circumstances under which a word can be decomposed into an element of a product space.

Suppose that $\mathbf{v}$ and $\mathbf{w}$ are words of Σ_{n+1}^* such that $\mathbf{v} \not\preccurlyeq \mathbf{w}$. Thus any attempt to embed $\mathbf{v}$ into $\mathbf{w}$ must fail. In particular, the attempt to construct a left-embedding must fail. Such an attempt to produce a left-embedding naturally gives rise to a decomposition of $\mathbf{w}$ into a tuple of subwords which can be identified with an element of the product $(\Sigma_n^*)^p$, where p is the number of letters in the word $\mathbf{v}$. To be precise, suppose that $\mathbf{v} = v_1 v_2 \dots v_p$, then $\mathbf{w}$ can be written as

$$\mathbf{w}_1 v_1 \mathbf{w}_2 v_2 \dots \mathbf{w}_i v_i \mathbf{w}_{i+1},$$

where $1 \leqslant i \leqslant p-1$ and, for each $1 \leqslant j \leqslant i+1$, $\mathbf{w}_j$ does not contain the letter v_j. We shall refer to these special occurrences of the letters $v_1, v_2, \dots, v_i$ in $\mathbf{w}$ as *separators* in $\mathbf{w}$. Note that the letters $v_1 v_2 \dots v_i$ of $\mathbf{v}$ which occur as separators in $\mathbf{w}$ always correspond to an *initial* subword of $\mathbf{v}$. Now, each word $\mathbf{w}_j$ is a word in $(\Sigma_{n+1} \backslash \{v_j\})^*$ and so, can be identified with a word $\mathbf{w}_j'$ in Σ_n^*. If necessary, we extend the length of the tuple obtained to p by appending copies of the empty word, Λ. In summary, we make the following transformations:

$$\begin{aligned} \mathbf{w} &\mapsto \mathbf{w}_1 v_1 \mathbf{w}_2 v_2 \dots \mathbf{w}_i v_i \mathbf{w}_{i+1} \\ &\mapsto (\mathbf{w}_1', \mathbf{w}_2', \dots, \mathbf{w}_i', \mathbf{w}_{i+1}') \in (\Sigma_n^*)^{i+1} \\ &\mapsto (\mathbf{w}_1', \mathbf{w}_2', \dots, \mathbf{w}_i', \mathbf{w}_{i+1}', \Lambda, \dots, \Lambda) \in (\Sigma_n^*)^p \end{aligned}$$

We denote this last p-tuple by $\mathscr{D}_{\mathbf{v}}(\mathbf{w})$.

The main point of this decomposition is given by the following lemma.

Lemma 6.1. *Suppose that $a_0, a_1, \dots, a_N$ is a sequence of Σ_{n+1}^*-words, and that, for every $i \in 1 \dots N$, $a_0 \not\preccurlyeq a_i$, where we suppose that a_0 has p letters. Using the method outlined above, we obtain a sequence $\mathscr{D}_{a_0}(a_1), \dots, \mathscr{D}_{a_0}(a_N)$ of elements of $(\Sigma_n^*)^p$, and we have: if $\mathscr{D}_{a_0}(a_1), \dots, \mathscr{D}_{a_0}(a_N)$ contains an increasing sequence of length $p \cdot r + 1$ then $a_1, \dots, a_N$ contains an increasing sequence of length $r+1$.*

Proof. If $\mathscr{D}_{a_0}(a_i) \preccurlyeq \mathscr{D}_{a_0}(a_j)$ then one easily deduces $a_i \preccurlyeq a_j$ provided that the same separators were used in generating $\mathscr{D}_{a_0}(a_i)$ and $\mathscr{D}_{a_0}(a_j)$. The number of separators is always less than p, so, by Lemma 5.1, in any sequence $\mathscr{D}_{a_0}(a_{i_1}), \dots, \mathscr{D}_{a_0}(a_{i_{p \cdot (r-1)+1}})$ at least r of these will have been obtained using the same separators. $\square$

Lemma 6.2. *An obvious property the Higman function with respect to product spaces is that if $p \leqslant q \in \mathbb{N}$ then $Hig((\Sigma_n^*)^p, r, g)(x) \leqslant Hig((\Sigma_n^*)^q, r, g)(x)$.*

