Trends in Logic 53

Peter M. Schuster
Monika Seisenberger
Andreas Weiermann *Editors*

# Well-Quasi Orders in Computation, Logic, Language and Reasoning

## A Unifying Concept of Proof Theory, Automata Theory, Formal Languages and Descriptive Set Theory

Springer

**Trends in Logic**

Volume 53

# TRENDS IN LOGIC
*Studia Logica Library*

## VOLUME 53



The book series Trends in Logic covers essentially the same areas as the journal Studia Logica, that is, contemporary formal logic and its applications and relations to other disciplines. The series aims at publishing monographs and thematically coherent volumes dealing with important developments in logic and presenting significant contributions to logical research.

Volumes of Trends in Logic may range from highly focused studies to presentations that make a subject accessible to a broader scientific community or offer new perspectives for research. The series is open to contributions devoted to topics ranging from algebraic logic, model theory, proof theory, philosophical logic, non-classical logic, and logic in computer science to mathematical linguistics and formal epistemology. This thematic spectrum is also reflected in the editorial board of Trends in Logic. Volumes may be devoted to specific logical systems, particular methods and techniques, fundamental concepts, challenging open problems, different approaches to logical consequence, combinations of logics, classes of algebras or other structures, or interconnections between various logic-related domains. Authors interested in proposing a completed book or a manuscript in progress or in conception can contact either christi.lue@springer.com or one of the Editors of the Series.

More information about this series at http://www.springer.com/series/6645

Peter M. Schuster · Monika Seisenberger ·
Andreas Weiermann
Editors

# Well-Quasi Orders in Computation, Logic, Language and Reasoning

## A Unifying Concept of Proof Theory, Automata Theory, Formal Languages and Descriptive Set Theory

Springer

*Editors*
Peter M. Schuster
Dipartimento di Informatica
Università degli Studi di Verona
Verona, Italy

Monika Seisenberger
Department of Computer Science
Swansea University
Swansea, UK

Andreas Weiermann
Vakgroep Wiskunde
Ghent University
Ghent, Belgium

ISSN 1572-6126 ISSN 2212-7313 (electronic)
Trends in Logic
ISBN 978-3-030-30228-3 ISBN 978-3-030-30229-0 (eBook)
https://doi.org/10.1007/978-3-030-30229-0



This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

The theory of well-quasi orders, also known as wqos, is a highly active branch of combinatorics deeply rooted in and between many fields of mathematics and logic, among which there are proof theory, commutative algebra, braid groups, graph theory, analytic combinatorics, theory of relations, reverse mathematics and subrecursive hierarchies. As a unifying concept for slick finiteness or termination proofs, wqos have been rediscovered in diverse contexts, and turned out utmost useful in computer science.

With this volume we intend to display the many facets of and recent developments about wqos, through chapters written by scholars from different areas. Last but not least we thus wish to transfer knowledge between different areas of logic, mathematics and computer science.

A special highlight of the present volume is Diana Schmidt's habilitation thesis 'Well-partial ordering and the maximal order type' at the University of Heidelberg from 1979. Since publication this thesis has been extremely influential but never published, not even in parts.

This volume grew out of the following two meetings: the minisymposium 'Well-quasi orders: from theory to applications' organised by Peter Schuster, Monika Seisenberger and Andreas Weiermann within the 'Jahrestagung 2015 der Deutschen Mathematiker-Vereinigung (DMV)' from 21 to 25 September 2015 in Hamburg, and the Dagstuhl Seminar 16031 'Well Quasi-Orders in Computer Science' organised by Jean Goubault-Larrecq, Monika Seisenberger, Victor Selivanov and Andreas Weiermann from 17 to 22 January 2016 in Schloss Dagstuhl. The related financial support by the 'Deutsche Vereinigung für Mathematische Logik und für Grundlagenforschung der exakten Wissenschaften (DVMLG)' and by 'Schloss Dagstuhl: Leibniz Zentrum für Informatik' is gratefully acknowledged.

Parts of the editing process took place during Schuster's participation in the John Templeton Foundation's project 'A New Dawn of Intuitionism: Mathematical and Philosophical Advances' (ID: 60842). The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

The editors wish to thank the speakers and further participants of the aforementioned meetings, and the authors and referees of the chapters of the present volume.

Verona, Italy — Peter M. Schuster
Swansea, UK — Monika Seisenberger
Ghent, Belgium — Andreas Weiermann
June 2019

# Contents

# Editors and Contributors

## About the Editors

**Peter M. Schuster** is Associate Professor for Mathematical Logic at the University of Verona. After both doctorate and habilitation in mathematics at the University of Munich, he was Lecturer at the University of Leeds and member of the Leeds Logic Group. Apart from constructive mathematics at large, his principal research interests are about the computational content of classical proofs in abstract algebra and related fields in which maximum or minimum principles are invoked.

**Monika Seisenberger** is Associate Professor in Computer Science at Swansea University. After her doctorate in the Graduate Programme 'Logic in Computer Science' at the University of Munich she took up a position as Research Assistant at Swansea University, and subsequently as Lecturer and Programme Director. Her research focuses on Logic as well as on theorem proving and verification.

**Andreas Weiermann** is Full Professor for Mathematics at Ghent University. After doctorate and habilitation in mathematics at the University of Münster, he held postdoctoral positions in Münster and Utrecht and became first Associate Professor and somewhat later Full Professor in Ghent. His research interests are proof theory, theoretical computer science and discrete mathematics.

## Contributors

**Raphaël Carroy** University of Vienna, Wien, Austria

**Mirna Džamonja** University of East Anglia, Norwich, UK

**Lev Gordeev** Universität Tübingen, Ghent University, Ghent, Belgium

**Jean Goubault-Larrecq** LSV, ENS Paris-Saclay, Université Paris-Saclay, Cachan, France

**Simon Halfon** LSV, ENS Paris-Saclay, Université Paris-Saclay, Cachan, France

**Prateek Karandikar** LSV, ENS Paris-Saclay, Université Paris-Saclay, Cachan, France;
Chennai Mathematical Institute, Chennai, India;
CNRS UMI ReLaX, Chennai, India

**Julia F. Knight** Department of Mathematics, University of Notre Dame, Notre Dame, IN, USA

**Martin Krombholz** School of Mathematics, University of Leeds, Leeds, UK

**Karen Lange** Department of Mathematics, Wellesley College, Wellesley, MA, USA

**Chun-Hung Liu** Department of Mathematics, Texas A&M University, College Station, TX, USA;
Department of Mathematics, Princeton University, Princeton, NJ, USA

**Alberto Marcone** Dipartimento di Scienze Matematiche, Informatiche e Fisiche, Università di Udine, Udine, Italy

**K. Narayan Kumar** Chennai Mathematical Institute, Chennai, India;
CNRS UMI ReLaX, Chennai, India

**Yann Pequignot** University of Vienna, Wien, Austria

**Thomas Powell** Fachbereich Mathematik, Technische Universität Darmstadt, Darmstadt, Germany

**Michael Rathjen** School of Mathematics, University of Leeds, Leeds, UK

**Diana Schmidt** Emeritus, Hochschule Heilbronn, Heilbronn, Germany

**Sylvain Schmitz** IRIF, Université de Paris & CNRS, Paris, France

**Philippe Schnoebelen** LSV, CNRS, ENS Paris-Saclay, Université Paris-Saclay, Cachan, France;
CNRS UMI ReLaX, Chennai, India

**Victor Selivanov** A.P. Ershov Institute of Informatics Systems SB RAS and Kazan Federal University, Novosibirsk, Russia

**Silvia Steila** Institut für Informatik, Universität Bern, Bern, Switzerland

**Christian Sternagel** University of Innsbruck, Innsbruck, Austria

# Well, Better and In-Between

**Raphaël Carroy and Yann Pequignot**

**Abstract** Starting from well-quasi-orders (WQOs), we motivate step by step the introduction of the complicated notion of better-quasi-order (BQO). We then discuss the equivalence between the two main approaches to defining BQO and state several essential results of BQO theory. After recalling the rôle played by the ideals of a WQO in its BQOness, we give a new presentation of known examples of WQOs which fail to be BQO. We also provide new forbidden pattern conditions ensuring that a quasi-order is a better quasi-order.

> It is the variety of these applications, rather than any depth in the results obtained, that suggests that the theorems may be interesting.
>
> Graham Higman [13]

While studying a generalization of the partial order of divisibility on the natural numbers for an abstract algebra, Higman [13] identified the following desirable property for a quasi-order (qo). A qo has the *finite basis property* if every upwards closed subset is the upward closure of a finite subset. He notices that this property is equivalent to that defining a well-quasi-order (WQO): being well-founded and having no infinite antichains. Higman proves the following essential fact: in order to be WQO it suffices to be generated by means of *finitary* operations from a WQO. He then proceeds to apply his theorem to solve a problem posed by Erdős, to provide a new proof of a theorem on power-series ring and also to the study of fully invariant subgroups of a free group. These were only the first instances of a long series of applications of this result that became known as Higman's Theorem. Pouzet [28] later commented on the possibilities and the limitations of that fruitful approach:

Raphaël Carroy was supported by FWF Grants $P28153$ and $P29999$. Yann Pequignot gratefully acknowledges the support of the Swiss National Science Foundation (SNF) through grant $P2LAP2\_164904$.

R. Carroy (✉) · Y. Pequignot
University of Vienna, Oskar-Morgenstern-Platz 1,1090 Wien, Austria
e-mail: raphael.carroy@univie.ac.at



P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_1

> In order to show that a certain class of posets (finite or infinite) is WQO, one tries first to see if the class can be constructed from some simpler class by means of some operations. If these operations are finitary, then it is possible that Higman's theorem can be applied. However, for infinite posets, these operations may very well also be infinitary and then there is no possibility of applying Higman's theorem since the obvious generalization of this is false for infinitary operations.

Pouzet here refers to the fact that well-founded quasi-orders, as well as WQOs, lack closure under certain *infinitary* operations as first proven by Rado [31]. This is explained in detail in Sect. 1 where we point out how it opens the way to the definition by Alvah Nash-Williams [1] of the concept of better-quasi-order (BQO): a stronger property than WQO which allows for an infinitary analogue to Higman's Theorem. We first provide a gentle introduction to the original definition of Nash-Williams, before presenting the more concise definition introduced by Simpson [33]. We use some new terminology with the hope that it makes it easier for the unacquainted reader to appreciate the respective advantages of these two complementary approaches to defining BQO.

To offer a few more words of introduction about this intriguing concept, we briefly comment on the emblematic case of the quasi-order $\mathrm{LIN}_{\aleph_0}$ of countable linear orders, equipped with the relation of embeddability. Fraïssé [10] conjectured that $\mathrm{LIN}_{\aleph_0}$ was well-founded, but the statement that became known as Fraïssé's Conjecture (FRA) is that $\mathrm{LIN}_{\aleph_0}$ is WQO; it follows from the famous theorem of Laver [16] that $\mathrm{LIN}_{\aleph_0}$ is in fact BQO. The reason for the use of the concept of BQO in Laver's proof of FRA was already alluded to in the above quote by Pouzet. While using Hausdorff's analysis of scattered linear orders and proceeding by induction is a very reasonable way to tackle FRA, the operations underlying this analysis are infinitary and this is a main obstacle when working with WQOs alone. This masterly use by Laver of the concept of BQO introduced by Nash-Williams inspired many other delightful results. But however successful this story is, it raises at least two questions.

Firstly, one may ask if the use of the concept of BQO in the proof of FRA is in a sense necessary. In the framework of Reverse Mathematics, one can formalize this question by asking for the exact proof-theoretic strength of FRA. The answer is still unknown despite many efforts, but important results have already been obtained (see [21] and more recently [23]).

Secondly, one may ask if this strategy for proving WQOness always works. On the one hand, many other quasi-orders were proved to be BQO in the subsequent years attesting to the effectiveness of this concept (see Sect. 3.3). On the other hand, there does exist a large range of examples of WQO that are not BQO (see Sect. 4). Nevertheless, these examples appear to have a somehow artificial flavor since as Kruskal [14, p. 302] observed in his very nice historical introduction to WQO: "all 'naturally occurring' WQO sets which are known are BQO".[1]

In a quest towards a deeper understanding of the discrepancy between WQO and BQO, we mention a result of ours on the role played by the ideals of a WQO in it being

[1]The minor relations on finite graphs, proved to be WQO by Robertson and Seymour [32], is to our knowledge the only naturally occurring WQO which is not *yet* known to be BQO.

BQO (see Theorem 3.23). We show the relevance of this rather singular theorem by giving two applications of it. First, we use it in Sect. 4 to give a new presentation of examples of WQOs that, while failing to be BQOs, still enjoy stronger and stronger properties. Finally, we use this theorem to give in Sect. 5 some new conditions under which the two notions of WQO and BQO coincide.

## 1 Well Is Not Good Enough

In the sequel, $(Q, \leq_Q)$ always stands for a quasi-order, qo for short, i.e. a reflexive and transitive relation $\leq_Q$ on a non-empty set $Q$. An antisymmetric qo is a *partial order*, or po. A sequence $(q_n)_{n\in\omega}$ in $Q$ is *bad* if and only if for all integers $m$ and $n$ such that $m < n$ we have $q_m \not\leq_Q q_n$.

The strict quasi-order associated to $\leq_Q$ is defined by $p <_Q q$ if and only if $p \leq_Q q$ and $q \not\leq_Q p$. We say that $Q$ is *well-founded* if there is no *infinite descending chain* in $Q$, i.e. no sequence $(q_n)_n$ such that $q_{n+1} <_Q q_n$ for every $n$. An *antichain* in $Q$ is a subset $A$ of $Q$ consisting of pairwise $\leq_Q$-incomparable elements, i.e. $p \neq q$ implies $p \not\leq_Q q$ for every $p, q \in A$.

A subset $D$ of $Q$ is a called a *downset*, if $q \in D$ and $p \leq_Q q$ implies $p \in D$. For any $S \subseteq Q$, we write $\downarrow S$ for the downset generated by $S$ in $Q$, i.e. the set $\{q \in Q \mid \exists p \in S\ q \leq_Q p\}$. We also write $\downarrow p$ for $\downarrow\{p\}$. Finally we denote by $\mathcal{D}(Q)$ the po of downsets of $Q$ under inclusion.

We start by proving the equivalence between three of the main characterizations of a WQO.

**Proposition 1.1** *A quasi-order $(Q, \leq_Q)$ is a* WQO *if and only if one of the following equivalent conditions is fulfilled:*

1. *there is no bad sequence in $(Q, \leq_Q)$,*
2. *$(Q, \leq_Q)$ is well-founded and contains no infinite antichain,*
3. *$(\mathcal{D}(Q), \subseteq)$ is well-founded.*

*Proof* **Item 2 ↔ Item 1** Notice that an infinite descending chain and a countably infinite antichain are both special cases of a bad sequence. Conversely if $(q_n)_n$ is a bad sequence in $Q$, then using Ramsey's theorem we obtain either an infinite descending chain or an infinite antichain.

**Item 1 → Item 3** By contraposition, suppose that $(D_n)_{n\in\omega}$ is an infinite descending chain inside $(\mathcal{D}(Q), \subseteq)$. Then for each $n \in \omega$ we can pick some $q_n \in D_n \setminus D_{n+1}$. Then $n \mapsto q_n$ is a bad sequence in $Q$. To see this, suppose towards a contradiction that for $m < n$ we have $q_m \leq q_n$. As $q_n \in D_n$ and $D_n$ is a downset, we have $q_m \in D_n$. But since $D_n \subset D_{m+1}$, we have $q_m \in D_{m+1}$, a contradiction with the choice of $q_m$.

**Item 3 → Item 1** By contraposition, suppose that $(q_n)_{n\in\omega}$ is a bad sequence in $Q$. Set $D_n = \downarrow\{q_k \mid n \leq k\}$, then $(D_n)_n$ is a descending chain in $\mathcal{D}(Q)$. Indeed for every $n$ we clearly have $D_{n+1} \subseteq D_n$ and since the sequence is bad, $k \geq n+1$ implies $q_n \not\leq q_k$ and so $q_n \in D_n$ while $q_n \notin D_{n+1}$. □

After Proposition 1.1, it is natural to ask if being well-founded and being WQO is actually equivalent for the partial order of downsets of any quasi-order. The answer is negative and the first example of a WQO with an antichain of downsets was identified by Richard Rado.

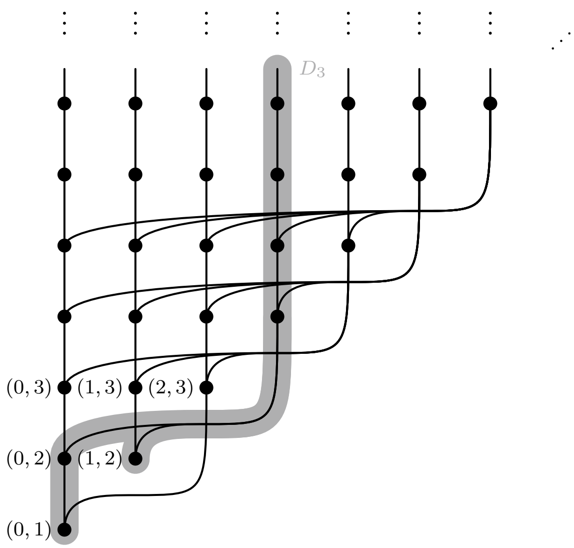
*Example 1.2* ([31]) Rado's partial order $\mathfrak{R}$ is the set $[\omega]^2$ of pairs of natural numbers, partially ordered by (cf. Fig. 1):

$$\{m, n\} \leq_{\mathfrak{R}} \{m', n'\} \quad \longleftrightarrow \quad \begin{cases} m = m' \text{ and } n \leq n', \text{ or} \\ n < m'. \end{cases}$$

where by convention a pair $\{m, n\}$ of natural numbers is always assumed to be written in increasing order ($m < n$).

The po $\mathfrak{R}$ is WQO. To see this, consider any map $f : \omega \to [\omega]^2$ and let $f(n) = \{f_0(n), f_1(n)\}$ for all $n \in \omega$. Now if $f_0$ is unbounded, then there exists $n > 0$ with $f_1(0) < f_0(n)$ and so $f(0) \leq_{\mathfrak{R}} f(n)$ by the second clause. So $f$ is good in this case. Next if $f_0$ is bounded, then by going to a subsequence we can assume that $f_0$ is constantly equal to some $k$. But then the restriction of $\mathfrak{R}$ to the pairs $\{k, n_1\}$ is simply $\omega$ which is WQO, so $f$ must be good in this case too.