Proof. It suffices to note that if $\boldsymbol{a}$ is a sequence taken from $(\Sigma_n^*)^q$ which contains an increasing sequence of length r, then the suppression of some coordinates in each element of $\boldsymbol{a}$ (the same coordinates) results in a sequence from $(\Sigma_n^*)^p$ for some $p \leqslant q$ which also contains an increasing subsequence of length r. □

7. The general case

Our aim is to not only define a bound for $Hig(\sigma, r, g)(x)$ but also to characterise as closely as possible the level of transfinite induction necessary for its proof of totality. In the sequel we shall establish a functional relationship between the ordinal term denoting a space and the ordinal index for the level of the Hardy length-hierarchy of Definition 3.6 at which one obtains bounds for the Higman function for that space. Thus we shall specify a function $\alpha : \mathscr{T} \times \mathbb{N} \mapsto \mathscr{T}$, written $\alpha_\sigma(r)$, such that, for every σ which denotes a space, for all $x \in \mathbb{N}$ and for all $1 \leqslant r \in \mathbb{N}$,

$$Hig(\sigma, r+1, g)(x) \leqslant g_{\alpha_\sigma(r)}(x) \tag{6}$$

Lemma 7.1. *Define the function F as follows*:

$$F(0) = 0,$$

$$F(k+1) - F(k) = Hig(\omega^n \cdot g^{F(k)}(x), g^{F(k)}(x) \cdot r + 1, g)(g^{F(k)}(x)).$$

Then $Hig(\omega^{n+1}, r+1, g)(x) \leqslant F(r)$.

Proof. Recall that ω^{n+1} is an abbreviation for the space Σ_{n+1}^*. So suppose that $a_0, a_1, \ldots$ is a sequence of words from Σ_{n+1}^* and is controlled by (g,x).

Let us first illustrate the argument for the case of an increasing sequence of length 3, that is, we take $r = 2$. This exemplifies the general case because r is a parameter throughout.

Case 1:

$$\nexists j \in 1 \ldots F(1)[a_0 \preccurlyeq a_j]$$

$$a_0 \quad \cdots\cdots\cdots\cdots\cdots\cdots \quad a_{F(1)}$$

The decomposition induced by a_0, the definition of $F(1)$ and the decomposition lemma (Lemma 6.1) guarantee an increasing subsequence of length 3 within the sequence $a_1, \ldots, a_{F(1)}$.

Case 2:

$$\exists j \in 1 \ldots F(1)[a_0 \preccurlyeq a_j] \qquad\qquad \nexists a_k \mid a_j \preccurlyeq a_k$$

$$a_0 \quad \cdots\cdots\cdots\cdots\cdots\cdots \quad a_{F(1)} \quad \cdots\cdots\cdots\cdots\cdots\cdots \quad a_{F(2)}$$

The decomposition induced by a_j, the definition of $F(2)$ and the decomposition lemma (Lemma 6.1) guarantee an increasing subsequence of length 3 in the interval $a_{j+1}, \ldots,$

$a_{F(2)}$. Note that our bound is obtained by treating a_j as if it were $a_{F(1)}$. That this leads to a correct bound follows from Lemma 6.2.

Case 3:

$$\begin{array}{ccccc} & \exists j \in 1 \ldots F(1)[a_0 \preccurlyeq a_j] & & \exists a_k \mid a_j \preccurlyeq a_k & \\ a_0 & \cdots\cdots\cdots\cdots & a_{F(1)} & \cdots\cdots\cdots\cdots & a_{F(2)} \end{array}$$

We have the increasing subsequence $a_0 \preccurlyeq a_j \preccurlyeq a_k$ of length 3.