**Fig. 1** Rado's poset $\mathfrak{R}$



However the map $n \mapsto D_n = \downarrow\{\{n,l\} \mid n < l\}$ is a bad sequence (in fact an infinite antichain) inside $\mathcal{D}(\mathfrak{R})$. Indeed whenever $m < n$ we have $\{m,n\} \in D_m$ while $\{m,n\} \notin D_n$, and so $D_m \nsubseteq D_n$.

Suppose we want to make sure that $\mathcal{D}(Q)$ is WQO. What condition on $Q$ could ensure this? In other words, what phenomenon are we to exclude inside $Q$ in order to rule out the existence of antichains inside $\mathcal{D}(Q)$? Forbidding bad sequences in $Q$ is certainly not enough, as shown by the existence of Rado's example. But here is what we can do.

Suppose that $(P_n)_{n\in\omega}$ is a bad sequence in $\mathcal{D}(Q)$. Fix some $m \in \omega$. Then whenever $m < n$ we have $P_m \nsubseteq P_n$ and we can choose a witness $q \in P_m \setminus P_n$. In general, no single $q \in P_m$ can witness that $P_m \nsubseteq P_n$ for all $n > m$, so we have to pick a whole sequence $f_m : \omega/m \to Q, n \mapsto q_m^n$ of witnesses[2]:

$$q_m^n \in P_m \quad \text{and} \quad q_m^n \notin P_n, \quad n \in \omega/m.$$

In this way we get a sequence $f_0, f_1, \ldots$ of sequences which is advantageously viewed as single map from $[\omega]^2$:

$$\begin{aligned} f : [\omega]^2 &\longrightarrow Q \\ \{m,n\} &\longmapsto f_m(n) = q_m^n. \end{aligned}$$

By our choices this sequence of sequences satisfies the following condition:

$$\forall m,n,l \in \omega \quad m < n < l \to q_m^n \nleq q_n^l.$$

To see this, suppose towards a contradiction that for $m < n < l$ we have $q_m^n \leq q_n^l$. Since $q_n^l \in P_n$ which is a downset, we would have $q_m^n \in P_n$, but we chose $q_m^n$ such that $q_m^n \notin P_n$.

Let us say that a *sequence of sequences* $f : [\omega]^2 \to Q$ is *bad* if for every $m,n,l \in \omega$, $m < n < l$ implies $f(\{m,n\}) \nleq f(\{n,l\})$. We have found the desired condition.

**Proposition 1.3** *Let $Q$ be a qo. Then $\mathcal{D}(Q)$ is* WQO *if and only if there is no bad sequence of sequences in $Q$.*

*Proof* As we have seen in the preceding discussion, if $\mathcal{D}(Q)$ is not WQO then from a bad sequence in $\mathcal{D}(Q)$ we can define a bad sequence of sequences in $Q$.

Conversely, if $f : [\omega]^2 \to Q$ is a bad sequence of sequences, then for each $m \in \omega$ we can consider the set $P_m = \{f(\{m,n\}) \mid n \in \omega/m\}$ consisting of the image of the $m^{\text{th}}$ sequence. Then the sequence $m \mapsto \downarrow P_m$ in $\mathcal{D}(Q)$ is a bad sequence. Indeed every time $m < n$ we have $f(\{m,n\}) \in P_m$ while $f(\{m,n\}) \notin \downarrow P_n$, since otherwise there would exist $l > n$ with $f(\{m,n\}) \leq f(\{n,l\})$, a contradiction with the fact that $f$ is a bad sequence of sequences. □

---

[2]where $\omega/m$ denotes the set $\{n \in \omega \mid m < n\}$.

Notice that in the case of Rado's partial order $\mathfrak{R}$, the identity map itself is a bad sequence of sequences witnessing that $\mathcal{D}(\mathfrak{R})$ is not WQO, since every time $m < n < l$ then $\{m, n\} \not\leq_{\mathfrak{R}} \{n, l\}$. This example is actually minimal in the following sense; if $Q$ is WQO but $\mathcal{D}(Q)$ is not WQO, then $\mathfrak{R}$ embeds into $Q$, as proved by Laver [17].

For now, let us just say that a *better-quasi-order* is a quasi-order $Q$ such that $\mathcal{D}(Q)$ is well-founded, $\mathcal{D}(\mathcal{D}(Q))$ is well founded, $\mathcal{D}(\mathcal{D}(\mathcal{D}(Q)))$ is well-founded, so on and so forth into the transfinite. While this idea can be formalized, we can already see that it cannot serve as a convenient definition.[3]

In the next Section, we introduce the super-sequences and the multi-sequences which are two equivalent way of generalizing the idea of sequence of sequences into the transfinite. This allows us to define better-quasi-orders in Sect. 3.

## 2 Super-Sequences Versus Multi-sequences

As the preceding section suggests, we are going to define a better-quasi-order as a quasi-order with no bad sequence, with no bad sequence of sequences, no bad sequence of sequences of sequences, so on and so forth, into the transfinite. In order to formalize this idea, we need a convenient notion of "index set" for a sequence of sequences of … of sequences, in short a *super-sequence*. We first describe the original combinatorial approach of Nash-Williams, before presenting the more condensed topological definition due to Simpson.

Let us first fix some notations. We adopt the set-theoretic convention that $m \in n$ for all natural numbers $m < n$. A *sequence* is a map from an initial segment of $\omega$ to some non-empty set, a *finite* sequence $s$ has domain an integer $n$, also called the *length* of $s$ and denoted by $|s|$. When $i < |s|$, $s(i)$ stands for the $i$-th element of the sequence $s$. If $A$ is a non-empty set, $A^{<\omega}$ stands for the set of finite sequences in $A$ and $A^\omega$ stands for the set of infinite sequences in $A$.

We write $[A]^{<\infty}$ for the set of finite subsets of $A$ and $[A]^\infty$ for the set of infinite subsets of $A$. We identify any subset of the natural numbers with its increasing enumeration.

In this way, we say that $t \subseteq \omega$ extends $s \subseteq \omega$, in symbol $s \sqsubseteq t$, exactly when this happens for the corresponding increasing enumerations.

**Definition 2.1** A family $F \subseteq [\omega]^{<\infty}$ is a *front on* $X \in [\omega]^\infty$ if

1. either $F = \{\emptyset\}$, or $\bigcup F = X$,
2. for all $s, t \in F$ $s \sqsubseteq t$ implies $s = t$,
3. (Density) for all $X' \in [X]^\infty$ there is an $s \in F$ such that $s \sqsubset X'$.

A *super-sequence* in a set $Q$ is a map from any front $F$ on $X \in [\omega]^\infty$ to $Q$.

[3] The reader who remains unconvinced can try to prove that the partial order $(3, =)$ satisfies this property.

Notice that, according to our definition, the trivial front $\{\emptyset\}$ is a front on $X$ for every $X \in [\omega]^\infty$. Except for this degenerate example, if a family $F \subseteq [X]^{<\infty}$ is a front on $X$, then necessarily $X$ is equal to $\bigcup F$, the set-theoretic union of the family $F$. For this reason we will sometimes say that $F$ is a front, without reference to any infinite subset $X$ of $\omega$. Moreover when $F$ is not trivial, we refer to the unique $X$ for which $F$ is a front on $X$, namely $\bigcup F$, as the *base* of $F$.

*Example 2.2* For all natural number $n$ the set $[\omega]^{n+1}$ is a non-trivial front. The family $\mathcal{S} := \{s \in [\omega]^{<\infty} : |s| = s(0) + 1\}$ is also a front, it is traditionally called the *Schreier barrier*.

A sequence of sequences is a super-sequence with domain $[\omega]^2$, a sequence of sequences of sequences is a super-sequence with domain $[\omega]^3$, and so on. A super-sequence with domain $\mathcal{S}$ is an example of a transfinite super-sequence.

A front can be profitably decomposed in a sequence of "simpler" fronts.

**Fact 2.3** *Let $F$ be a non trivial front on $X \in [\omega]^\infty$. For all $n \in X$ the set $F_n := \{s \in [\omega/n]^{<\infty} \mid \{n\} \cup s \in F\}$ is a front on $X/n$ called the* ray *of $F$ at $n$. Moreover $F = \bigcup_{n \in X} \{\{n\} \cup s \mid s \in F_n\}$.*

*Proof* Let $n \in X$. For every $Y \in [X/n]^\infty$ there exists $s \in F$ with $s \sqsubset \{n\} \cup Y$. Since $F$ is non trivial, $s \neq \emptyset$ and so $n \in s$. Therefore $s' = s \setminus \{n\} \in F_n$ with $s' \sqsubset Y$, and $F_n$ satisfies Item 3. Now if $F_n$ is not trivial and $k \in X/n$, there is $s \in F_n$ with $s \sqsubset \{k\} \cup X/k$ and necessarily $k \in s \subseteq \bigcup F_n$. Hence $\bigcup F_n = X/n$, so Item 1 is met. To see Item 2, let $s, t \in F_n$ with $s \sqsubseteq t$. Then for $s' = \{n\} \cup s$ and $t' = \{n\} \cup t$ we have $s', t' \in F$ and $s' \sqsubseteq t'$, so $s' = t'$ and $s = t$, as desired. The last statement is obvious. □

Generalizing Ramsey's theorem, Nash-Williams proved that fronts enjoy a fundamental property: any time we partition a front into finitely many pieces, at least one of the pieces contains a front. We are now going to introduce the necessary tools to prove this result: first, an ordinal rank on fronts that allows for inductive proofs, then a characterization of what a sub-front looks like.

## 2.1 The Ordinal Rank

We define a rank on fronts by associating to every given front a well-founded tree. We first need some classical notations and definitions about trees.

**Definitions 2.4** 1. A *tree* $T$ on a set $A$ is a subset of $A^{<\omega}$ that is closed under prefixes, i.e. $u \sqsubseteq v$ and $v \in T$ implies $u \in T$.
2. A tree $T$ on $A$ is called *well-founded* if $T$ has no infinite branch, i.e. if there is no infinite sequence $x \in A^\omega$ such that $x\restriction_n \in T$ holds for all $n \in \omega$. In other words, a tree $T$ is well-founded if $(T, \sqsupseteq)$ is a well-founded partial order.

3. When $T$ is a non-empty well-founded tree we can define a strictly decreasing function $\rho_T$ from $T$ to the ordinals by transfinite recursion:

$$\rho_T(t) = \sup\{\rho_T(s) + 1 \mid t \sqsubset s \in T\} \quad \text{for all } t \in T.$$

It is easily shown to be equivalent to

$$\rho_T(t) = \sup\{\rho_T(t \frown (a)) + 1 \mid a \in A \text{ and } t \frown (a) \in T\} \quad \text{for all } t \in T.$$

The *rank* of the non-empty well-founded tree $T$ is the ordinal $\rho_T(\emptyset)$.

Through the identification of a set of natural numbers with its increasing enumeration we can consider the tree generated by a front. For any front $F$, we let $T(F)$ be the smallest tree on $\omega$ containing $F$, i.e.

$$T(F) = \{s \in \omega^{<\omega} \mid \exists t \in F \; s \sqsubseteq t\}.$$

For convenience, we henceforth write $X/s$ to denote the set $\{k \in X \mid \max s < k\}$.

**Lemma 2.5** *For every front $F$, the tree $T(F)$ is well-founded.*

*Proof* If $x$ is an infinite branch of $T(F)$, then $x$ enumerates an infinite subset $X$ of $\bigcup F$ such that for every $u \sqsubset X$ there exists $t \in F$ with $u \sqsubseteq t$. Since $F$ is a front there exists a (unique) $s \in F$ with $s \sqsubset X$. But for $n = \min X/s$ and $u = s \cup \{n\}$, there is $t \in F$ with $u \sqsubseteq t$. But then $F \ni s \sqsubset u \sqsubseteq t \in F$ contradicting Item 2 in the definition of a front. □

**Definition 2.6** Let $F$ be a front. The *rank* of $F$, denoted by rk $F$, is the rank of the tree $T(F)$.

*Example 2.7* Notice that the family $\{\emptyset\}$ is the only front of null rank, and for all positive integer $n$, the front $[\omega]^n$ has rank $n$. Moreover the Schreier barrier $\mathcal{S}$ has rank $\omega$.

We now observe that the rank of $F$ is closely related to the rank of its rays $F_n$, $n \in X$. Given $F$ a non trivial front on $X \in [\omega]^\infty$, notice that the tree $T(F_n)$ of the front $F_n$ is naturally isomorphic to the subset

$$\{s \in T(F) \mid \{n\} \sqsubseteq s\}$$

of $T(F)$. The rank of the front $F$ is therefore related to the ranks of its rays through the following formula:

$$\operatorname{rk} F = \sup\{\operatorname{rk}(F_n) + 1 \mid n \in X\}.$$

In particular, $\operatorname{rk} F_n < \operatorname{rk} F$ for all $n \in X$.

This simple remark allows one to prove results on fronts by induction on the rank by applying the induction hypothesis to the rays, as it was first done in [29].

## 2.2 Sub-fronts

By analogy with classical sequences let us make the following definition.

**Definition 2.8** A sub-super-sequence of a super-sequence $f : F \to E$ is a restriction $f{\restriction_G} : G \to E$ to some front $G$ included in $F$.

The following important operation is quite useful when dealing with *sub-fronts* of a given front, i.e. sub-families of a front which are themselves fronts. For a family $F \subseteq \mathcal{P}(\omega)$ and some $X \in [\omega]^\infty$, we define the sub-family

$$F|X := \{s \in F \mid s \subseteq X\}.$$

**Proposition 2.9** *Let $F$ be a front on $X$. Then a family $F' \subseteq F$ is a front if and only if there exists $Y \in [X]^\infty$ such that $F|Y = F'$.*

*Proof* The claim is obvious if $F$ is trivial so suppose $F$ is non-trivial.

- $\to$ Let $F' \subseteq F$ be a front on $Y$. Since $F'$ is not trivial either, $Y = \bigcup F' \subseteq \bigcup F = X$. Now if $s \in F'$ then clearly $s \in F|Y$. Conversely if $s \in F|Y$ then there exists a unique $t \in F'$ with $t \sqsubset s \cup Y/s$ and so either $s \sqsubseteq t$ or $t \sqsubseteq s$. Since $F$ is a front and $s, t \in F$, necessarily $s = t$ and so $s \in F'$. Therefore $F' = F|Y$.
- $\leftarrow$ If $Y \in [X]^\infty$ then the family $F|Y$ is a front on $Y$. Clearly $F|Y$ satisfies Item 2. If $Z \in [Y]^\infty$ then since $Y \subseteq X$, then $Z \in [X]^\infty$ and so there exists $s \in F$ with $s \sqsubset Z$. But then $s \subseteq Z \subseteq Y$, so in fact $s \in F|Y$ and therefore $F|Y$ satisfies Item 3. For Item 1, notice that $\bigcup F|Y \subseteq Y$ by definition and that if $n \in Y$, then as we have already seen there exists $s \in F|Y$ with $s \sqsubset \{n\} \cup Y/n$, so $n \in s$ and $n \in \bigcup F|Y$. □

Observe that the operation of restriction commutes with the taking of rays.

**Fact 2.10** *Let $F \subseteq \mathcal{P}(\omega)$ and $X \in [\omega]^\infty$. For every $n \in X$ we have*

$$F_n|X = (F|X)_n.$$

Notice also the following simple important fact. If $F'$ is a sub-front of a front $F$, then the tree $T(F')$ is included in the tree $T(F)$ and so $\operatorname{rk} F' \leq \operatorname{rk} F$.

## 2.3 A Ramsey Property for Fronts: Nash-Williams's Theorem

We now prove this theorem to give a simple example of a proof by induction on the rank of a front, an extremely fruitful technique.

**Theorem 2.11** (Nash-Williams) *Let $F$ be a front. For any subset $S$ of $F$ there exists a front $F' \subseteq F$ such that either $F' \subseteq S$ or $F' \cap S = \emptyset$.*

*Proof* The claim is obvious for the trivial front whose only subsets are the empty set and the whole trivial front. So suppose that the claim holds for every front of rank smaller than $\alpha$. Let $F$ be a front on $X$ with rk $F = \alpha$ and $S \subseteq F$. For every $n \in X$ let $S_n$ be the subset of the ray $F_n$ given by $S_n = \{s \in F_n \mid \{n\} \cup s \in S\}$.

Set $X_{-1} = X$ and $n_0 = \min X_{-1}$. Since rk $F_{n_0} < \alpha$ there exists by induction hypothesis some $X_0 \in [X_{-1}/n_0]^\infty$ such that

$$\text{either } F_{n_0}|X_0 \subseteq S_{n_0}, \quad \text{or } F_{n_0}|X_0 \cap S_{n_0} = \emptyset.$$

Set $n_1 = \min X_0$. Now applying the induction hypothesis to $F_{n_1}|(X_0/n_0)$ and $S_{n_1}$ we get an $X_1 \in [X_0/n_0]^\infty$ such that either $F_{n_1}|X_1 \subseteq S_{n_1}$, or $F_{n_1}|X_1 \cap S_{n_1} = \emptyset$. Continuing in this fashion, we obtain a sequence $X_k$ together with $n_k = \min X_{k-1}$ such that for all $k$ we have $X_k \in [X_{k-1}/n_k]^\infty$ and

$$\text{either } F_{n_k}|X_k \subseteq S_{n_k}, \quad \text{or } F_{n_k}|X_k \cap S_{n_k} = \emptyset.$$

Now there exists $Y \in [\omega]^\infty$ such that either $F_{n_k}|X_k \subseteq S_{n_k}$ for all $k \in Y$, or $F_{n_k}|X_k \cap S_{n_k} = \emptyset$ for all $k \in Y$. Let $X = \{n_k \mid k \in Y\}$ then $F|X$ is as desired. Indeed for all $s \in F|X$ we have $\min s = n_k$ for some $k \in Y$ and $s \setminus \{n_k\} \in F_{n_k}|X_k$. Hence by the choice of $Y$, either $s \setminus \{\min s\} \in S_{\min s}$ for all $s \in F|X$, or $s \setminus \{\min s\} \notin S_{\min s}$ for all $s \in F|X$. Therefore either $F|X \subseteq S$ or $F|X \cap S = \emptyset$. □

Nash-Williams' Theorem 2.11 is easily seen to be equivalent to the following statement.

**Theorem 2.12** *Let $E$ be a finite set. Then every super-sequence $f : F \to E$ admits a constant sub-super-sequence.*

The above result obviously does not hold in general for an infinite set $E$ (consider for example any injective super-sequence). However Pudlák and Rödl [29] proved an interesting theorem in this context. In a different direction the authors also obtained in [4] that when $E$ is a compact metric space, then every super-sequence $f : F \to E$ admits a sub-super-sequence which is a so-called *Cauchy* super-sequence.