Coming back to the general case, let $a_{i_0}, \ldots, a_{i_r}$ be a subsequence (of length $r+1$) of $a_0, a_1, \ldots$ where, for $0 \leqslant k \leqslant r$, $i_k \leqslant F(k)$. Then, we have two possibilities:

either $a_{i_0} \preccurlyeq \cdots \preccurlyeq a_{i_r}$ and we are done,

or for some $k<r$, the elements indexed by the interval $[i_k+1, F(k+1)]$ form a subset of $\{b \in \Sigma^*_{n+1} \mid a_{i_k} \not\preccurlyeq b\}$. Thus, in particular, the elements indexed by the interval $[F(k)+1, F(k+1)]$ also form such a subset. Now, the size of this last interval is $F(k+1) - F(k)$, and this value has been chosen, in defining F, precisely so as to be able to apply Lemma 6.1, noting that $|a_{i_k}| \leqslant g^{F(k)}(x)$. Thus the interval $[a_{F(k)+1}, a_{F(k+1)}]$ contains an increasing subsequence of length $r+1$. □

Lemma 7.2.

$$Hig(\omega^{n+1}, r+1, g)(x) \leqslant g_{\Delta_{X\to\omega}[\alpha_{\omega^n \cdot X}(X \cdot r)]\cdot r}(x)$$

Proof. We have shown that

$$Hig(\omega^{n+1}, r+1, g)(x) \leqslant F(r).$$

We now give a Hardy bound for $F(r)$. We can suppose inductively that, for any $\sigma \prec \omega^{n+1}$, for any $1 \leqslant r \in \mathbb{N}$ and for any $x \in \mathbb{N}$, $Hig(\sigma, r+1, g)(x) \leqslant g_{\alpha_\sigma(r)}(x)$.

$$\begin{aligned} F(k+1) - F(k) &= Hig(\omega^n \cdot g^{F(k)}(x), g^{F(k)}(x) \cdot r + 1, g)(g^{F(k)}(x)) \\ &\leqslant_{IH} g_{\alpha_{\omega^n \cdot g^{F(k)}(x)}(g^{F(k)}(x)\cdot r)}(g^{F(k)}(x)) \end{aligned}$$

Thus, we have

$$\begin{aligned} g^{F(k+1)}(x) &\leqslant g^{g_{\alpha_{\omega^n \cdot g^{F(k)}(x)}(g^{F(k)}(x)\cdot r)}(g^{F(k)}(x)) + F(k)}(x) \\ &= g^{g_{\alpha_{\omega^n \cdot g^{F(k)}(x)}(g^{F(k)}(x)\cdot r)}(g^{F(k)}(x))}(g^{F(k)}(x)) \\ &= g^{\alpha_{\omega^n \cdot g^{F(k)}(x)}(g^{F(k)}(x)\cdot r)}(g^{F(k)}(x)), \quad \text{by Eq. (4)} \\ &\leqslant g^{\Delta(r)}(g^{F(k)}(x)), \text{ where } \Delta(r) := \Delta_{X\to\omega}[\alpha_{\omega^n \cdot X}(X \cdot r)]. \end{aligned}$$

From this we obtain

$$g^{F(k)}(x) \leqslant (g^{\Delta(r)})^k(x) = g^{\Delta(r)\cdot k}(x).$$

Thus $g^{F(r)}(x) \leqslant g^{\Delta(r)\cdot r}(x)$ and hence, by Eq. (4), $F(r) \leqslant g_{\Delta(r)\cdot r}(x)$, as required. □

7.1. Bounds for product spaces

The two lemmas which follow deal with the bounding of product spaces. In order to obtain more insight into this problem, let us first consider how a naïve argument might go. For the sake of simplicity, suppose that $a_0, a_1, \dots$ is a sequence of *pairs* taken from $(\Sigma_n^*)^2$. If we concentrate on the first co-ordinate of each pair, we can (inductively) suppose that we can extract, for any given r, an increasing subsequence of length r, that is, we have a subsequence $a_{i_1}, \dots, a_{i_r}$ for which $(a_{i_1})_0 \preccurlyeq \cdots \preccurlyeq (a_{i_r})_0$. If r is taken large enough then presumably we can guarantee that the subsequence obtained in the second co-ordinates, that is, $(a_{i_1})_1, \dots, (a_{i_r})_1$ contains an increasing subsequence of length 2, say. This then gives a subsequence of length 2 in the subsequence $a_{i_1}, \dots, a_{i_r}$ and hence in the original sequence, $a_0, a_1, \dots$.

The problem with this argument is that in order to specify a bounding function we need control information for the extracted subsequence $a_{i_1}, \dots, a_{i_r}$. This is difficult to obtain and we have thus been led to a modified argument.