## *2.4 Continuous Definition: Multi-sequences*

We now present another fruitful approach to the definition of better-quasi-orders, initiated by [33], and we relate it to super-sequences.

Let $E$ be any set, and $f : F \to E$ be a super-sequence with $F$ a front on $X$. For every $Y \in [X]^\infty$ there exists a unique $s \in F$ with $s \sqsubset Y$. We can therefore define a map $f^\uparrow : [X]^\infty \to E$ defined by $f^\uparrow(Y) = f(s)$ where $s$ is the unique member of $F$ with $s \sqsubset Y$.

**Definition 2.13** A *multi-sequence* in some set $E$ is a map $h : [X]^\infty \to E$ for some $X \in [\omega]^\infty$. A *sub-multi-sequence* of $h : [X]^\infty \to E$ is a restriction of $h$ to $[Y]^\infty$ for some $Y \in [X]^\infty$.

For every $X \in [\omega]^\infty$ we endow $[X]^\infty$ with the topology induced by the Cantor space, identifying once again subsets of the natural numbers with their characteristic functions. As a topological space $[X]^\infty$ is homeomorphic to the Baire space $\omega^\omega$. This homeomorphism is conveniently realized via the embedding of $[X]^\infty$ into $\omega^\omega$ which maps each $Y \in [X]^\infty$ to its injective and increasing enumeration $e_Y : \omega \to Y$. We henceforth identify the space $[X]^\infty$ with the closed subset of $\omega^\omega$ of injective and increasing sequences in $X$. We thus obtain a countable basis of clopen sets for $[X]^\infty$:

$$M_s = N_s \cap [X]^\infty = \{Y \in [X]^\infty \mid s \sqsubset Y\}, \quad \text{for } s \in [X]^{<\infty}.$$

**Definition 2.14** A multi-sequence $h : [X]^\infty \to E$ is *locally constant* if for all $Y \in [X]^\infty$ there exists $s \in [X]^{<\infty}$ such that $Y \in M_s$ and $h$ is constant on $M_s$, i.e. for every $Y \in [X]^\infty$ there exists $s \sqsubset Y$ such that for every $Z \in [X]^\infty$, $s \sqsubset Z$ implies $h(Z) = h(Y)$.

Clearly for every super-sequence $f : F \to E$ where $F$ is a front on $X$ the map $f^\uparrow : [X]^\infty \to E$ is locally constant.

Conversely for any locally constant multi-sequence $h : [X]^\infty \to E$, let

$$S^h = \{s \in [X]^{<\infty} \mid h \text{ is constant on } M_s\}.$$

**Lemma 2.15** *The set $F^h$ of $\sqsubseteq$-minimal elements of $S^h$ is a front on $X$.*

*Proof* By $\sqsubseteq$-minimality if $s, t \in F^h$ and $s \sqsubseteq t$, then $s = t$. For every $Y \in [X]^\infty$, since $h$ is locally constant there exists $s \sqsubset Y$ such that $h$ is constant on $M_s$. Hence there exists $t \in F^h$ with $t \sqsubseteq s$, and so $t \sqsubset Y$ too. To see that either $F^h$ is trivial or $\bigcup F^h = X$, notice that $h$ is constant if and only if $F^h$ is the trivial front if and only if $\emptyset \in F^h$. So if $F^h$ is not trivial, then for every $n \in X$ there exists $s \in F^h$ with $s \sqsubset \{n\} \cup X/n$ and since $s \neq \emptyset$, we get $n \in s$ and $n \in \bigcup F^h$. □

We can therefore associate to every locally constant multi-sequence $h : [X]^\infty \to E$ a super-sequence $h^\downarrow : F^h \to E$ by letting, in the obvious way, $h^\downarrow(s)$ be equal to the unique value taken by $h$ on $M_s$ for every $s \in F^h$.

**Corollary 2.16** *Let $E$ be a finite set. Then every locally constant multi-sequence $f : [X]^\infty \to E$ admits a constant sub-multi-sequence.*

# 3 Well, Here Is Better!

As promised, we now give the two main definitions of BQO available in the literature. Proceeding in unchronological order, we start by the one due to Simpson which makes use of multi-sequences before stating the original one due to Nash-Williams based on super-sequences. In both definitions, we only miss one last ingredient: a suitable generalization of the usual order on the natural numbers.

## 3.1 Two Equivalent Definitions

For every $N \in [\omega]^\infty$ we call the *shift of* $N$, denoted by ${}_*N$, the set $N \setminus \{\min N\}$. Notice that $N \mapsto {}_*N$ is a continuous map from $[\omega]^\infty$ to itself.

**Definition 3.1** Let $Q$ be a qo and $h : [X]^\infty \to Q$ a multi-sequence.

1. We say that $h$ is *bad* if $h(N) \not\leq h({}_*N)$ for every $N \in [X]^\infty$,
2. We say that $h$ is *good* if there exists $N \in [X]^\infty$ with $h(N) \leq h({}_*N)$,

At last, we present the deep definition due to Nash-Williams here in its modern "Simpsonian" reformulation.

**Definition 3.2** A quasi-order $Q$ is a *better-quasi-order* (BQO) if there is no bad locally constant multi-sequence in $Q$.

*Remark 3.3* The reader familiar with Descriptive Set Theory may suspect that this is not the most general definition. Simpson indeed considers *Borel multi-sequences*, namely multi-sequences whose range is countable and such that the preimage of any singleton is Borel. By the Galvin-Prikry Theorem [12], which is the Borel generalization of Theorem 2.11, any such multi-sequence admits a locally constant sub-multi-sequence. One can therefore safely replace "locally constant" by "Borel" in the above definition. While this result is very convenient in certain constructions and essential to some proofs, we shall not use it in this article.

Of course the definition of better-quasi-order can be formulated in terms of super-sequences as Nash-Williams originally did. The only missing ingredient is a counterpart of the shift map $N \mapsto {}_*N$ on finite subsets of natural numbers.

**Definition 3.4** For $s, t \in [\omega]^{<\infty}$ we say that $t$ is a shift of $s$ and write $s \lhd t$ if there exists $X \in [\omega]^\infty$ such that

$$s \sqsubset X \text{ and } t \sqsubset {}_*X.$$

**Definitions 3.5** Let $Q$ be a qo and $f : F \to Q$ be a super-sequence.

1. We say that $f$ is *bad* if whenever $s \lhd t$ in $F$, we have $f(s) \not\leq f(t)$.
2. We say that $f$ is *good* if there exists $s, t \in F$ with $s \lhd t$ and $f(s) \leq f(t)$.

Using the notations introduced in Sect. 2.4.

**Lemma 3.6** *Let $Q$ be a quasi-order.*

1. *If $h : [\omega]^\infty \to Q$ is locally constant and bad, then $h^\downarrow : F^h \to Q$ is a bad super-sequence.*
2. *If $f : F \to Q$ is a bad super-sequence from a front on $X$, then $f^\uparrow : [X]^\infty \to Q$ is a bad locally constant multi-sequence.*

*Proof* 1. Suppose $h : [X]^\infty \to Q$ is locally constant and bad. Let us show that $h^\downarrow : F^h \to Q$ is bad. If $s, t \in F^h$ with $s \lhd t$, i.e. there exists $Y \in [X]^\infty$ such that $s \sqsubset Y$ and $t \sqsubset {}_*Y$. Then $h^\downarrow(s) = h(Y)$ and $h^\downarrow(t) = h({}_*Y)$ and since $h$ is assumed to be bad, we have $h^\downarrow(s) \not\leq h^\downarrow(t)$.
2. Suppose $f : F \to Q$ is bad from a front on $X$ and let $Y \in [X]^\infty$. There are unique $s, t \in F$ such that $s \sqsubset Y$ and $t \sqsubset {}_*Y$, and clearly $f^\uparrow(Y) = f(s)$, $f^\uparrow({}_*Y) = f(t)$, and $s \lhd t$. Therefore $f^\uparrow(X) \not\leq f^\uparrow({}_*X)$ holds. □

We finally have the equivalence between both definitions.

**Corollary 3.7** *A quasi-order $Q$ is a* BQO *if and only if there is no bad super-sequence in $Q$.*

In particular, since a bad sequence is an instance of a bad super-sequence, it follows that every BQO is WQO.

## 3.2 First Examples and Finite Stability

Every constant super-sequence is good, so Theorem 2.12 can be reformulated as follows:

*Example 3.8* Every finite quasi-order is a better-quasi-order.

If $Q$ is a well-order and $h : [X]^\infty \to Q$ is a multi-sequence in $Q$, we can consider the sequence $(X_n)_{n\in\omega}$ in $[X]^\infty$ defined inductively by $X_0 = X$ and $X_{n+1} = {}_*X_n$. As $Q$ is a well-order, $h(X_n) \not\leq_Q h(X_{n+1})$ implies $h(X_{n+1}) <_Q h(X_n)$. Since the sequence $(h(X_n))_{n\in\omega}$ cannot be strictly decreasing in $Q$, there exists $n$ such that $h(X_n) \leq_Q h(X_{n+1}) = h({}_*X_n)$. We have obtained the following:

*Example 3.9* Every well-order is a better-quasi-order.

**Lemma 3.10** *Suppose $h : (P, \leq_P) \to (Q, \leq_Q)$ is map such that $h(p) \leq h(p')$ implies $p \leq p'$ for all $p, p' \in P$. If $Q$ is* BQO, *then $P$ is* BQO.

*Proof* If $f : F \to P$ is a bad super-sequence in $P$, then $h \circ f : F \to Q$ is a bad super-sequence in $Q$. □

Theorem 2.11 also gives the following easy closure property.

**Proposition 3.11** *If* $(Q, \leq_Q)$ *and* $(P, \leq_P)$ *are* BQO*, then so is* $(P \cup Q, \leq_P \cup \leq_Q)$*. In other words, a finite union of* BQO *is still* BQO*.*

*Proof* Take a super-sequence $f : F \to P \cup Q$. Apply Theorem 2.11 to the partition $\{f^{-1}(P), f^{-1}(Q)\}$ of $F$ to get a sub-front $G$ of $F$ that defines a sub-super-sequence which ranges either in $P$ or in $Q$. Since both $P$ and $Q$ are BQO, $g$ is good, and in turn so is $f$. □

The following dichotomy is an easy corollary of Theorem 2.11 which turns out to be quite useful when dealing with multi-sequences or super-sequences.

**Proposition 3.12** *Let* $E$ *be a set and* $R$ *a binary relation on* $E$*. Then every multi-sequence* $f : [X]^\infty \to E$ *admits a sub-multi-sequence* $g : [Z]^\infty \to E$ *such that either* $g(Y)\ R\ g({}_*Y)$ *holds for all* $Y \in [Z]^\infty$*, or* $g(Y)\ R\ g({}_*Y)$ *holds for no* $Y \in [Z]^\infty$*.*

*Proof* Define $c : [X]^\infty \to 2$ by $c(Y) = 1$ if and only if $f(Y)\ R\ f({}_*Y)$. Clearly $c$ is locally constant so by Corollary 3.7 there exists an infinite subset $Z$ of $X$ such that $c : [Z]^\infty \to 2$ is constant. The corresponding sub-multi-sequence $g = f\restriction_{[Z]^\infty} : [Z]^\infty \to E$ is as desired. □

The analogue result for super-sequences also holds and easily follows from the result for multi-sequences:

**Proposition 3.13** *Let* $E$ *be a set and* $R$ *a binary relation on* $E$*. Then every super-sequence* $f : F \to E$ *admits a sub-super-sequence* $g : G \to E$ *such that*

**either** *for all* $s, t \in G$*,* $s \lhd t$ *implies* $g(s)\ R\ g(t)$*,*
**or** *for all* $s, t \in G$*,* $s \lhd t$ *implies that* $g(s)\ R\ g(t)$ *does not hold.*

When $(P, \leq_P)$ and $(Q, \leq_Q)$ are two quasi-orders, $\leq_P \times \leq_Q$ stands for the *product* quasi-order on $P \times Q$, that is

$$(p, q) \leq_P \times \leq_Q (p', q') \quad \longleftrightarrow \quad p \leq_P p' \text{ and } q \leq_Q q'.$$

**Proposition 3.14** *If* $(Q, \leq_Q)$ *and* $(P, \leq_P)$ *are* BQO*, then so is* $(P \times Q, \leq_P \times \leq_Q)$*. Therefore a finite product of* BQOs *is still* BQO*.*

*Proof* Suppose that $f : [X]^\infty \to P \times Q$ is a bad multi-sequence and write $f(Y) = (f_P(Y), f_Q(Y))$. Then for every $Y \in [X]^\infty$ either $f_P(Y) \not\leq_P f_P({}_*Y)$ or $f_Q(Y) \not\leq_Q f_Q({}_*Y)$. Applying Proposition 3.12 to $f$ and the binary relation $(p, q)\ R\ (p', q')$ iff $p \leq_P p'$ we obtain either that $f_Q$ is a bad multi-sequence in $(Q, \leq_Q)$ or that $f_P$ is a bad multi-sequence in $(P, \leq_P)$. □

### *3.3 The Real Deal: Infinite Stability*

We now turn to stability under infinitary operations, which was the original motivation behind the introduction of BQO; to quote Marcone[4] "the general pattern [is] that if a finitary operation preserves WQO then its infinitary version preserves BQO." More specifically, we are interested in operations $Q \mapsto O(Q)$ which are infinitary in the sense that each member of $O(Q)$ can be thought of as an infinite structure labeled by elements of $Q$. As an example we already have seen $Q \mapsto \mathcal{D}(Q)$. While Higman's Theorem ensures that finite sequences in a WQO again form a WQO, we now turn to the infinitary analogue: the operation of taking infinite sequences. Let us define what we mean by that, assuming basic knowledge concerning ordinals.[5]

Let $(Q, \leq_Q)$ be a qo. A *transfinite sequence* $\bar{q}$ in $Q$ is a map from an ordinal $\alpha$ to $Q$, $\alpha$ being then the length of $\bar{q}$, denoted by $|\bar{q}|$. The notation $Q^\alpha$ stands for the sequences in $Q$ of length $\alpha$, and $Q^{\mathsf{ON}}$ denotes the class of all transfinite sequences in $Q$. Given two sequences $\bar{q}$ and $\bar{p}$ in $Q^{\mathsf{ON}}$ of respective lengths $\alpha$ and $\beta$, we write $\bar{q} \leq_{Q^{\mathsf{ON}}} \bar{p}$ if there is an increasing injection $\iota : \alpha \to \beta$ satisfying $\bar{q}(\xi) \leq_Q \bar{p}(\iota(\xi))$ for all $\xi$ in $\alpha$. Notice in particular that $\bar{q} \leq_{Q^{\mathsf{ON}}} \bar{p}$ implies $|\bar{q}| \leq |\bar{p}|$.

Observe that for Rado's partial order $\mathfrak{R}$ from Example 1.2 $\mathfrak{R}^\omega$ is not WQO, hence $Q \mapsto Q^{\mathsf{ON}}$ does not preserve WQOness. However $Q^{\mathsf{ON}}$ is BQO whenever $Q$ is BQO and we now outline the proof this result. The central element to this proof is the so-called *Minimal Bad Lemma*, that we state without proof. It is a key result in many theorems concerning BQO.

**Definition 3.15** Let $(Q, \leq_Q)$ be a qo.

- A *partial ranking* of $(Q, \leq_Q)$ is a well-founded quasi-order $\leq'$ on $Q$ such that $p \leq' q$ implies $p \leq_Q q$.
- Given any qo $(Q, \leq')$ and multi-sequences $f : [X]^\infty \to Q$ and $g : [Y]^\infty \to Q$ we write $f \leq' g$ (resp. $f <' g$) when we have both $X \subseteq Y$ and $f(Z) \leq' g(Z)$ (resp. $f(Z) <' g(Z)$) for all $Z \in [X]^\infty$.
- Given a partial ranking $\leq'$ of $(Q, \leq_Q)$, a locally constant multi-sequence $g : [Y]^\infty \to Q$ that is bad with respect to $\leq_Q$ is *minimal bad* if every locally constant multi-sequence $f$ satisfying $f <' g$ is good.

**Theorem 3.16** (Minimal Bad Lemma) *Let $\leq'$ be a partial ranking of a quasi-order $(Q, \leq_Q)$. If $f : [X]^\infty \to Q$ is a locally constant bad multi-sequence, then there is a locally constant multi-sequence $g \leq' f$ that is minimal bad.*

We refer the interested reader to [35] (see also [20]) for a condensed proof of that result. Note however that we restricted ourselves here to talking about locally constant multi-sequences where in both of the above references the authors deal with Borel multi-sequences. This is not an issue since as we explained in Remark 3.3 every Borel multi-sequence admits a locally constant sub-multi-sequence.

We borrow the proof of the following lemma to [33].

---

[4] See the introduction of [21].

[5] As treated for instance in any introduction to set theory.

**Lemma 3.17** *Given $(Q, \leq_Q)$ a qo, if $\bar{q}$ and $\bar{p}$ in $Q^{\mathsf{ON}}$ satisfy $\bar{q} \not\leq_{Q^{\mathsf{ON}}} \bar{p}$ then there is $\theta < |\bar{q}|$ such that $\bar{q}\restriction_\theta \leq_{Q^{\mathsf{ON}}} \bar{p}$ but $\bar{q}\restriction_{\theta+1} \not\leq_{Q^{\mathsf{ON}}} \bar{p}$.*

*Proof* Define a map $h : |\bar{q}| \to |\bar{p}| + 1$ by induction: $h(\alpha)$ is the minimal $\xi < |\bar{p}|$ such that $\bar{q}(\alpha) \leq_Q \bar{p}(\xi)$ and $\xi > h(\beta)$ for all $\beta < \alpha$, if such a $\xi$ exists; and $h(\alpha) = |\bar{p}|$ otherwise. Notice that $\bar{q} \leq_{Q^{\mathsf{ON}}} \bar{p}$ iff $h(\alpha) < |\bar{p}|$ for all $\alpha < |\bar{q}|$.

Now since $\bar{q} \not\leq_{Q^{\mathsf{ON}}} \bar{p}$, there is a minimal $\theta$ such that $h(\theta) = |\bar{p}|$. By minimality of $\theta$ we have $\bar{q}\restriction_\theta \leq_{Q^{\mathsf{ON}}} \bar{p}$, and by definition of $h$ we have $\bar{q}\restriction_{\theta+1} \not\leq_{Q^{\mathsf{ON}}} \bar{p}$. □

We are ready to prove stability under the taking of tranfinite sequences. As a matter of fact, we prove a stronger property that Louveau and Saint-Raymond in [18] call, in our terminology, *reflection of bad multi-sequences*.