The key to both of the following lemmas is a function F on indices which is defined so as to cater for two situations; *either* there exists an increasing subsequence, of suitable length, in some chosen co-ordinate for which F provides control information (though F itself is not the control function), thus enabling us to apply the naïve argument above *correctly*, *or* this fails, and the information obtained from this failure enables us to reduce the problem to that for a less complex space. The definition of F is made so that the size of each interval $[F(i), F(i+1)]$ is sufficient to solve such a reduced problem.

Lemma 7.3. *Define the function F as follows*:

$$F(0) = 0$$

$$F(k+1) - F(k) = Hig(\sigma,\ g^{F(k)}(x) \cdot r + 1,\ g)(g^{F(k)}(x))$$

Then

$$Hig(\sigma + \omega, r+1, g)(x) \leqslant F(Hig(\sigma, r+1, G)(x)),$$

where G is a new control function, induced by F. From this we obtain

$$Hig(\sigma + \omega,\ r+1,\ g)(x) \leqslant g_{\alpha_{\sigma+\omega}(r)}(x)$$

where $\alpha_{\sigma+\omega}(r) := \Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)] \cdot \alpha_\sigma(r)$.

Proof. $\sigma + \omega$ denotes a product space whose last component is Σ_1^*. Suppose that $a_0, a_1, \dots$ is a sequence in $\sigma + \omega$. We concentrate on the last component. The projection in the last component of this sequence is a sequence in Σ_1^*, which we call a *principal* sequence. We call the sequence obtained by omitting the last component the *passenger* sequence. We can extract an increasing subsequence of arbitrary length from the principal sequence. Now,

either a sufficiently long initial part of this principal subsequence is such that, for each i, the ith element occurs at or before position $F(i)$ and hence is controlled by (G,x), in which case the corresponding passenger subsequence is also controlled by (G,x). Taking "sufficiently long" to mean the value of $Hig(\sigma, r+1, G)(x)$ will ensure that the passenger subsequence has an increasing subsequence of length $r+1$, thus we obtain the conclusion above.

or the principal subsequence fails to be controlled by (G,x) before the $Hig(\sigma, r+1, G)(x)$th element. But this means that in some interval $[F(k), F(k+1)]$ all the elements of the principal sequence with indices in this interval are bounded in size by the constant $g^{F(k)}(x)$. The size of this interval has been chosen to guarantee that the passenger sequence contains an increasing sequence of length $g^{F(k)}(x) \cdot r + 1$. The corresponding elements of the principal sequence will, by Lemma 5.1, contain a constant sequence of length r and our first claim will follow once we have specified G. Now,

$$\begin{aligned} F(k+1) - F(k) &= Hig(\sigma, g^{F(k)}(x) \cdot r + 1, g)(g^{F(k)}(x)) \\ &\leqslant_{IH} g_{\alpha_\sigma(g^{F(k)}(x) \cdot r)}(g^{F(k)}(x)) \\ &\leqslant g_{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)]}(g^{F(k)}(x)) \end{aligned}$$

Thus, we have

$$\begin{aligned} g^{F(k+1)}(x) &\leqslant g^{g_{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)]}(g^{F(k)}(x)) + F(k)}(x) \\ &= g^{g_{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)]}(g^{F(k)}(x))}(g^{F(k)}(x)) \\ &= g^{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)]}(g^{F(k)}(x)) \quad \text{by Eq. (4).} \end{aligned}$$

From this, we see that a suitable control function G is $g^{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)]}$, and we have

$$\begin{aligned} g^{F(k)}(x) &\leqslant (g^{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)]})^k(x) \\ &= g^{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)] \cdot k}(x) \end{aligned}$$

A final calculation gives the result:

$$\begin{aligned} g^{F(Hig(\sigma, r+1, g^{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)]})(x))}(x) &\leqslant_{IH} g^{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)] \cdot (g^{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)]})_{\alpha_\sigma(r)}(x)}(x) \\ &= (g^{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)]})^{(g^{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)]})_{\alpha_\sigma(r)}(x)}(x) \\ &= (g^{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)]})^{\alpha_\sigma(r)}(x), \quad \text{by Eq. (4)} \\ &= g^{\Delta_{X \to \omega}[\alpha_\sigma(X \cdot r)] \cdot \alpha_\sigma(r)}(x). \end{aligned}$$

and the lemma follows. □

Lemma 7.4. *Define the function F as follows*:

$$F(0)=0,$$

$$F(k+1)-F(k)=Hig(\sigma+\omega^n\cdot g^{F(k)}(x),\ g^{F(k)}(x)\cdot r+1,\ g)(g^{F(k)}(x)).$$

Then

$$Hig(\sigma+\omega^{n+1},\ r+1,\ g)(x)\leqslant F(Hig(\sigma,\ r+1,\ G)(x)),$$

where G is a new control function, induced by F. From this we obtain

$$Hig(\sigma+\omega^{n+1},\ r+1,\ g)(x)\leqslant g_{\alpha_{\sigma+\omega^{n+1}}(r)}(x),$$

where $\alpha_{\sigma+\omega^{n+1}}(r):=\Delta_{X\to\omega}[\alpha_{\sigma+\omega^n\cdot X}(X\cdot r)]\cdot\alpha_\sigma(r)$.

Proof. The argument is similar to that of the previous lemma. We concentrate on the component ω^{n+1}. In the case where there is a principal controlled increasing subsequence the reasoning is the same as before. The difference here occurs in case the principal increasing subsequence fails to be controlled by (G,x), since we cannot deduce that the elements of the interval $[F(k),F(k+1)]$ are bounded in size. But the last element, e, of the controlled part of the principal increasing subsequence occurs at or before position $F(k)$ and so the elements of the principal sequence with indices in the interval $[F(k),F(k+1)]$ form a subset of $\{b\mid e\not\leqslant b\}$. These elements can therefore be decomposed as described in Section 6. The size of the interval has been chosen to solve the problem for the reduced space determined by *combining the spaces for the decomposed sequence and the passenger sequence.*

Now,

$$\begin{aligned}F(k+1)-F(k)&=Hig(\sigma+\omega^n\cdot g^{F(k)}(x),\ g^{F(k)}(x)\cdot r+1,g)(g^{F(k)}(x))\\&\leqslant_{IH} g_{\alpha_{\sigma+\omega^n\cdot g^{F(k)}(x)}(g^{F(k)}(x)\cdot r)}(g^{F(k)}(x))\end{aligned}$$

Thus, we have

$$\begin{aligned}g^{F(k+1)}(x)&\leqslant g^{g_{\alpha_{\sigma+\omega^n\cdot g^{F(k)}(x)}(g^{F(k)}(x)\cdot r)}(g^{F(k)}(x))+F(k)}(x)\\&=g^{g_{\alpha_{\sigma+\omega^n\cdot g^{F(k)}(x)}(g^{F(k)}(x)\cdot r)}(g^{F(k)}(x))}(g^{F(k)}(x))\\&=g^{\alpha_{\sigma+\omega^n\cdot g^{F(k)}(x)}(g^{F(k)}(x)\cdot r)}(g^{F(k)}(x))\quad\text{by Eq. (4)}\\&\leqslant g^{\Delta(r)}(g^{F(k)}(x)),\quad\text{where }\Delta(r)=\Delta_{X\to\omega}[\alpha_{\sigma+\omega^n\cdot X}(X\cdot r)].\end{aligned}$$

From this, we see that G is $g^{\Delta(r)}$, and we have

$$g^{F(k)}(x)\leqslant(g^{\Delta(r)})^k(x)=g^{\Delta(r)\cdot k}(x)$$

and

$$Hig(\sigma, r+1, g^{\Delta(r)})(x) \leqslant (g^{\Delta(r)})_{\alpha_\sigma(r)}(x).$$

Now

$$\begin{aligned} g^{F(Hig(\sigma,r+1,g^{\Delta(r)})(x))}(x) &\leqslant g^{\Delta(r)\cdot(g^{\Delta(r)})_{\alpha_\sigma(r)}(x)}(x) \\ &= (g^{\Delta(r)})^{(g^{\Delta(r)})_{\alpha_\sigma(r)}(x)}(x) \\ &= g^{\Delta(r)\cdot\alpha_\sigma(r)}(x) \quad \text{by Eq. (4).} \end{aligned}$$