**Theorem 3.18** (Nash-Williams) *If $(Q, \leq_Q)$ is a quasi-order and $f : [X]^\infty \to Q^{\mathsf{ON}}$ is a locally constant bad multi-sequence, then there is $Y \in [X]^\infty$ and a locally constant multi-sequence $\varphi$ from $[Y]^\infty$ to the ordinals such that $f \circ \varphi$ is a (locally constant) bad multi-sequence in $Q$.*

*Proof* We first define a partial ranking of $(Q^{\mathsf{ON}}, \leq_{Q^{\mathsf{ON}}})$. For $\bar{q}$ and $\bar{p}$ in $Q^{\mathsf{ON}}$ define $\bar{q} \leq' \bar{p}$ if and only if $\bar{q}$ is a prefix of $\bar{p}$, that is: there is $\theta \leq |\bar{p}|$ such that $\bar{q} = \bar{p}\restriction_\theta$. Clearly $\leq'$ is a partial ranking.

Take a locally constant bad multi-sequence $f : [X]^\infty \to Q^{\mathsf{ON}}$, and apply Theorem 3.16 to get $X' \in [X]^\infty$ and a locally constant minimal bad multi-sequence $g : [X']^\infty \to Q^{\mathsf{ON}}$ such that $g \leq' f$. For every $Z \in [X']^\infty$ let $\varphi(Z)$ be the unique ordinal $< |g(Z)|$ such that $g(Z)\restriction_{\varphi(Z)} \leq_{Q^{\mathsf{ON}}} g({}_*Z)$ but $g(Z)\restriction_{\varphi(Z)+1} \not\leq_{Q^{\mathsf{ON}}} g({}_*Z)$, the existence of which is granted by Lemma 3.17.

Let us check that the map $\varphi : [X']^\infty \to \mathsf{ON}$ is locally constant. For any $Z \in [X']^\infty$, since $g$ is locally constant there are basic open sets $U \ni Z$ and $V \ni {}_*Z$ on which $g$ is constant. By continuity of the shift map, we can find a basic open set $W$ with $Z \in W \subseteq U$ such that ${}_*Y \in V$ for all $Y \in W$. It follows that $\varphi$ is constant on $W$, as desired.

Notice that by definition of $\varphi(Z)$ we have $g(Z)\restriction_{\varphi(Z)} <' g(Z)$ for all $Z \in [X']^\infty$, so by minimality of $g$ the multi-sequence $Z \mapsto g(Z)\restriction_{\varphi(Z)}$ is good. We can now apply Proposition 3.12 to obtain $Y \in [X']^\infty$ such that in fact $g(N)\restriction_{\varphi(N)} \leq_{Q^{\mathsf{ON}}} g({}_*N)\restriction_{\varphi(N)}$ for every $N \in [Y]^\infty$. So to sum up, we have for all $Z \in [Y]^\infty$:

$$g(Z)\restriction_{\varphi(Z)} \leq_{Q^{\mathsf{ON}}} g({}_*Z)\restriction_{\varphi({}_*Z)} \text{ but } g(Z)\restriction_{\varphi(Z)+1} \not\leq_{Q^{\mathsf{ON}}} g({}_*Z)\restriction_{\varphi({}_*Z)+1}.$$

This implies that for all $Z \in [Y]^\infty$ we have $g(Z)(\varphi(Z)) \not\leq_Q g({}_*Z)(\varphi({}_*Z))$. Notice finally that as $g(Z)$ is a prefix of $f(Z)$, we have $f \circ \varphi = g \circ \varphi$ which concludes the proof. □

The previous result enjoys a converse, proven by Pouzet in [26]:

**Theorem 3.19** (Pouzet) *A quasi-order $(Q, \leq_Q)$ is* BQO *if and only if $(Q^{\mathsf{ON}}, \leq_{Q^{\mathsf{ON}}})$ is* WQO.

This technique was instrumental in finding many examples of BQOs, the first and arguably most famous of these examples was found by Laver, a couple of years only after Nash-Williams' results. We recall it here, along with some others, in chronological order. We leave the interested reader to look for precise definitions in the references.

**Theorem 3.20** *The following quasi-orders are* BQO*s:*

1. *Embeddability between $\sigma$-scattered linear orders (*Laver *[16]).*
2. *Surjective homomorphism between countable linear orders (*Landraitis *[15]).*
3. *Embeddability between countable trees (*Corominas *[5]).*
4. *Continuous embeddability between countable linear orders (*van Engelen-Miller-Steel *[35]).*
5. *Embeddability between countable $N$-free partial orders (*Thomassé *[34]).*

We would like to point out that stronger versions of the above results are available, generally in the same articles as those cited. The general method for these proofs follows indeed the same pattern as Theorem 3.18, and what is proven is in general reflection of bad multi-sequences, which gives a BQO result for BQO-*labelled* structures. For more on this, see [18, Sect. 3].

That being said, there are other ways to prove that a certain class is BQO. The most famous example of this involves games, it is due to Wadge and is called *continuous reducibility*. Once again, we mention some of them and leave the reader look for the specifics.

**Theorem 3.21** *The following quasi-orders are* BQO*s:*

1. *Continuous reducibility between Borel subsets of 0-dimensional Polish spaces (*Wadge, Martin, van Engelen-Miller-Steel, see for instance *[35]).*
2. *Embeddability between Borel sub-orders of $\mathbb{R}^2$ (*Louveau-Saint Raymond, *[18]).*
3. *Assuming Projective Determinacy, embeddability between Borel sub-orders of $\mathbb{R}^n$ for all $n \in \omega$ (*Louveau-Saint Raymond, *[18]).*
4. *Topological embeddability between 0-dimensional Polish spaces (*see *[3]).*

In all these examples, BQO is used to prove that some quasi-order is WQO. We would now like to turn to some results of a somewhat different nature, focusing on yet another operation: passing from a quasi-order to the quasi-order of ideals. This is a central notion for the last two sections of this article.

**Definition 3.22** Let $P$ be a partial order. An *ideal* of $P$ is a subset $I \subseteq P$ such that

1. $I$ is non empty;
2. $I$ is a downset;
3. for every $p, q \in I$ there exists $r \in I$ with $p \leq r$ and $q \leq r$.

We write $\mathrm{Id}(P)$ for the set of ideals of $P$ partially ordered by inclusion.

Equivalently, a subset $I$ of a po $P$ is an ideal if $I$ a downset and $I$ is *directed*, namely every (possibly empty) finite subset $F \subseteq I$ admits an *upper bound* in $I$, i.e. there exists $q \in I$ with $F \subseteq \downarrow q$. For every $p \in P$, the set $\downarrow p$ is an ideal called a *principal* ideal.

For each quasi-order $(Q, \leq_Q)$ we have the embedding

$$\begin{aligned} Q &\longmapsto \mathrm{Id}(Q) \\ q &\longmapsto \downarrow q, \end{aligned}$$

and we henceforth identify each element $p$ with the corresponding principal ideal $\downarrow p$. In particular we have the inclusions $P \subseteq \mathrm{Id}(P) \subseteq \mathcal{D}(P)$ as partial orders.

We observe that we cannot replace $\mathcal{D}(P)$ by $\mathrm{Id}(P)$ in Proposition 1.1, Item 3, i.e. it is *not* true that a po $P$ is WQO if and only $\mathrm{Id}(P)$ is well-founded. The simplest example is given by the antichain $A = (\omega, =)$. The partial order $\mathrm{Id}(A)$ is equal to $A$ so, in particular, even though $A$ is not WQO, $\mathrm{Id}(A)$ is well-founded. Nonetheless, when $P$ is WQO then $\mathrm{Id}(P)$ is well-founded.

A *non-principal* ideal of $P$ is an ideal which is not of the form $\downarrow p$ for some $p \in P$. We write $\mathrm{Id}^*(P)$ for the partial order of non-principal ideals, i.e. $\mathrm{Id}^*(P) = \mathrm{Id}(P) \setminus P$. The partial order $\mathrm{Id}^*(P)$ is therefore the remainder of the ideal completion of $P$.

The following result was conjectured by [27] and proved by the authors in [4].

**Theorem 3.23** *Let $Q$ be* WQO. *If* $\mathrm{Id}^*(Q)$ *is* BQO, *then $Q$ is* BQO.

The first corollary of Theorem 3.23 that we mention is:

**Corollary 3.24** *If $Q$ is* WQO *and* $\mathrm{Id}^*(Q)$ *is finite, then $Q$ is* BQO.

This result is due to Pouzet [27] and a direct proof is presented by Fraïssé [11, Chap. 7, Sect. 7.7.8].

This first simple corollary already allows us to prove the following proposition, a particular case of which was used by the first author in [3].

**Proposition 3.25** *Let $\varphi : \omega \to \omega$ be progressive, i.e. such that $n \leq \varphi(n)$ for every $n \in \omega$. Then the partial order $\leq_\varphi$ on $\omega$ defined by*

$$m \leq_\varphi n \quad \longleftrightarrow \quad m = n \text{ or } \varphi(m) < n.$$

*is a better-quasi-order.*

*Proof* Let $g : \omega \to \omega$ be any sequence. Then either $g$ is bounded in the usual order and so $g$ is good for $\leq_\varphi$, or $g$ is unbounded in the usual order and so there exists $n$ such that $g(n) > \varphi(g(0))$ and so $g$ is good for $\leq_\varphi$. Hence $(\omega, \leq_\varphi)$ is WQO.

Now let $I$ be a non principal ideal in $(\omega, \leq_\varphi)$. In particular $I$ is an infinite subset of $\omega$, so for every $m \in \omega$ there exists $n \in I$ such that $\varphi(m) < n$ and so $m \leq_\varphi n \in I$. Therefore $I = \omega$ and so there is exactly one non-principal ideal of $(\omega, \leq_\varphi)$. It follows by Corollary 3.24 that $(\omega, \leq_\varphi)$ is BQO. □

## 4 In-Between: An Intractable Diversity of Inconspicuous Orders

The following classes of WQOs are sometimes considered as approximations of the concept of BQO.

**Definition 4.1** Let $Q$ be a quasi-order and $1 \leq \alpha < \omega_1$. We say that $Q$ is $\alpha$-BQO if and only if every super-sequence $f : F \to Q$ with $\text{rk}\, F \leq \alpha$ is good.

*Remark 4.2* The authors in [19, 30], for example, use a different definition of $\alpha$-BQO which is easily seen to be equivalent to ours.

Clearly a qo is WQO if and only if it is 1-BQO, and it is BQO if and only if its $\alpha$-BQO for every $\alpha < \omega_1$. Rado's poset, as shown in Example 1.2 is WQO but is not 2-BQO.

Marcone showed in [19] that these notions are all distinct.

**Theorem 4.3** *For every countable ordinal $\alpha$ there exists a quasi-order that is $\beta$-*BQO *for all $\beta < \alpha$ but that is not $\alpha$-*BQO.

One natural attempt to prove this theorem consists of considering on any front of rank equal to $\alpha$ the complement of the binary relation $\lhd$. This binary relation however fails to be a quasi-order.[6] Marcone's proof actually amounts to first showing that on "well chosen" fronts the complement of $\lhd$ lacks only transitivity in order to be the desired counter-example. Then one chooses an enumeration of the front at stake before using a result due to Pouzet to fix the transitivity issue (see [19, Theorem 1.8]).

In this section we present examples of quasi-orders which are $n$-BQO but not $(n+1)$-BQO for each $n \geq 1$ and an example of a quasi-order which is not $\omega$-BQO but is $n$-BQO for all $n < \omega$. While their definition is simple and it is easy to see that they fail to be BQO, it does require some work to show they do enjoy a fair share of BQOness. Here we follow a new approach based on ideals and which relies on the following easy refinement of Theorem 3.23 whose proof can be found in the second author's Ph.D. thesis [25, Theorem 4.43]:

**Theorem 4.4** *For every $n \in \omega$, if $Q$ is* WQO *and* $\text{Id}^*(Q)$ *is $n$-*BQO *then $Q$ is $(1+n)$-*BQO.

Before going further, let us stop on our crucial example once again, to illustrate the rôle played by ideals.

*Example 4.5* (*Rado's poset continued*) We continue on Rado's partial order $\mathfrak{R}$ defined in Example 1.2 and compute $\text{Id}(\mathfrak{R})$. We claim that $\text{Id}(\mathcal{R}) = \mathcal{R} \cup \{I_n \mid n \in \omega\} \cup \{\top\}$ where $I_n = \downarrow\big\{\{n, k\} \mid n < k\big\}$ for $n \in \omega$ and $\top = \mathcal{R}$. We have $\{m, n\} \leq I_k$

[6] The only notable exception is $[\omega]^1$ where both $\lhd$ and its complement are actually transitive. If $F$ is a front on $X = \{x_0, x_1, x_2, \ldots\}$, $Y = \{x_3, x_4, x_5, \ldots\}$ and if $s, t, u \in F$ are such that $s \sqsubset \{x_0\} \cup Y$, $|s| \geq 2$, $t \sqsubset \{x_1, x_2\} \cup Y$, $|t| \geq 2$ and $u \sqsubset Y$. Then $s \lhd u$, while neither $s \lhd t$ nor $t \lhd u$.

if and only if $m = k$ or $n < k$, and $a \leq \top$ for all $a \in \mathrm{Id}(\mathcal{R})$. The non principal ideals are the $I_n$s and $\top$. We show there are no other ideals. Let $I$ be an ideal of $\mathcal{R}$. First suppose for all $k \in \omega$ there exists $\{m, n\} \in I$ with $k < m$, then $I = \top$. Suppose now that there exists $m = \max\{k \mid \exists l \ \{k, l\} \in I\}$. If there is infinitely many $n$ such that $\{m, n\} \in I$ then $I = I_m$. Otherwise $I = \downarrow\{m, n\}$ for $n = \max\{l \mid \{m, l\} \in I\}$.

Observe that $(I_n)_{n\in\omega}$ is an antichain in $\mathrm{Id}(\mathcal{R})$, hence a bad sequence witness to the fact that $\mathrm{Id}(\mathcal{R})$ is not WQO.

Here is the definition of the explicit counter-examples that we know. The quasi-orders defined on fronts of finite ranks first appeared in [26] and the quasi-order on $\mathcal{S}$ was defined by Assous–Pouzet [24].

**Definition 4.6** For every $n \geq 1$ we let $\mathfrak{R}_n = ([\omega]^n, R_n)$ where $R_n$ is the binary relation defined by

$$s \ R_n \ t \quad \longleftrightarrow \quad \begin{cases} s(i) \leq t(i), \text{ for all } i < n, \text{ and} \\ \text{if } s(0) < t(0), \text{ then there is } j > 0 \text{ with } s(j) < t(j-1). \end{cases}$$

Furthermore, we let $\mathfrak{R}_\omega = (\mathcal{S}, R_\omega)$ where $\mathcal{S} = \{s \in [\omega]^{<\infty} : |s| = s(0) + 1\}$ is the Schreier barrier and:

$$s \ R_\omega \ t \quad \longleftrightarrow \quad \begin{cases} \text{if } s(0) = t(0), \text{ then } s(i) \leq t(i) \text{ for all } i < |s|, \text{ and} \\ \text{if } s(0) < t(0), \text{ then there is } n \leq |s| \text{ such that } s{\restriction_n} \ R_n \ t{\restriction_n}. \end{cases}$$

Notice that $\mathfrak{R}_1$ is simply $([\omega]^1, =)$ and that $\mathfrak{R}_2$ is exactly Rado's poset $\mathfrak{R}$ from Example 1.2. Moreover it easy to check that the binary relations we defined are included in the complement of the binary relation $\lhd$.

**Lemma 4.7** *Each of the* $\mathfrak{R}_n$ *for* $n \geq 2$ *as well as* $\mathfrak{R}_\omega$ *are* WQO*s.*

*Proof* One needs to check that $R_\omega$ and each $R_n$ is transitive. We only treat one specific case which is a main obstacle to generalizing this idea to arbitrary fronts. Let $s, t, u \in \mathcal{S}$ with $s \ R_\omega \ t$ and $t \ R_\omega \ u$. Clearly $|s| \leq |u|$ and $s(i) \leq u(i)$ for all $i < |s|$. Now suppose that $s(0) = t(0)$ and $t(0) < u(0)$, so $s(0) < u(0)$. Then there exists $n \leq |t|$ such that $t{\restriction_n} R_n u{\restriction_n}$. By definition of $\mathcal{S}$, $s(0) = t(0)$ implies $|s| = |t|$, so we also have $n \leq |s|$. Since $s{\restriction_n} \ R_n \ t{\restriction_n}$ by transitivity of $R_n$ we get $s{\restriction_n} \ R_n \ u{\restriction_n}$ and so $s \ R_\omega \ u$ in this case, as desired.

Next we show that $\mathfrak{R}_\omega$ is WQO. The case of $\mathfrak{R}_n$ for $n \geq 2$ is similar. Suppose towards a contradiction that $(s_n)_{n\in\omega}$ is a bad sequence in $\mathfrak{R}_\omega$ and let $m_n = \min s_n$ for each $n$. By possibly going to a subsequence, we can assume that either $(m_n)_n$ is unbounded or else $(m_n)_n$ is constant equal to some $k$. In the latter case, it follows that the sequence $t_n = s_n \setminus \{m_n\}$, $n \in \omega$, is a bad sequence in $\omega^k$ with the pointwise ordering, a contradiction. In the former case, we can find $i$ with $m_i = |s_i| \geq 2$ and then $j > i$ such that $m_j > \max s_i$, so that we get $s_i \ R_{m_i} \ s_j{\restriction_{m_i}}$. This implies that $s_i \ R_\omega \ s_j$, a contradiction again. $\square$

While it is very tempting to try to generalize the above definition to arbitrary fronts, one should notice that very specific properties of the fronts $[\omega]^n$ and $\mathcal{S}$ are needed to prove their transitivity. As a matter of fact, so far any attempt to generalize these examples has failed to be transitive.

We now compute the ideals of $\mathfrak{R}_{n+1}$.

**Definition 4.8** For every $n \geq 1$, we define $\mathfrak{I}_n = ([\omega]^{\leq n}, S_n)$ where

$$s\ S_n\ t \quad \longleftrightarrow \quad \begin{cases} |s| \geq |t| \text{ and } s(i) \leq t(i) \text{ for all } i < |t|, \text{ and} \\ \text{if } |s| = |t| = n, \text{ then } s\ R_n\ t. \end{cases}$$

where $R_n$ is the relation from Definition 4.6.

**Proposition 4.9** *For ever natural number $n \geq 1$, $\mathrm{Id}^*(\mathfrak{R}_{n+1})$ is isomorphic to $\mathfrak{I}_n$.*

*Proof* For $s \in [\omega]^{\leq n}$ we say that a sequence $(s_i)_{i\in\omega}$ in $\mathfrak{R}_{n+1}$ is *good for $s$* if

1. $s \sqsubset s_i$ for all $i$,
2. $(s_i(|s|))_{i\in\omega}$ is strictly increasing, and
3. $(s_i)_{i\in\omega}$ is strictly increasing in $\mathfrak{R}_{n+1}$.

Let $s, t \in [\omega]^{\leq n}$ and suppose that $(s_i)_{i\in\omega}$ and $(t_i)_{i\in\omega}$ are good for $s$ and $t$, respectively. We claim that if $\downarrow\{s_i \mid i \in \omega\} \subseteq \downarrow\{t_i \mid i \in \omega\}$ (where the downward closure is taken in $\mathfrak{R}_{n+1}$) then $s\ S_n\ t$. So in particular if $\downarrow\{s_i \mid i \in \omega\} = \downarrow\{t_i \mid i \in \omega\}$, then $s = t$.

First suppose towards a contradiction that $|s| < |t|$. There exists $k$ such that $s_k(|s|) > t(|s|)$ and so $s_k \notin \downarrow\{t_i \mid i \in \omega\}$ since for no $j$ we have $s_k R_{n+1} t_j$. A contradiction, hence $|s| \geq |t|$. Since there exists $i$ such that $s_0 R_{n+1} t_i$, it follows in particular that $s(j) \leq t(j)$ for all $j < |t|$. Finally assume that $|s| = |t| = n$ and that $s(0) < t(0)$. There is $i$ such that $s_i(|s|) \geq t(|s| - 1)$ and there is $j$ such that $s_i R_{n+1} t_j$. By definition of $R_{n+1}$ there is $k > 0$ such that $s_i(k) < t_j(k-1)$. Since necessarily $k < n$, it follows that $k$ is also a witness to the fact that $s\ R_n\ t$. This proves the claim.

Next we see that for every $s \in [\omega]^{\leq n}$ there exists a sequence $(s_i)_{i\in\omega}$ which is good for $s$. One easily checks that $s_i = \{2i, 2i+1, \ldots, 2i+n\}$, $i \in \omega$, is good for $\emptyset$. Moreover if $s \in [\omega]^{\leq n}$ is not empty and we let $s_i \in [\omega]^{n+1}$ with $s_i \sqsubset s \cup \omega/(i + \max(s))$ for each $i \in \omega$, then $(s_i)_{i\in\omega}$ is good for $s$.

For every $s \in [\omega]^{\leq n}$, let $I_s = \downarrow\{s_i \mid i \in \omega\}$ for some sequence $(s_i)_{i\in\omega}$ that is good for $s$. By the claim this is a well defined map $\mathfrak{I}_n \to \mathrm{Id}^*(\mathfrak{R}_{n+1})$ and moreover $I_s \subseteq I_t$ implies $s\ S_n\ t$.

To see this map is surjective, let $I$ be a non principal ideal of $\mathfrak{R}_{n+1}$. Since $I$ is a (countable) non principal ideal, there exists a strictly $R_{n+1}$-increasing sequence $(s_i)_{i\in\omega}$ such that $I = \downarrow\{s_i \mid i \in \omega\}$. By repeated application of the Ramsey Theorem and possibly going to a subsequence, we can assume that for all $j < n+1$ the sequence $(s_i(j))_{i\in\omega}$ is either constant or strictly increasing. Since $(s_i)_{i\in\omega}$ has no constant subsequence, this implies that there exists $s_I \in [\omega]^{\leq n}$ such that $(s_i)_{i\in\omega}$ is good for $s_I$, as desired.

Finally we prove that $s\ S_n\ t$ implies $I_s \subseteq I_t$. Let $(s_i)_{i\in\omega}$ and $(t_i)_{i\in\omega}$ be good for $s$ and $t$ respectively. We distinguish two cases. First assume that $|t| < n$, then for every $i$ there exists $j$ with $s_i(n) < t_j(|t|)$, which together with $s\ S_n\ t$ implies $s_i\ R_{n+1}\ t_j$ and so $I_s \subseteq I_t$. Next assume that $|s| = |t| = n$ and so $s\ R_n\ t$. Then for all $i$ there is $j$ with $s_i(n) \leq t_j(n)$, and this implies that $s_i\ R_{n+1}\ t_j$. □

We can now prove that each $\mathfrak{R}_n$ is the counter-example we wanted.

**Theorem 4.10** (Pouzet) *For all $n \geq 1$, $\mathfrak{R}_{n+1}$ is $n$-BQO but not $(n+1)$-BQO.*

*Proof* As we already observed, the identity map $[\omega]^{n+1} \to [\omega]^{n+1}$ is a bad super-sequence of rank $n+1$ in $\mathfrak{R}_{n+1}$, so it is not $(n+1)$-BQO. We prove that $\mathfrak{R}_{n+1}$ is $n$-BQO by induction on $n$. For $n = 1$, we already know that $\mathfrak{R}_2 = \mathfrak{R}$ is WQO. So let $n > 1$ and assume that $\mathfrak{R}_n$ is $(n-1)$-BQO. We showed in Lemma 4.7 that $\mathfrak{R}_{n+1}$ is WQO and by Proposition 4.9 $\mathrm{Id}^*(\mathfrak{R}_{n+1})$ is isomorphic to $\mathfrak{I}_n$. We show that $\mathfrak{I}_n$ is $(n-1)$-BQO and we then conclude the proof by Theorem 4.4.

Notice that if $P_k = ([\omega]^k, \leq^k)$ denotes the pointwise ordering where $s \leq^k t$ iff $s(i) \leq t(i)$ for all $i < k$, then $P_k$ is easily seen to be BQO by Lemma 3.10 and Proposition 3.14. Moreover by induction hypothesis $\mathfrak{R}_n$ is $(n-1)$-BQO, and therefore as in Proposition 3.11 it follows that $\mathfrak{R}_n \cup \bigcup_{k=0}^{n-1} P_k$ is also $(n-1)$-BQO. One easily checks that the identity map $\mathfrak{I}_n \to \mathfrak{R}_n \cup \bigcup_{k=0}^{n-1} P_k$ allows to conclude as in Lemma 3.10 that $\mathfrak{I}_n$ is $(n-1)$-BQO, as desired. □

Using the above result we can prove that $\mathfrak{R}_\omega$ is $n$-BQO for all $n$.

**Theorem 4.11** (Pouzet-Assous) *The qo $\mathfrak{R}_\omega$ is $n$-BQO for all $n \in \omega$, but it is not $\omega$-BQO.*

*Proof* As before, the identity map is a bad super-sequence of rank $\omega$ in $\mathfrak{R}_\omega$, so it remains to show that $\mathfrak{R}_\omega$ is $n$-BQO for all $n$. Take $F$ a front of rank $n$, $f : F \to \mathcal{S}$ a super-sequence in $\mathfrak{R}_\omega$, and let us prove that it is good. Using the fact that $\omega$ is BQO and applying Proposition 3.12, we can assume by possibly going to a sub-super-sequence that $s \lhd t$ in $F$ implies $f(s)(0) \leq f(t)(0)$. Applying Proposition 3.12 again, we can further assume that either $s \lhd t$ in $F$ implies $f(s)(0) = f(t)(0)$, or $s \lhd t$ in $F$ implies $f(s)(0) < f(t)(0)$.

In the former case, there actually exists $i$ such that $f(s)(0) = i$ for all $s \in F$. This is because the transitive closure $\lhd^*$ of $\lhd$ inside $F$ is directed; notice indeed that for any $s$ and $t$ in $F$, if $u \in F$ satisfies $\max(s \cup t) < \min(u)$ then both $s \lhd^* u$ and $t \lhd^* u$ hold. But for each $i$ the restriction of $R_\omega$ to $\{s \in \mathcal{S} \mid s(0) = i\}$ is BQO since $\omega$ is BQO and any finite product of BQOs is BQO by Proposition 3.14.

We now suppose that $s \lhd t$ in $F$ implies $f(s)(0) < f(t)(0)$. Applying Theorem 2.11 to the subset of $F$ given by $\{s \in F \mid f(s)(0) < n\}$ where $n$ is the rank of $F$, we obtain a sub-super-sequence which falls into one of the two following cases.

Either $f(s)(0) < n$ for every $s \in F$ and so $f$ is good since for each $i < n$ the restriction of $R_\omega$ to $\{s \in \mathcal{S} \mid s(0) = i\}$ is BQO and a finite union of BQOs is BQO by Proposition 3.11.

Or $f(s)(0) \geq n$ for every $s \in F$, and so $|f(s)| \geq n+1$ by definition of the Schreier barrier $\mathcal{S}$. In this case, we consider the super-sequence $g : F \to \mathfrak{R}_{n+1}$ given by $g(s) = f(s)\restriction_{n+1}$. Since $\mathfrak{R}_{n+1}$ is $n$-BQO by Theorem 4.10, it follows that $g$ is good and there exists $s, t \in F$ with $s \lhd t$ and $g(s)\ R_{n+1}\ g(t)$. As $f(s)(0) < f(t)(0)$ and $f(s)\restriction_{n+1}\ R_{n+1}\ f(t)\restriction_{n+1}$, we have $f(s)\ R_\omega\ f(t)$ and therefore $f$ is good as desired. □

## 5 Classes of Bqos Definable by Forbidden Pattern

We have seen in the last Section that a great variety of quasi-orders distinguish the notion of WQO from that of BQO. Importantly, these two concepts also are of distinct descriptive complexity: while the set of WQOs on $\omega$ is a $\Pi^1_1$-complete subset of the Cantor space, that of BQOs is $\Pi^1_2$-complete as proved by Marcone [19, 20]. In particular, while $Q$ is WQO iff neither $(\omega, =)$ nor $\omega^{\mathrm{op}}$ (the opposite of $\omega$) embeds into $Q$, BQO cannot be defined by forbidding pattern, at least not in any relevant way.[7]

However in some particular cases, forbidding finitely many patterns do imply BQOness. In this section, we deduce results of this kind from Theorem 3.23.

### 5.1 *Interval Orders*

Suppose that $Q$ is a WQO such that $\mathrm{Id}^*(Q)$ is a well-order. Since by Example 3.9 well-orders are BQO, such quasi-orders are BQO by Theorem 3.23.

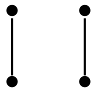
Observe that when $Q$ is WQO, since ideals are downsets and $\mathcal{D}(Q)$ is well-founded, $\mathrm{Id}^*(Q)$ is well-founded too. Hence, if $Q$ is WQO then $\mathrm{Id}^*(Q)$ is linearly ordered if and only if $\mathrm{Id}^*(Q)$ is a well-order.

What are the quasi-orders whose non principal ideals are linearly ordered? Well, assume $Q$ is a quasi-order and that $I, J \in \mathrm{Id}^*(Q)$ are incomparable for inclusion. Let $p \in I \setminus J$ and $q \in J \setminus I$. Then $p$ is incomparable with $q$. Forbidding antichains of size 2 in $Q$ is simply asking that $Q$ is a linear order, and of course well-orders are BQO. But we can do better: since $I$ and $J$ are non principal, there are $p' \in I$ with $p < p'$ and $q' \in J$ with $q < q'$. The restriction of the quasi-order on $Q$ to $\{p, q, p', q'\}$ is isomorphic to the partial order:

[7]Suppose that $\mathcal{B}$ is a basis for the set of BQOs on $\omega$, i.e. $B \subseteq 2^{\omega\times\omega}$ and for every qo $Q$ on $\omega$ we have $Q$ is BQO iff for no $B \in \mathcal{B}$ there exists an embedding from $B$ to $Q$. Then $\mathcal{B}$ is not analytic. Otherwise

$$Q \text{ is BQO} \quad \longleftrightarrow \quad \text{there exists no } B \in \mathcal{B} \text{ such that } B \text{ embeds in } Q$$

is a co-analytic definition of the set of BQOs on $\omega$, a contradiction with Marcone's Theorem.

$2 \oplus 2 =$ 

and therefore $2 \oplus 2$ embeds into $Q$. We are naturally led to following definition which appears frequently in the literature.

**Definition 5.1** A partial order $P$ is an *interval order* if the partial order $2 \oplus 2$ does not embed into $P$. In other words for every $p, q, x, y \in P$, $p < x$ and $q < y$ imply $p < y$ or $q < x$.

The preceding discussion yields the following which is already stated in [30].

**Theorem 5.2** *An interval order is* BQO *if and only if it is* WQO.

Notice that this theorem can be rephrased as follows: a partial order $P$ such that neither $(\omega, =)$, nor $\omega^{\mathrm{op}}$, nor $2 \oplus 2$ embeds into $P$ is a better-quasi-order.

It appears that the notion of *interval order* was first studied by the twenty-years-old Norbert Wiener [36] who credits Bertrand Russell for suggesting the subject [9]. Wiener was later acknowledged as the originator of cybernetics [6]. The Reverse Mathematics of interval orders is studied in [22].

For $p \in P$, let $\mathrm{Pred}(p) = \{q \in P \mid q < p\}$. It is easy to see that a partial order $P$ is an interval order if and only if the set $\{\mathrm{Pred}(p) \mid p \in P\}$ is linearly ordered by inclusion.

The terminology "interval order" was introduced by [8] and stems from the following characterisation.

A non trivial *closed interval* of a partial order $Q$ is a set of the form $[a, b] = \{q \in Q \mid a \leq q \leq b\}$ for some $a, b \in Q$ with $a < b$. We partially order the set $\mathrm{Int}(Q)$ of non trivial closed intervals of $Q$ by $[a, b] \leq [c, d]$ if and only if $a = c$ and $b = d$ or $b \leq c$.

For a partial order $P$ let us say that a map $I : P \to \mathrm{Int}(Q)$ is an *interval representation* of $P$ in $Q$ if for every $x, y \in P$ we have $x < y \leftrightarrow I(x) < I(y)$.

Let us first see that any partial order $P$ admits an interval representation. Let $\mathrm{Pred}^+(p) = \bigcap_{p<x} \mathrm{Pred}(x)$ and

$$Q_P = \{\mathrm{Pred}(p) \mid p \in P\} \cup \{\mathrm{Pred}^+(p) \mid p \in P\}$$

be partially ordered by inclusion.

**Proposition 5.3** ([2]) *Let $P$ be a partial order. The map*

$$\begin{aligned} I : P &\longrightarrow \mathrm{Int}(Q_P) \\ p &\longmapsto I_p = \Big(\mathrm{Pred}(p), \mathrm{Pred}^+(p)\Big) \end{aligned}$$

*is an interval representation of $P$ in $Q_P$.*

*Proof* First observe that for every $p \in P$ we have $\mathrm{Pred}(p) \subset \mathrm{Pred}^+(p)$ since $q < p$ imply $q \in \mathrm{Pred}(x)$ for all $x > p$, and in fact $p \in \mathrm{Pred}^+(p) \setminus \mathrm{Pred}(p)$. So $I$ is well defined. If $p < q$, then $\mathrm{Pred}^+(p) = \bigcap_{p<x} \mathrm{Pred}(x) \subseteq \mathrm{Pred}(q)$, and so $I_p < I_q$. Conversely if $I_p < I_q$, then $\mathrm{Pred}^+(p) \subseteq \mathrm{Pred}(q)$ and since $p \in \mathrm{Pred}^+(p)$ we have $p < q$. Hence $I$ is an interval representation of $P$. □

The following is a slight generalisation of a theorem by [8]. The proof we give here is due to [2].

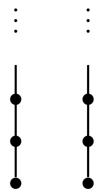
**Proposition 5.4** *A partial order $P$ is an interval order if and only if there exists an interval representation of $P$ in some linear order.*

*Proof* Suppose $I : P \to \mathrm{Int}(L)$ is an interval representation of $P$ in a linear order $L$ and let $p_0 < p_1$ and $q_0 < q_1$ in $P$. If $I(p_i) = [l_i, r_i]$ and $I(q_i) = [m_i, s_i]$ then $r_0 \leq l_1$ and $s_0 \leq m_1$. Since $L$ is linearly ordered, either $r_0 \leq m_1$ and so $p_0 < q_1$, or $m_1 \leq r_0$ and so $q_0 < p_1$. Therefore $P$ is an interval order.

Conversely, suppose $P$ is an interval order. By Proposition 5.3, it suffices to prove the $Q_P$ is linearly ordered. But $\{\mathrm{Pred}(p) \mid p \in P\}$ is linearly ordered and $\mathrm{Pred}^+(p) = \bigcap_{p<x} \mathrm{Pred}(x)$ is incomparable for the inclusion with some $X \in Q_P$ if and only if $\mathrm{Pred}(x)$ is incomparable with $X$ for some $x > p$. □

## 5.2 *More Classes of Better-Quasi-orders via Forbidden Patterns*

In fact continuing the above discussion we find that for any qo $Q$, $\mathrm{Id}^*(Q)$ is linearly ordered if and only if the po

$\omega \oplus \omega =$ 

does not embed into $Q$. We therefore have the following:

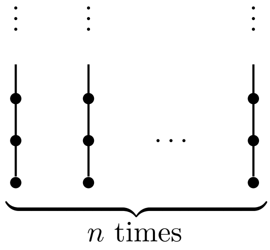
**Theorem 5.5** *If neither $(\omega, =)$, nor $\omega^{\mathrm{op}}$, nor $\omega \oplus \omega$ embed into $Q$, then $Q$ is* BQO.

Suppose now for a partial order $P$ that there exists a natural number $n$ such that the size of every antichain of $P$ is bounded by $n$. Then, by a theorem due to Dilworth [7], for $A$ an antichain of maximum size, say $n$, there exist subsets $P_i$, $i \in n$, such that $|P_i \cap A| = 1$, $P_i$ is linearly ordered and $\bigcup_{i \in n} P_i = P$ (see also [11, Sect. 4.14.1,

p. 141]). In particular, if $P$ is further assumed to be well-founded, then $P$ is BQO as a finite union of well-orders.