Thus we read off the equation

$$\alpha_{\sigma+\omega^{n+1}}(r) := \Delta_{X\to\omega}[\alpha_{\sigma+\omega^n\cdot X}(X\cdot r)]\cdot\alpha_\sigma(r). \qquad \square$$

8. Summary

The arguments of the preceding sections summarise as follows. We have shown that $Hig(\sigma, r+1, g)(x) \leqslant g_{\alpha_\sigma(r)}(x)$ where

$$\alpha_\omega(r) = \omega\cdot r,$$

$$\alpha_{\sigma+\omega}(r) = \Delta_{X\to\omega}[\alpha_\sigma(X\cdot r)]\cdot\alpha_\sigma(r),$$

$$\alpha_{\omega^{n+1}}(r) = \Delta_{X\to\omega}[\alpha_{\omega^n\cdot X}(X\cdot r)]\cdot r,$$

$$\alpha_{\sigma+\omega^{n+1}}(r) = \Delta_{X\to\omega}[\alpha_{\sigma+\omega^n\cdot X}(X\cdot r)]\cdot\alpha_\sigma(r).$$

The above equations are independent of the particular choice of control function g, which therefore occurs as a parameter in our proof. It now remains to solve these equations.

The terms generated by the functions $\{\alpha_\sigma\}$ quickly become unmanageable. We therefore introduce a hierarchy $\{\beta_\sigma\}$ of functions whose values are easier to describe. This results in a slight coarsening of the bounds for the Higman function.

Note that the functions $\{\beta_\sigma\}$ could have been used in Lemmas 7.1, 7.3 and 7.4 in place of the functions $\{\alpha_\sigma\}$.

Definition 8.1.

$$\beta_0(\delta) = \delta$$

$$\beta_{\sigma+1}(\delta) = [\beta_\sigma(\omega\delta)]^2$$

$$\beta_\lambda(\delta) = (\Delta_{X\to\omega}[\beta_{\lambda_X}(X\delta)])^2.$$

Lemma 8.2. *Writing* $p_0(\omega)=1,\ p_{n+1}(\omega)=p_n(\omega)\cdot\omega+1$, *we have*:

$$\beta_{\sigma+\omega^n m}(\delta)=\beta_\sigma(\omega^{p_n(\omega)m}\delta)^{2^{p_n(\omega)m}}$$

Proof. The proof is by induction on m with a secondary induction on n:

$m=0$. The result follows trivially.

$m>0,\ n=0$. By a straightforward induction on $n<\omega$ one shows that $\beta_{\sigma+n}(\delta)=\beta_\sigma(\omega^n\delta)^{2^n}$.

$m+1,\ n+1$ We have

$$\begin{aligned}
\beta_{\sigma+\omega^{n+1}.(m+1)}(\delta) &= \beta_{\sigma+\omega^{n+1}.m+\omega^{n+1}}(\delta)\\
&= (\varDelta_{X\to\omega}[\beta_{\sigma+\omega^{n+1}.m+\omega^n.X}(X.\delta)])^2\\
&=_{IH(n)} (\varDelta_{X\to\omega}[\beta_{\sigma+\omega^{n+1}.m}(\omega^{p_n(\omega).X}.X.\delta)^{2^{p_n(\omega).X}}])^2\\
&= \beta_{\sigma+\omega^{n+1}.m}(\omega^{p_{n+1}(\omega)}.\delta)^{2^{p_{n+1}(\omega)}}\\
&=_{IH(m)} \beta_\sigma(\omega^{p_{n+1}(\omega).m}.\omega^{p_{n+1}(\omega)}.\delta)^{2^{p_{n+1}(\omega).m}.2^{p_{n+1}(\omega)}}\\
&= \beta_\sigma(\omega^{p_{n+1}(\omega).(m+1)}.\delta)^{2^{p_{n+1}(\omega).(m+1)}} \qquad \Box
\end{aligned}$$

Lemma 8.3. *For all* $\sigma<\omega^\omega$, $g_{\alpha_{\omega.\sigma}(\delta)}\leqslant g_{\beta_\sigma(\delta)}$.