Continuing further the discussion of the previous subsection, we see that if there exists an antichain $A$ of size $n$ among the non principal ideals of a qo $Q$, then the partial order

$n \otimes \omega =$




embeds into $Q$. Indeed, assume that $\{I_i \mid i \in n\}, n \geq 2$ is an antichain of non principal ideals of a qo $Q$. For each $i \in n$ and every $j \in n$ with $i \neq j$, since $I_i \nsubseteq I_j$ we can pick $q_j \in I_i \setminus I_j$ and by the fact that $I_i$ is directed there is $q^i \in I_i$ with $q^i \notin I_j$ for every $j \neq i$. Now since each $I_i$ is non principal there exists a strictly increasing sequence $(q^i_k)_{k \in \omega}$ in $I_i$ with $q^i_0 = q^i$. This clearly yields and embedding of $n \otimes \omega$ into $Q$. Therefore

**Theorem 5.6** *Let $n \geq 1$. If neither $(\omega, =)$, nor $\omega^{\text{op}}$, nor $n \otimes \omega$ embed into $Q$, then $Q$ is* BQO.

In this theorem, for each $n \geq 1$, we have a class of BQO which is defined by finitely many forbidden patterns. Examples of classes of BQOs defined by mean of forbidden patterns—left alone by finitely many—are quite rare. In fact to our knowledge the previous theorem is the best result of this sort.

## References


1. Alvah Nash-Williams, C. S. J. (1965). On well-quasi-ordering transfinite sequences. In *Proceedings of Cambridge Philosophical Society* (Vol. 61, pp. 33–39). Cambridge: Cambridge University Press.
2. Bogart, K. P. (1993). An obvious proof of Fishburn's interval order theorem. *Discrete Mathematics*, *118*(1), 239–242.
3. Carroy, R. (2013). A quasi-order on continuous functions. *Journal of Symbolic Logic*, *78*(2), 633–648.
4. Carroy, R., & Pequignot, Y. (2014). From well to better, the space of ideals. *Fundamenta Mathematicae*, *227*(3), 247–270.
5. Corominas, E. (1985). On better quasi-ordering countable trees. *Discrete Mathematics*, *53*, 35–53.
6. Conway, F. D., & Siegelman, J. (2006). *Dark hero of the information age: In search of Norbert Wiener—The father of cybernetics*. Basic Books.
7. Dilworth, R. P. (1950). A decomposition theorem for partially ordered sets. *Annals of Mathematics*, 161–166.

8. Fishburn, P. C. (1970). Intransitive indifference with unequal indifference intervals. *Journal of Mathematical Psychology*, *7*(1), 144–149.
9. Fishburn, P., & Monjardet, B. (1992). Norbert Wiener on the theory of measurement 1914 1915 1921. *Journal of Mathematical Psychology*, *36*(2), 165–184.
10. Fraïssé, R. (1948). Sur la comparaison des types d'ordres. *Comptes rendus de l'Académie des Sciences*, *226*, 1330–1331.
11. Fraïssé, R. (2000). *Theory of relations*. Studies in Logic and the Foundations of Mathematics Elsevier Science.
12. Galvin, F., & Prikry, K. (1973). Borel sets and Ramsey's theorem. *The Journal of Symbolic Logic*, *38*, 193–198.
13. Higman, G. (1952). Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society*, *3*(1), 326–336.
14. Kruskal, J. B. (1972). The theory of well-quasi-ordering: A frequently discovered concept. *Journal of Combinatorial Theory, Series A*, *13*(3), 297–305.
15. Landraitis, C. (1979). A combinatorial property of the homomorphism relation between countable order types. *Journal of Symbolic Logic*, *44*, 403–411.
16. Laver, R. (1971). On Fraïssé's order type conjecture. *The Annals of Mathematics*, *93*(1), 89–111.
17. Laver, R. (1976). Well-quasi-orderings and sets of finite sequences. *Mathematical Proceedings of the Cambridge Philosophical Society*, *79*(01), 1–10.
18. Louveau, A., & Saint-Raymond, J. (1990). On the quasi-ordering of Borel linear orders under embeddability. *Journal of Symbolic Logic*, *55*(2), 537–560.
19. Marcone, A. (1994). Foundations of bqo theory. *Transactions of the American Mathematical Society*, *345*(2), 641–660.
20. Marcone, A. (1995). The set of better quasi orderings is $\pi^1_2$-complete. *Mathematical Logic Quarterly*, *41*, 373–383.
21. Marcone, A. (2005). Wqo and bqo theory in subsystems of second order arithmetic. In S. G. Simpson (Ed.) *Reverse mathematics 2001* (Vol. 21). Association for Symbolic Logic.
22. Marcone, A. (2007). Interval orders and reverse mathematics. *Notre Dame Journal of Formal Logic*, *48*(3), 425–448.
23. Montalbán, A. (2017). Fraïssé's conjecture in Pi-1-1-comprehension. Preprint.
24. Pouzet, M., & Assous, R. Personal communication.
25. Pequignot, Y. (2015). *Better-quasi-order: ideals and spaces*. Ph.D. thesis, Université de Lausanne and Université Paris Diderot–Paris 7.
26. Pouzet, M. (1972). Sur les prémeilleurordres. *Annales de l'institut Fourier (Grenoble)*, *22*, 1–20.
27. Pouzet, M. (1978). *Sur la théorie des relations*. Thèse d'état: Université Claude Bernard.
28. Pouzet, M. (1985). Applications of well quasi-ordering and better quasi-ordering. In *Graphs and order*, pp. 503–519. Dordrecht: Springer.
29. Pudlák, P., & Rödl, V. (1982). Partition theorems for systems of finite subsets of integers. *Discrete Mathematics*, *39*(1), 67–73.
30. Pouzet, M., & Sauer, N. (2006). From well-quasi-ordered sets to better-quasi-ordered sets. *Journal of Combinatorics*, *13*(4), 101.
31. Rado, R. (1954). Partial well-ordering of sets of vectors. *Mathematika*, *1*(02), 89–95.
32. Robertson, N.,& Seymour, P. D. (2004). Graph minors. XX. Wagner's conjecture. *Journal of Combinatorial Theory, Series B*, *92*(2), 325–357. Special Issue Dedicated to Professor W.T. Tutte.
33. Simpson, S. G. (1985). Bqo theory and Fraïssé's conjecture. In R. B. Mansfield & G. Weitkamp (Eds.), *Recursive aspects of descriptive set theory* (pp. 124–138). Oxford: Oxford University Press.
34. Thomassé, S. (2000). On better-quasi-ordering countable series-parallel orders. *Transactions of the American Mathematical Society*, *352*(6), 2491–2505.
35. van Engelen, F., Miller, A. W., & Steel, J. (1987). Rigid borel sets and better quasiorder theory. *Contemporary Mathematics*, *65*.
36. Wiener, N. (1914). A contribution to the theory of relative position. *Proceedings of the Cambridge Philosophical Society*, *17*, 441–449.

# On Ordinal Invariants in Well Quasi Orders and Finite Antichain Orders

**Mirna Džamonja, Sylvain Schmitz and Philippe Schnoebelen**

**Abstract** We investigate the ordinal invariants height, length, and width of well quasi orders (WQO), with particular emphasis on width, an invariant of interest for the larger class of orders with finite antichain condition (FAC). We show that the width in the class of FAC orders is completely determined by the width in the class of WQOs, in the sense that if we know how to calculate the width of any WQO then we have a procedure to calculate the width of any given FAC order. We show how the width of WQO orders obtained via some classical constructions can sometimes be computed in a compositional way. In particular this allows proving that every ordinal can be obtained as the width of some WQO poset. One of the difficult questions is to give a complete formula for the width of Cartesian products of WQOs. Even the width of the product of two ordinals is only known through a complex recursive formula. Although we have not given a complete answer to this question we have advanced the state of knowledge by considering some more complex special cases and in particular by calculating the width of certain products containing three factors. In the course of writing the paper we have discovered that some of the relevant literature was written

Mirna Džamonja thanks the Leverhulme Trust for a Research Fellowship for the academic year 2014/2015, the Simons Foundation for a Visiting Fellowship in the Autumn of 2015 and the Isaac Newton Institute in Cambridge for their support during the HIF programme in the Autumn of 2015, supported by EPSRC Grant Number EP/K032208/1. She most gratefully acknowledges the support of the *Institut d'histoire et de philosophie des sciences et des techniques (IHPST)* at University Paris 1, where she is as an Associate Member. The three authors thank the London Mathematical Society for their support through a Scheme 3 Grant.

M. Džamonja (✉)
University of East Anglia, Norwich, UK
e-mail: m.dzamonja@uea.ac.uk

S. Schmitz
IRIF, Université de Paris & CNRS, Paris, France
e-mail: schmitz@irif.fr

Ph. Schnoebelen
LSV, ENS Paris-Saclay & CNRS, Université Paris-Saclay, Cachan, France
e-mail: phs@lsv.fr



P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_2

on cross-purposes and some of the notions re-discovered several times. Therefore we also use the occasion to give a unified presentation of the known results.

**Keywords** WQO · Width of WQO · Ordinal invariants

## 1 Introduction

In the finite case, a partial order—also called a *poset*—$(P, \leq)$ has natural cardinal invariants: a *width*, which is the cardinal of its maximal antichains, and a *height*, which is the cardinal of its maximal chains. The width and height are notably the subject of the theorems of Dilworth [5] and Mirsky [17] respectively; see West [27] for a survey of these *extremal* problems. In the infinite case, cardinal invariants are however less informative—especially for countable posets—while the theorems of Dilworth [5] and Mirsky [17] are well-known to fail [18, 22].

When the poset at hand enjoys additional conditions, the corresponding *ordinal invariants* offer a richer theory, as studied for instance by Kříž and Thomas [9]. Namely, if $(P, \leq)$ has *the finite antichain condition* (FAC), meaning that its antichains are finite, then the tree

$$\mathrm{Inc}(P) \stackrel{\text{def}}{=} \left\{\langle x_0, x_1, \ldots, x_n\rangle \in P^{<\omega} \;:\; 0 \leq n < \omega \wedge \forall 0 \leq i < j \leq n,\; x_i \perp x_j\right\}$$

of all non-empty (finite) sequences of pairwise *incomparable* elements of $P$ ordered by initial segments has no infinite branches. Note that the tree $(\mathrm{Inc}(P), \lhd)$ does not necessarily have a single root and that the empty sequence is excluded (the latter is a matter of aesthetics, but it does make various arguments run more smoothly by not having to consider the case of the empty sequence separately). Therefore, $\mathrm{Inc}(P)$ has a rank, which is the smallest ordinal $\gamma$ such that there is a function $f : \mathrm{Inc}(P) \to \gamma$ with $s \lhd t \implies f(s) > f(t)$ for all $s, t \in \mathrm{Inc}(P)$. This ordinal is called the *width* of $P$ and in this paper we denote it by $\boldsymbol{w}(P)$—it was denoted by $\mathrm{wd}(P)$ by Kříž and Thomas [9].

Similarly, if $(P, \leq)$ is *well-founded* (WF), also called Artinian, meaning that its descending sequences are finite, then the tree

$$\mathrm{Dec}(P) \stackrel{\text{def}}{=} \left\{\langle x_0, x_1, \ldots, x_n\rangle \in P^{<\omega} \;:\; 0 \leq n < \omega \wedge \forall 0 \leq i < j \leq n,\; x_i > x_j\right\}$$

of non-empty strictly descending sequences has an ordinal rank, which we denote by $\boldsymbol{h}(P)$ ([9] denote it by $\mathrm{ht}(P)$) and call the *height* of $P$.

Finally, if $(P, \leq)$ is both well-founded and FAC, i.e., is a *well partial order* (WPO), then the tree

$$\mathrm{Bad}(P) \stackrel{\text{def}}{=} \left\{\langle x_0, x_1, \ldots, x_n\rangle \in P^{<\omega} \;:\; 0 \leq n < \omega \wedge \forall 0 \leq i < j \leq n,\; x_i \not\leq x_j\right\}$$

of non-empty *bad sequences* of $P$ has an ordinal rank, which we denote by $\boldsymbol{o}(P)$ and call the *maximal order type* of $P$ after [8, 23] ([9] denote it by $c(P)$, [4] call it the *stature* of $P$). In the finite case, this invariant is simply the cardinal of the poset.

Quite some work has already been devoted to heights and maximal order types, and to their computation. Widths are however not that well-understood: as Kříž and Thomas [9, Remark 4.14] point out, they do not enjoy nice characterisations like heights and maximal order types do, and the range of available results and techniques on width computations is currently very limited.

Our purpose in this paper is to explore to what extent we can find such a characterisation, and provide formulæ for the behaviour of the width function under various classically defined operations with partial orders. Regarding the first point, we first show in Sect. 3 that the width coincides with the *antichain rank* defined by Abraham and Bonnet [2], which is the height of the chains of antichains; however, unlike the height and maximal order type of WPOs, the width might not be attained (Remark 3.7). Regarding the second point, we first show in Sect. 2.6 that computing widths in the class of FAC orders reduces to computing widths in the class of WPOs. We recall several techniques for computing ordinal invariants, and apply them in Sect. 4 to obtain closed formulæ for the width of sums of posets, and for the finite multisets, finite sequences, and tree extensions of WPOs. One of the main questions is to give a complete formula for the width of the Cartesian products of WPOs. Even the width of the product of two ordinals is only known through a complex recursive formula (due to Abraham, see Sect. 4.4) and we only have partial answers to the general question.

The three ordinal invariants appear in different streams of the literature, often unaware of the results appearing in one another, and using different definitions and notations. Another motivation of this paper is then to provide a unified presentation of the state of the knowledge on the subject, and we also recall the corresponding results for heights and maximal order types as we progress through the paper.

## 2 Background and Basic Results

### *2.1 Posets and Quasi-orders*

We consider posets and, more generally, quasi-orders (QO). When $(Q, \leq_Q)$ is a QO, we write $x <_Q y$ when $x \leq_Q y$ and $y \not\leq_Q x$. We write $x \perp_Q y$ when $x \not\leq_Q y$ and $y \not\leq_Q x$, and say that $a$ and $b$ are *incomparable*. We write $x \equiv_Q y$ when $x \leq_Q y \wedge y \leq_Q x$: this is an equivalence and the quotient $(Q, \leq_Q)/\equiv_Q$ is a poset that, as far as ordinal invariants are concerned, is indistinguishable from $Q$. Therefore we restrict our attention to posets for technical reasons but without any loss of generality. Note that some constructions on posets (e.g., taking powersets) yield quasi-orders that are not posets. A QO $Q$ is *total* if for all $x, y$ in $Q$, $x \leq_Q y$ or $x \geq_Q y$; a total poset is also called a *chain*.

When a QO does not have infinite antichains, we say that it satisfies the *Finite Antichain Condition*, or simply that it is FAC. A QO that does not have any infinite (strictly) decreasing sequence is said to be well-founded (or WF). A *well-quasi order* (or WQO) is a QO that is both WF and FAC: it is well-known that a QO is WQO if and only if it does not have any infinite bad sequence [11, 15], where a sequence $\langle x_0, x_1, x_2, \ldots \rangle$ is *good* if $x_i \leq x_j$ for some positions $i < j$, and is *bad* otherwise.

For a QO $(Q, \leq)$ we define the *reverse* QO $Q^*$ as $(Q, \geq)$, that is to say, $x \leq_{Q^*} y$ if and only if $x \geq_Q y$. An *augmentation* of $(Q, \leq)$ is a QO $(Q, \leq')$ such that $x \leq y \implies x \leq' y$, i.e., $\leq$ is a subset of $\leq'$. A *substructure* of a QO $(Q, \leq)$ is a QO $(Q', \leq')$ such that $Q' \subseteq Q$ and $\leq' \subseteq \leq$. In this case, we write $Q' \leq Q$.

## 2.2 Rankings and Well-Founded Trees

Recall that for every WF poset $P$ there exist ordinals $\gamma$ and order preserving functions $f: P \to \gamma$, that is, such that $x <_P y \implies f(x) < f(y)$ for all $x, y \in P$. The smallest such ordinal $\gamma$ is called the *rank* of $P$; one can obtain the associated *ranking function* $r: P \to \gamma$ by defining inductively $r(x) = \sup\{r(y) + 1 : y <_P x\}$, and the rank turns out to be equal to its height $\boldsymbol{h}(P)$ (see Sect. 2.3). When $P$ is total, i.e., is a chain, then its rank is also called its *order type*.

Traditionally, for a tree $(T, \leq_T)$, one says that it is well-founded if it *does not have an infinite branch*, which with the notation above amounts to saying that the reverse partial order $(T, \geq_T)$ is well-founded. This somewhat confusing notation, implies that for rooted well-founded trees, the root(s) have the largest rank, and the leaves have rank 0. In our definitions of ordinal invariants given in the introduction, we considered trees of non-empty finite sequences, ordered by initial segments: if $s = \langle x_0, x_1, \ldots, x_n \rangle$ and $t = \langle y_0, y_1, \ldots, y_m \rangle$, we write $s \trianglelefteq t$ and say that $s$ is an *initial segment* of $t$, when $n \leq m$ and $s = \langle y_0, \ldots, y_n \rangle$. Equivalently, the associated strict ordering $s \triangleleft t$ means that $t$ can be obtained by appending some sequence $t'$ after $s$, denoted $t = s \frown t'$.

We also make an easy but important observation regarding substructures: When $P$ is embedded in $Q$ as an induced substructure, then $\boldsymbol{w}(P) \leq \boldsymbol{w}(Q)$, and similarly for $\boldsymbol{o}$ and $\boldsymbol{h}$. Indeed, every antichain (bad sequence, decreasing sequence, resp.) of $P$ is an antichain (bad sequence, decreasing sequence, resp.) of $Q$, so the ranks of the corresponding trees can only increase when going from $P$ to $Q$.

## 2.3 Residual Characterisation

For a poset $(P, \leq)$, $x \in P$, and $* \in \{\perp, <, \not\geq\}$, we define the $*$-*residual* of $P$ at $x$ as the induced poset defined by

$$P_{*x} \stackrel{\text{def}}{=} \{y \in P \ : \ y * x\} \ . \tag{1}$$

Since this is an induced substructure of $P$, $P_{*x}$ is FAC (resp. WF, WPO) whenever $P$ is FAC (resp. WF, WPO).

The interest of $\perp$-residuals (resp. $<$-residuals, $\not\geq$-residuals) is that they provide the range of choices for continuing incomparable (resp. descending, bad) sequences once element $x$ has been chosen as first element: the suffix of the sequence should belong to $P_{*x}$, and we have recursively reduced the problem to measuring the rank of the tree $\text{Inc}(P_{\perp x})$ (resp. $\text{Dec}(P_{<x})$, $\text{Bad}(P_{\not\geq x})$).

The following lemma shows precisely how we can extract the rank from such a recursive decomposition of the tree.

**Lemma 2.1**

1. *Suppose that $\{T_i \ : \ i \in I\}$ is a family of well-founded trees and let $T$ be their disjoint union. Then $T$ is a well-founded tree and it has rank $\rho(T) = \sup_{i\in I} \rho(T_i)$.*
2. *Let $T = t^\frown F$ denote a tree rooted at $t$ with $F = T \setminus t$ and suppose that $F$ is well-founded of rank $\rho(F)$. Then so is $T$, and $\rho(T) = \rho(F) + 1$.*

*Proof (of 1)* It is clear that $T$ is well founded. For each $i \in I$, let $f_i \colon T_i \to \rho(T_i)$ be a function witnessing the rank of $T_i$. Then $f \stackrel{\text{def}}{=} \bigcup_{i\in I} f_i$ is an order reversing function from $T$ to $\gamma \stackrel{\text{def}}{=} \sup_{i\in I} \rho(T_i)$, showing $\rho(T) \leq \gamma$.

Conversely, if $f \colon T \to \rho(T)$ is a witness function for the rank of $T$, its restriction to any $T_i$ is order reversing, showing that $\rho(T_i) \leq \rho(T)$.

*Proof (of 2)* Clearly $T$ is well-founded. Let $\rho^* \stackrel{\text{def}}{=} \rho(F) + 1 = \big(\sup_{\alpha<\rho(F)}(\alpha + 1)\big) + 1$. Consider the ranking function $r \colon F \to \rho(F)$, and let $f \colon T \to \rho^*$ be given by

$$f(s) \stackrel{\text{def}}{=} \begin{cases} r(s) & \text{if } s \in F \ , \\ \sup_{\alpha<\rho(F)}(\alpha + 1) & \text{if } s = t. \end{cases}$$

It is clear that $f$ is an order reversing function, witnessing $h(T) \leq \rho^*$. Suppose that $\beta < \rho^*$ and that $h \colon T \to \beta$ is an order reversing function. In particular, $h(r) < f(r)$, so let $\alpha < \rho(F)$ be such that $h(r) < \alpha + 1$. Let $s \in F$ be such that $f(s) = \alpha$. Hence $h(r) \leq h(s)$, yet $r <_T s$, a contradiction. ★2.1

Lemma 2.1 yields the equations:

$$\boldsymbol{w}(P) = \sup_{x\in P}\{\boldsymbol{w}(P_{\perp x}) + 1\} \ , \quad \boldsymbol{h}(P) = \sup_{x\in P}\{\boldsymbol{h}(P_{<x}) + 1\} \ , \quad \boldsymbol{o}(P) = \sup_{x\in P}\{\boldsymbol{o}(P_{\not\geq x}) + 1\} \ , \tag{2}$$

that hold for any FAC, WF, or WPO, poset $P$ respectively. Note that it yields $\boldsymbol{w}(\emptyset) = \boldsymbol{h}(\emptyset) = \boldsymbol{o}(\emptyset) = 0$.

Equation (2) is used very frequently in the literature and provides for a method for computing ordinal invariants recursively, which we call the *method of residuals*.

Equation (2) further shows that the function $r(x) \stackrel{\text{def}}{=} \boldsymbol{h}(P_{<x})$ is the optimal ranking function of $P$. Thus $\boldsymbol{h}(P)$ is the rank of $P$, i.e. the minimal $\gamma$ such that there exists a strict order-preserving $f\colon P \to \gamma$ (recall Sect. 2.2).

## 2.4 Games for WQO Invariants

One limitation of the method of residuals is that it tends to produce recursive rather than closed formulæ, see, e.g., [25]. Another proof technique adopts a game-theoretical point of view. This is based on [4, Sect. 3], which in turn can be seen as an application of a classical game for the rank of trees to the specific trees used for the ordinal invariants. We shall use this technique to obtain results about special products of more than two orders, see for example Theorem 4.18.

The general setting is as follows. For a WQO $P$ and an ordinal $\alpha$, the game $G^{*}_{P,\alpha}$—where $*$ is one of $\boldsymbol{h}, \boldsymbol{o}, \boldsymbol{w}$—is a two-player game where positions are pairs $(\beta, S)$ of an ordinal and a sequence over $P$. We start in the initial position $(\alpha, \langle\rangle)$. At each turn, and in position $(\beta, S)$, Player 1 picks an ordinal $\beta' < \beta$ and Player 2 answers by extending $S$ with an element $x$ from $P$. Player 2 is only allowed to pick $x$ so that the extended $S' = S \frown x$ is a decreasing sequence (or a bad sequence, or an antichain) when $* = \boldsymbol{h}$ (resp. $* = \boldsymbol{o}$, or $* = \boldsymbol{w}$) and he loses the game if he cannot answer Player 1's move. After Player 2's move, the new position is $(\beta', S')$ and the game continues. Player 2 wins when the position has $\beta = 0$ and hence Player 1 has no possible move. The game cannot run forever so one player has a winning strategy. Applying [4, Prop. 23] we deduce that Player 2 wins in $G^{*}_{P,\alpha}$ iff $*(P) \geq \alpha$. As we are mostly interested in the invariant $\boldsymbol{w}$, we shall adopt the notation $G_{P,\alpha}$ for $G^{\boldsymbol{w}}_{P,\alpha}$.

## 2.5 Cardinal Invariants

We can connect the ordinal invariants with cardinal measures but this does not lead to very fine bounds. Here are two examples of what can be said.

**Lemma 2.2** *Suppose that $Q$ is a FAC quasi-order of cardinal $\kappa \geq \aleph_0$. Then $\boldsymbol{w}(Q) < \kappa^{+}$, the cardinal successor of $|Q|$.*

*Proof* The tree $\mathrm{Inc}(Q)$ has size equal to $\kappa$ and therefore its rank is an ordinal $\gamma < \kappa^{+}$. ★2.2

**Theorem 2.3** (Dushnik-Miller) *Suppose that $P$ is a WPO of cardinal $\kappa \geq \aleph_0$. Then $\boldsymbol{h}(P) \geq \kappa$.*

*Proof* This is an easy consequence of Theorem 5.25 in [6]. By the definition of $\boldsymbol{h}$, it suffices to show that $P$ has a chain of size $\kappa$. Define a colouring $c$ on the set $[P]^2$ of pairs of $P$ by saying $c(x, y) \stackrel{\text{def}}{=} 0$ if $x$ is comparable to $y$ and $c(x, y) \stackrel{\text{def}}{=} 1$ otherwise.

Then use the relation $\kappa \longrightarrow (\kappa, \aleph_0)^2$, meaning that $P$ has a chain of cardinal $\kappa$ or an antichain of cardinal $\aleph_0$, which for $\kappa = \aleph_0$ is the Ramsey Theorem, and for $\kappa > \aleph_0$ is the Dushnik-Miller Theorem. Since $P$ is FAC, we must have a chain of order type at least $\kappa$. ★2.3

Such results are however of little help when the poset at hand is countable, because they only tell us that the invariants are countable infinite, as expected. This justifies the use of ordinal invariants rather than cardinal ones.

## *2.6 WPOs as a Basis for FAC Posets*

A *lexicographic sum* of posets in some family $\{P_i : i \in Q\}$ of disjoint orders *along* a poset $(Q, \leq_Q)$, denoted by $\sum_{i \in Q} P_i$, is defined as the order $\leq$ on the disjoint union $P$ of $\{P_i : i \in Q\}$ such that for all $x, y \in P$ we have $x \leq y$ iff $x, y \in P_i$ for some $i \in Q$ and $x \leq_{P_i} y$, or $x \in P_i$ and $y \in P_j$ for some $i, j \in Q$ satisfying $i <_Q j$.

The lexicographic sum of copies of $P$ along $Q$ is denoted by $P \cdot Q$ and called the *direct product* of $P$ and $Q$. The *disjoint sum* of posets in $\{P_i : i \in Q\}$ is defined as the union of the orders $\leq_{P_i}$: this is just a special case of a lexicographic sum, where the sum is taken over an antichain $Q$. In the case of two orders $P_1, P_2$, the lexicographic sum is denoted by $P_1 \sqcup P_2$.

As a consequence of Theorem 7.3 of Abraham et al. [3] (by taking the union over all infinite cardinals $\kappa$), one obtains the following classification theorem.

**Theorem 2.4** ([3]) *Let $\mathscr{BP}$ be the class of posets which are either a WPO, the reverse of a WPO, or a linear order. Let $\mathscr{P}$ be the closure of $\mathscr{BP}$ under lexicographic sums with index set in $\mathscr{BP}$ and augmentation. Then $\mathscr{P}$ is exactly the class of all FAC posets.*

We will use the classification in Theorem 2.4 to see that if we know how to calculate $\boldsymbol{w}(P)$ for $P$ an arbitrary WPO, then we can bound $\boldsymbol{w}(P)$ for any FAC poset $P$. This in fact follows from some simple observations concerning the orders in the class $\mathscr{BP}$.

**Lemma 2.5** *(1) If $P$ is total, then $\boldsymbol{w}(P) = 1$. In general, if all the antichains in a poset $P$ are of length $\leq n$ for some $n < \omega$, then $\boldsymbol{w}(P) \leq n$, and $\boldsymbol{w}(P) = n$ in the case that there are antichains of length $n$.*
*(2) For any poset $P$, $Inc(P) = Inc(P^*)$ and hence in the case of FAC posets we have $\boldsymbol{w}(P^*) = \boldsymbol{w}(P)$.*
*(3) If $P'$ is an augmentation of a FAC poset $P$, then $Inc(P')$ is a subtree of $Inc(P)$ and therefore $\boldsymbol{w}(P') \leq \boldsymbol{w}(P)$.*
*(4) Let $P$ be the lexicographic sum of posets $\{P_i : i \in L\}$ along some linear order $L$. Then $Inc(P) = \bigcup_{i \in L} Inc(P_i)$ and in the case of FAC posets we have $\boldsymbol{w}(P) = \sup_{i \in L} \boldsymbol{w}(P_i)$.*

*Proof* (1) The only non-empty sequences of antichains in a linear order $P$ are the singleton sequences. It is clear that the resulting tree $\mathrm{Inc}(P)$ has rank 1, by assigning the value 0 to any singleton sequence. The more general statement is proved in the same way, namely if all the antichains in a poset $P$ are of length $< n$ for some $n < \omega$ then it suffices to define $f : \mathrm{Inc}(P) \to n$ by letting $f(s) \stackrel{\text{def}}{=} n - |s|$.
(2), (3) Obvious.
(4) This is the same argument as in Lemma 4.1.(3). ★2.5

In conjunction with Theorem 2.4, we conclude that the problem of bounding the width of any given FAC poset is reduced to knowing how to calculate the width of WQO posets. This is the consideration of the second part of this article, starting with Sect. 4.

## 3 Characterisations of Ordinal Invariants

We recall in this section the known characterisations of ordinal invariants. With the method of residuals we can follow [9] and show that the height and maximal order types of WPOs also correspond to their maximal chain heights (Sect. 3.1) and maximal linearisation heights (Sect. 3.2), relying on results of [8, 28] to show that these maxima are indeed attained. In a similar spirit, the width of a FAC poset is equal to its antichain rank (Sect. 3.4), an invariant studied by Abraham and Bonnet [2]—but this time it is not necessarily attained. Finally, in Sect. 3.5 we recall an inequality relating all three invariants and shown by Kříž and Thomas [9].

### *3.1 Height and Maximal Chains*

Given a WF poset $P$, let $\mathscr{C}(P)$ denote its set of non-empty chains. Each chain $C$ from $\mathscr{C}(P)$ is well-founded and has a rank $\boldsymbol{h}(C)$; we denote the supremum of these ranks by $\mathrm{rk}_{\mathscr{C}}\, P \stackrel{\text{def}}{=} \sup_{C \in \mathscr{C}(P)} \boldsymbol{h}(C)$. As explained for example by Kříž and Thomas [9, Theorem 4.9], we have

$$\mathrm{rk}_{\mathscr{C}}\, P \leq \boldsymbol{h}(P) \tag{3}$$

and this can be shown, for instance, by induction on the height using the method of residuals. Indeed, (3) holds when $P = \emptyset$, and for the induction step

$$\sup_{C \in \mathscr{C}(P)} \boldsymbol{h}(C) \stackrel{(2)}{=} \sup_{C \in \mathscr{C}(P)} \sup_{x \in C} (\sup\{\boldsymbol{h}(C_{<x}) + 1\}) \leq \sup_{x \in P}\{( \sup_{C' \in \mathscr{C}(P_{<x})} \boldsymbol{h}(C')) + 1\}$$

because $C_{<x}$ is a chain in $\mathscr{C}(P_{<x})$, and then by induction hypothesis (3)

$$\sup_{C\in\mathscr{C}(P)} \boldsymbol{h}(C) \leq \sup_{x\in P}\{\boldsymbol{h}(P_{<x})+1\} \stackrel{(2)}{=} \boldsymbol{h}(P)\ .$$

*Remark 3.1* The inequality in (3) can be strict. For instance, consider the forest $F$ defined by the disjoint union $\{C_n : n \in \mathbb{N}\}$ along $(\mathbb{N}, =)$, where each $C_n$ is a chain of height $n$, and add a new top element $t$ yielding $P \stackrel{\text{def}}{=} t\frown F$. Then $P$ is WF (but not FAC and is thus not a WPO). Note that $\boldsymbol{h}(P) = \boldsymbol{h}(F) + 1 = \omega + 1$. However, every chain $C$ in $\mathscr{C}(P)$ is included in $t\frown C_n$ for some $n$ and has height bounded by $n + 1$, while $\mathrm{rk}_{\mathscr{C}}(P) = \omega < \boldsymbol{h}(P)$. ★3.1

Wolk [28, Theorem 9] further shows that, when $P$ is a WPO, the supremum is attained, i.e. there is a chain $C$ with rank $\boldsymbol{h}(C) = \mathrm{rk}_{\mathscr{C}} P$. In such a case, (3) can be strengthened to

$$\max_{C\in\mathscr{C}(P)} \boldsymbol{h}(C) = \mathrm{rk}_{\mathscr{C}} P = \boldsymbol{h}(P) \tag{4}$$

as can be checked by well-founded induction with

$$\boldsymbol{h}(P) \stackrel{(2)}{=} \sup_{x\in P}\{\boldsymbol{h}(P_{<x})+1\} \leq \sup_{x\in P}\{\boldsymbol{h}(C_x)+1\} \leq \sup_{x\in P} \boldsymbol{h}(C_x\cup\{x\}) \leq \sup_{C\in\mathscr{C}(P)} \boldsymbol{h}(C)$$

where $C_x$ is a chain of $P_{<x}$ witnessing (4) by induction hypothesis, and $C_x \cup \{x\}$ is therefore a chain in $\mathscr{C}(P)$ of height $\boldsymbol{h}(C_x) + 1$.

**Theorem 3.2** ([9, 28]) *Let $P$ be a WPO. Then $\boldsymbol{h}(P) = \mathrm{rk}_{\mathscr{C}} P = \max_{C\in\mathscr{C}(P)} \boldsymbol{h}(C)$ is the maximal height of the non-empty chains of $P$.*

More generally, the WPO condition in Theorem 3.2 can be relaxed using the following result proven in [16, 19, 24].

**Theorem 3.3** ([16, 19, 24]) *Let $P$ be a WF poset. Then*

- either *$\mathrm{rk}_{\mathscr{C}} P = \max_{C\in\mathscr{C}(P)} \boldsymbol{h}(C)$, i.e. there exist chains of maximal height,*
- or *there exists an antichain $A$ of $P$ such that the set of heights $\{\boldsymbol{h}(P_{<x}) : x \in A\}$ is infinite.*

## 3.2 Maximal Order Types and Linearisations

A *linearisation* of a poset $(P, \leq)$ is an augmentation $L = (P, \preceq)$ which is a total order: $x \leq y$ implies $x \preceq y$. We let $\mathscr{L}(P)$ denote the set of linearisations of $P$. As stated by de Jongh and Parikh [8], a poset is a WPO if and only if all its linearisations are well-founded. de Jongh and Parikh [8] furthermore considered the supremum $\sup_{L\in\mathscr{L}(P)} \boldsymbol{h}(L)$ of the order types of the linearisations of $P$, and showed that this supremum was attained [8, Theorem 2.13]; this is also the subject of [4, Theorem 10].

**Theorem 3.4** ([8, 9]) *Let $Q$ be a WQO. Then $\boldsymbol{o}(Q) = \max_{L\in\mathscr{L}(Q)} \boldsymbol{h}(L)$ is the maximal height of the linearisations of $Q$.*

### 3.3 Maximal Order Types and Height of Downwards-Closed Sets

A subset $D$ of a WQO $(Q, \leq)$ is *downwards-closed* if, for all $y$ in $D$ and $x \leq y$, $x$ also belongs to $D$. We let $\mathscr{D}(Q)$ denote the set of downwards-closed subsets of $Q$. For instance, when $Q = \omega$, $\mathscr{D}(\omega)$ is isomorphic to $\omega + 1$.

It is well-known that a quasi-order $Q$ is WQO if and only if it satisfies the descending chain condition, meaning that $(\mathscr{D}(Q), \subseteq)$ is well-founded. Therefore $\mathscr{D}(Q)$ has a rank $\boldsymbol{h}(\mathscr{D}(Q))$ when $Q$ is WQO. As shown by Blass and Gurevich [4, Prop. 31], this can be compared to the maximal order type of $Q$.

**Theorem 3.5** ([4]) *Let $Q$ be a WQO. Then $\boldsymbol{o}(Q) + 1 = \boldsymbol{h}(\mathscr{D}(Q))$.*

### 3.4 Width and Antichain Rank

Abraham and Bonnet [2] consider a structure similar to the tree $\mathrm{Inc}(P)$ for FAC posets $P$, namely the poset $\mathscr{A}(P)$ of all non-empty antichains of $P$. In the case of a FAC poset, the poset $(\mathscr{A}(P), \supseteq)$ is well-founded. Let us call its height the *antichain rank* of $P$ and denote it by $\mathrm{rk}_{\mathscr{A}} P \stackrel{\text{def}}{=} \boldsymbol{h}(\mathscr{A}(P))$; this is the smallest ordinal $\gamma$ such that there is a strict order-preserving function from $\mathscr{A}(P)$ to $\gamma$.

In fact the antichain rank and the width function we study have the same values, as we now show. Thus one can reason about the width $\boldsymbol{w}(P)$ by looking at the tree $\mathrm{Inc}(P)$ or at $(\mathscr{A}(P), \supseteq)$, a different structure.