Corollary 8.4. *A bound for Higman's theorem for* Σ_{n+1}^* *is obtained by taking* $\beta_{\omega^n}(1)$:

$$\beta_{\omega^n}(1)=(\omega^{p_n(\omega)})^{2^{p_n(\omega)}}\leqslant(\omega^{2^{p_n(\omega)}})^{2^{p_n(\omega)}}=\omega^{2^{p_n(\omega)+1}}.$$

The maximal order type of Higman's theorem on strings has been known for some time from [3]. For Σ_{n+1}^* this is ω^{ω^n}. This result leads to the intuition that an appropriate bound should be given by the function $g_{\omega^{\omega^n}}$, for some reasonable g.

Our approach has not generated results quite as precise as this. Our characterisations are sensitive to the combinatorial arguments used, thus it would be necessary to review these to obtain better ordinal bounds. Furthermore, it is not clear to us whether or not a significant improvement should be possible. Imposing the restriction that the Higman embedding ordering be contained in a total ordering on strings (as is the case in [3]) leads to a reduction of the set of bad sequences. So the problem is different and we are at present unable to say whether this is the reason for their lower complexity bound. Nevertheless, the result presented here is, modulo provability within fragments of arithmetic, comparable to that of De Jongh and Parikh [3].

References

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

[2] E.A. Cichon and H. Touzet, An ordinal calculus for proving termination in term rewriting, *Proc. Conf. on Trees in Algebra and Programming – CAAP 96*, Lecture Notes in Computer Science, Vol. 1059 (1996).

[3] D.H.J. De Jongh and R. Parikh, Well partial orderings and hierarchies, *Reprint from the Proc. Koninklijke Nederlandse Akademie van Wetenschappen, Amsterdam*, Ser. A, Vol. 80(3) (1977).

[4] E. Dennis-Jones and S.S. Wainer, Subrecursive hierarchies via direct limits, in: Richter, Borger, Oberschelp, Schinzel and Thomas, eds., *Logic Colloquium 83*, Lecture Notes in Mathematics, Vol. 1104 (Springer, Berlin, 1984) 117–128.

[5] M.V.H. Fairtlough and S.S. Wainer, Ordinal Complexity of Recursive Definitions, *Inform. and Comput.* **99**(2) (1992).

[6] J. Gallier, What's so special about Kruskal's theorem and the ordinal Γ_0. A survey of some results in proof theory, *Ann. Pure Appl. Logic* **53** (1991).

[7] J-Y. Girard, Π^1_2 Logic, Part 1, *Ann. Math. Logic* **21** (1981) 75–219.

[8] A. Grzegorczyk, Some classes of recursive functions, *Rozprawy Matematyczne* **4** (1953).

[9] P. Jullien, Analyse combinatoire: sur un théorème d'extension dans la théorie des mots, *C.R. Acad. Sci Paris, Ser.* A **266** (1968).

[10] L.A.S. Kirby and J. Paris, Accessible independence results for Peano arithmetic, *Bull. London Math. Soc.* **14** (1982).

[11] C.R. Murthy and J.R. Russell, A constructive proof of Higman's lemma, Proceedings of the fifth annual IEEE Symposium on Logic in Computer Science (1990) 257–267.

[12] R. Péter, *Recursive Functions* (Academic Press, New York, 1967).

[13] J.W. Robbin, Subrecursive hierarchies, PhD thesis, Princeton, 1965.

[14] K. Schütte and S.G. Simpson, Ein in der reinen Zahlentheorie unbeweisbarer Satz über endliche Folgen von natürlichen Zahlen, *Archiv für mathematische Logik und Grundlagenforschung* **25** (1985) 75–89.

[15] H. Schwichtenberg, Rekursionzahlen und die Grzegorczyk-Hierarchie, *Arch. Math. Logik* **12** (1969) 85–97.

[16] H. Schwichtenberg, Eine Klassifikation der ε_0-rekursiven Funktionen, *Zeit. Math. Logik* **17** (1971) 61–74.

[17] S.S. Wainer, A classification of the ordinal recursive functions, *Arch. Math. Logik* **13** (1970) 136–153.

[18] S.S. Wainer, Ordinal recursion and a refinement of the extended Grzegorczyk hierarchy, *J. Symbol. Logic* **37** (1972) 281–292.