**Theorem 3.6** *Let $P$ be a FAC poset. Then $\boldsymbol{w}(P) = rk_{\mathscr{A}} P$.*

*Proof* Let $\gamma = \mathrm{rk}_{\mathscr{A}} P$ and let $r: \mathscr{A}(P) \to \gamma$ be such that $S \supsetneq T \implies r(S) < r(T)$ for all non-empty antichains $S, T$. Define $f: \mathrm{Inc}(P) \to \gamma$ by letting for $s$ non-empty $f(s) \stackrel{\text{def}}{=} r(S)$, where $S$ is the set of elements of $s$. This function satisfies $s \lhd t \implies f(s) > f(t)$ and hence $\boldsymbol{w}(P) \leq \mathrm{rk}_{\mathscr{A}} P$.

Conversely, let $\gamma = \boldsymbol{w}(P)$ and $f: \mathrm{Inc}(P) \to \gamma$ be such that $s \lhd t \implies f(s) > f(t)$. For a non-empty antichain $S \in \mathscr{A}(P)$, observe that there exist finitely many—precisely $|S|!$—sequences $s$ in $\mathrm{Inc}(P)$ with support set $S$. Call this set $\mathrm{Lin}(S)$ and define $r: \mathscr{A}(P) \to \gamma$ by $r(S) \stackrel{\text{def}}{=} \min_{s \in \mathrm{Lin}(S)} f(s)$. Consider now an antichain $S$ with $r(S) = f(s)$ for some $s \in \mathrm{Lin}(S)$, and an antichain $T$ with $T \supsetneq S$: then there exists an extension $t$ of $s$ in $\mathrm{Lin}(T)$, which is therefore such that $f(s) > f(t)$, and hence $r(S) = f(s) > f(t) \geq r(T)$. Thus $\boldsymbol{w}(P) \geq \mathrm{rk}_{\mathscr{A}} P$. ★3.6

*Remark 3.7* The width $\boldsymbol{w}(P)$ is in general not attained, i.e., there might not exist any chain of antichains of height $\boldsymbol{w}(P)$. First note that even when $P$ is a WPO, $(\mathscr{A}(P), \supseteq)$ is in general not a WPO, hence Theorem 3.2 does not apply. In fact, examples of FAC posets where the width is not attained abound. Consider indeed any FAC poset $P$ with $\boldsymbol{w}(P) \geq \omega$, and any non-empty chain $C$ in $\mathscr{C}(\mathscr{A}(P))$. As $C$ is well-founded for $\supseteq$, it has a minimal element, which is an antichain $A \in \mathscr{A}(P)$ such that, for all $A' \neq A$ in $C$, $A' \subsetneq A$. Since $P$ is FAC, $A$ is finite, and $C$ is therefore finite as well: ★3.7

## 3.5 Relationship Between Width, Height and Maximal Order Type

As we have seen in the previous discussion, $\boldsymbol{w}(P) = \boldsymbol{h}(\mathscr{A}(P))$ the antichain rank (where antichains are ordered by reverse inclusion). Kříž and Thomas [9, Theorem 4.13] proved that there is another connection between the ordinal functions discussed here and the width function.

The statement uses natural products of ordinals. Recall for this that the Cantor normal form (CNF) of an ordinal $\alpha$

$$\alpha = \omega^{\alpha_0} \cdot m_0 + \cdots + \omega^{\alpha_\ell} \cdot m_\ell$$

is determined by a non-empty decreasing sequence $\alpha_0 > \alpha_1 \cdots > \alpha_\ell \geq 0$ of ordinals and a sequence of natural numbers $m_i > 0$. Cantor proved that every ordinal has a unique representation in this form. Two well-known operations can be defined based on this representation: the *natural or Hessenberg sum* $\alpha \oplus \beta$ is defined by adding the coefficients of the normal forms of $\alpha$ and $\beta$ as though these were polynomials in $\omega$. The *natural or Hessenberg product* $\alpha \otimes \beta$ is obtained when the normal forms of $\alpha$ and $\beta$ are viewed as polynomials in $\omega$ and multiplied accordingly.

**Theorem 3.8** (Kříž and Thomas) *For any WQO* $(Q, \leq)$ *the following holds:*

$$\boldsymbol{w}(Q) \leq \boldsymbol{o}(Q) \leq \boldsymbol{h}(Q) \otimes \boldsymbol{w}(Q)\,. \tag{5}$$

For completeness, we give a detailed proof.

*Proof* For the first inequality, clearly any antichain in $Q$ can be linearised in an arbitrary way in a linearisation of $Q$. So $\boldsymbol{w}(Q)$ is certainly bounded above by the length of the maximal such linearisation, which by Theorem 3.4 is exactly the value of $\boldsymbol{o}(Q)$.

For the second inequality, let $\alpha = \boldsymbol{w}(Q)$ and let $g : \mathrm{Inc}(Q) \to \alpha$ be a function witnessing that. Also, let $\beta = \boldsymbol{h}(Q)$ and let $\rho :\ Q \to \beta$ be the rank function.

For any bad sequence $\langle q_0, q_1, \ldots, q_n \rangle$ in $Q$ we know that $i < j \leq n$ implies that either $q_i$ is incomparable with $q_j$ or $q_i > q_j$ and hence, in the latter case $\rho(q_i) > \rho(q_j)$. Fixing a bad sequence $s = \langle q_0, q_1, \ldots, q_n \rangle$, consider the set

$$S_s \stackrel{\text{def}}{=} \{\langle q_{i_0}, q_{i_1}, \ldots, q_{i_m} \rangle : i_0 < i_1 \cdots < i_m = n \wedge \rho(q_{i_0}) \leq \rho(q_{i_1}) \cdots \leq \rho(q_{i_m})\}.$$

In other words, $S_s$ consists of subsequences of $s$ that end with $q_n$ and where all elements are incomparable. So for each $t \in S_s$ the value $g(t)$ is defined. We define that $\varphi(s)$ is the minimum over all $g(t)$ for $t \in S_s$. The intuition here is that $\varphi$ is an ordinal measure for the longest incomparable sequence within a bad sequence. Now we are going to combine $\rho$ and $\varphi$ into a function $f$ defined on bad sequences. Given such a sequence $s = \langle q_0, q_1, \ldots, q_n \rangle$, we let

$$f(s) \stackrel{\text{def}}{=} \big\langle \big(\rho(q_0), \varphi(\langle q_0 \rangle)\big), \big(\rho(q_1), \varphi(\langle q_0, q_1 \rangle)\big), \ldots, \big(\rho(q_n), \varphi(\langle q_0, q_1, \ldots, q_n \rangle)\big) \big\rangle.$$

Noticing that every non-empty subsequence of a bad sequence is bad, we see that $f$ is a well-defined function which maps $\mathrm{Bad}(Q)$ into the set of finite sequences from $\alpha \times \beta$. Moreover, let us notice that every sequence in the image of $f$ is a bad sequence in $\alpha \times \beta$: if $i < j$ and $\rho(q_i) \leq \rho(q_j)$, let $t$ be a sequence from $S_{\langle q_0, q_1, q_2, \ldots q_i \rangle}$ such that $g(t) = \varphi(\langle q_0, q_1, q_2, \ldots q_i \rangle)$. Hence $t$ includes $q_i$ and for every $q_k \in t$ we have $\rho(q_k) \leq \rho(q_i) \leq \rho(q_j)$. Therefore $t \frown q_j$ was taken into account when calculating $\varphi(\langle q_0, q_1, q_2, \ldots q_j \rangle)$. In particular,

$$\varphi(\langle q_0, q_1, q_2, \ldots q_j \rangle) \leq g(t \frown q_j) < g(t) = \varphi(\langle q_0, q_1, q_2, \ldots q_i \rangle) \,. \tag{6}$$

Then $(\rho(q_i), \varphi(\langle q_0, q_1, q_2, \ldots q_i \rangle)) \not\leq (\rho(q_j), \varphi(\langle q_0, q_1, q_2, \ldots q_j \rangle))$. Another possibility when $i < j$ is that $\rho(q_i) > \rho(q_j)$ and it yields the same conclusion. We have therefore shown that $f : \mathrm{Bad}(Q) \to \mathrm{Bad}(\alpha \times \beta)$. Let us also convince ourselves that $f$ is a tree homomorphism, meaning a function that preserves the strict tree order. The tree $\mathrm{Bad}(Q)$ is ordered by initial segments, the order which we have denoted by $\lhd$. If $s \lhd t$, then obviously $f(s) \lhd f(t)$. Given that it is well known and easy to see that tree homomorphisms can only increase the rank of a tree, we have that $\boldsymbol{o}(Q) \leq \boldsymbol{o}(\alpha \times \beta)$. The latter, as shown by de Jongh and Parikh [8], is equal to $\alpha \otimes \beta = \boldsymbol{w}(Q) \otimes \boldsymbol{h}(Q)$ (note that $\otimes$ is commutative). ★3.8

From Theorem 3.8 we derive a useful consequence. Recall that $\alpha$ is *additive* (or *multiplicative*) *principal* if $\beta, \gamma < \alpha$ implies $\beta + \gamma < \alpha$ (respectively implies $\beta \cdot \gamma < \alpha$). These implications also hold for natural sums and products.

**Corollary 3.9** *Assume that $\boldsymbol{o}(Q)$ is a principal multiplicative ordinal and that $\boldsymbol{h}(Q) < \boldsymbol{o}(Q)$. Then $\boldsymbol{w}(Q) = \boldsymbol{o}(Q)$.*

*Proof* Assume, by way of contradiction, that $\boldsymbol{w}(Q) < \boldsymbol{o}(Q)$. From $\boldsymbol{h}(Q) < \boldsymbol{o}(Q)$ we deduce $\boldsymbol{h}(Q) \otimes \boldsymbol{w}(Q) < \boldsymbol{o}(Q)$ (since $\boldsymbol{o}(Q)$ is multiplicative principal), contradicting the inequality (5) in Theorem 3.8. Hence $\boldsymbol{w}(Q) \geq \boldsymbol{o}(Q)$, and necessarily $\boldsymbol{w}(Q) = \boldsymbol{o}(Q)$, again by (5). ★3.9

# 4 Computing the Invariants of Common WQOs

We now consider WQOs obtained in various well-known ways and address the question of computing their width, and recall along the way what is known about their height and maximal order type.

In the ideal case, there would be a means of defining well-quasi-orders as the closure of some simple orders, in the 'Hausdorff-like' spirit of Theorem 2.4. Unfortunately, no such result is known and indeed it is unclear which class of orders one could use as a base—for example how would one obtain Rado's example (see Sect. 4.6) from a base of any 'reasonable orders.' Therefore, our study of the width of WQO orders will have to be somewhat pedestrian, concentrating on concrete situations.

## *4.1 Lexicographic Sums*

In the case of lexicographic sums along an ordinal (defined in Sect. 2.6), we have the following result.

**Lemma 4.1** *Suppose that for an ordinal $\alpha$ we have a family of WQOs $\{P_i : i < \alpha\}$. Then $\Sigma_{i<\alpha} P_i$ is a WQO, and:*

1. $\boldsymbol{o}(\Sigma_{i<\alpha} P_i) = \Sigma_{i<\alpha} \boldsymbol{o}(P_i)$,
2. $\boldsymbol{h}(\Sigma_{i<\alpha} P_i) = \Sigma_{i<\alpha} \boldsymbol{h}(P_i)$,
3. $\boldsymbol{w}(\Sigma_{i<\alpha} P_i) = \sup_{i<\alpha} \boldsymbol{w}(P_i)$.

*Proof* First note that any infinite bad sequence in $\Sigma_{i<\alpha} P_i$ would either have an infinite projection to $\alpha$ or an infinite projection to some $P_i$, which is impossible. Hence $\Sigma_{i<\alpha} P_i$ is a WQO. Therefore the values $\boldsymbol{w}(\Sigma_{i<\alpha} P_i)$, $\boldsymbol{o}(\Sigma_{i<\alpha} P_i)$ and $\boldsymbol{h}(\Sigma_{i<\alpha} P_i)$ are well defined.

1. We use Theorem 3.2. Let $\alpha_i \stackrel{\text{def}}{=} \boldsymbol{o}(P_i)$, then $\Sigma_{i<\alpha}\alpha_i$ is isomorphic to a linearisation of $\Sigma_{i<\alpha} P_i$. Hence $\boldsymbol{o}(\Sigma_{i<\alpha} P_i) \geq \Sigma_{i<\alpha} \boldsymbol{o}(P_i)$. Suppose that $L$ is a linearisation of $\Sigma_{i<\alpha} P_i$ (necessarily a well order), then the projection of $L$ to each $P_i$ is a linearisation of $P_i$ and hence it has type $\leq \alpha_i$. This gives that the type of $L$ is $\leq \Sigma_{i<\alpha}\alpha_i$, proving the other side of the desired inequality.
2. We use Theorem 3.4. Any chain $C$ in $\Sigma_{i<\alpha} P_i$ can be obtained as $C = \Sigma_{i<\alpha} C_i$, where $C_i$ is the projection of $C$ on the coordinate $i$. The conclusion follows as in the case of $\boldsymbol{o}$.
3. Every non-empty sequence of incomparable elements in $P$ must come from one and only one $P_i$, hence $\mathrm{Inc}(P) = \bigcup_{i\in L} \mathrm{Inc}(P_i)$, and therefore $\boldsymbol{w}(P_i) = \sup_{i<\alpha} \boldsymbol{w}(P_i)$ by Lemma 2.1. ★4.1

## 4.2 Disjoint Sums

We also defined disjoint sums in Sect. 2.6 as sums along an antichain.

**Lemma 4.2** *Suppose that $P_1, P_2, \ldots$ is a family of WQOs.*

1. $\boldsymbol{o}(P_1 \sqcup P_2) = \boldsymbol{o}(P_1) \oplus \boldsymbol{o}(P_2)$,
2. $\boldsymbol{h}(\bigsqcup_i P_i) = \sup\{\boldsymbol{h}(P_i)\}_i$,
3. $\boldsymbol{w}(P_1 \sqcup P_2) = \boldsymbol{w}(P_1) \oplus \boldsymbol{w}(P_2)$.

*Proof* (1) is Theorem 3.4 from [8].
(2) is clear since, for an arbitrary family $P_i$ of WQOs, $\mathrm{Dec}(\bigcup_i P_i)$ is isomorphic to $\bigsqcup_i \mathrm{Dec}(P_i)$. We observe that, for infinite families, $\bigsqcup_i P_i$ is not WQO, but it is still well-founded hence has a well-defined height.
(3) is Lemma 1.10 from [2] about antichain rank, which translates to widths thanks to Theorem 3.6. ★4.2

We can apply lexicographic sums to obtain the existence of WQO posets of every width.

**Corollary 4.3** *For every ordinal $\alpha$, there is a WQO poset $P_\alpha$ such that $\boldsymbol{w}(P_\alpha) = \alpha$.*

*Proof* The proof is by induction on $\alpha$. For $\alpha$ finite, the conclusion is exemplified by an antichain of length $\alpha$. For $\alpha$ a limit ordinal let us fix for each $\beta < \alpha$ a WPO $P_\beta$ satisfying $\boldsymbol{w}(P_\beta) = \beta$. Then $\boldsymbol{w}(\Sigma_{\beta<\alpha} P_\beta) = \sup_{\beta<\alpha} \beta = \alpha$, as follows by Lemma 4.1. For $\alpha = \beta + 1$, we take $P_\alpha = P_\beta \sqcup 1$, i.e., $P_\beta$ with an extra (incomparable) element added, and rely on $\boldsymbol{w}(Q \sqcup 1) = \boldsymbol{w}(Q) \oplus 1 = \boldsymbol{w}(Q) + 1$ shown in Lemma 4.2. ★4.3

## 4.3 Direct Products

Direct products are again a particular case of lexicographic sums along a poset $Q$, this time of the same poset $P$. While the cases of $\boldsymbol{o}$ and $\boldsymbol{h}$ are mostly folklore, the width of $P \cdot Q$ is not so easily understood, and its computation in Lemma 1.11 from [2] uses the notion of *Heisenberg products* $\alpha \odot \beta$, defined for any ordinal $\alpha$ by induction on the ordinal $\beta$:

$$\alpha \odot 0 \stackrel{\text{def}}{=} 0\,, \quad \alpha \odot (\beta + 1) \stackrel{\text{def}}{=} (\alpha \odot \beta) \oplus \alpha\,, \quad \alpha \odot \lambda \stackrel{\text{def}}{=} \sup\{(\alpha \odot \gamma) + 1 : \gamma < \lambda\}$$

where $\lambda$ is a limit ordinal. Note that this differs from the natural product, and is not commutative: $2 \odot \omega = \omega$ but $\omega \odot 2 = \omega \cdot 2$.

**Lemma 4.4** ([2]) *Suppose that $P$ and $Q$ are two WPOs.*

1. $\boldsymbol{o}(P \cdot Q) = \boldsymbol{o}(P) \cdot \boldsymbol{o}(Q)$,
2. $\boldsymbol{h}(P \cdot Q) = \boldsymbol{h}(P) \cdot \boldsymbol{h}(Q)$,
3. $\boldsymbol{w}(P \cdot Q) = \boldsymbol{w}(P) \odot \boldsymbol{w}(Q)$.

### *4.4 Cartesian Products*

The next simplest operation on WQOs is their Cartesian product. It turns out that the simplicity of the operation is deceptive and that the height and, especially, the width of a product $P \times Q$ are not as simple as we would like. As a consequence, this section only provides partial results and is unexpectedly long.

To recall, the product order $P \times Q$ of two partial orders is defined on the pairs $(p, q)$ with $p \in P$ and $q \in Q$ so that $(p, q) \leq (p', q')$ iff $p \leq_P p'$ and $q \leq_Q q'$. It is easy to check, and well known, that product of WQOs is WQO and similarly for FAC and WF orders.

The formula for calculating $\boldsymbol{o}(P \times Q)$ is still simple. It was first established by de Jongh and Parikh [8, Theorem 3.5]; see also [4, Theorem 6].

**Lemma 4.5** ([8]) *Suppose that P and Q are two WQOs. Then* $\boldsymbol{o}(P \times Q) = \boldsymbol{o}(P) \otimes \boldsymbol{o}(Q)$.

The question of the height of products is also well studied and a complete answer appears in [1], where it is stated that the theorem is well known. The following statement is a reformulation of Lemma 1.8 of [1].

**Lemma 4.6** (Abraham; folklore) *If* $\rho_P : P \to \boldsymbol{h}(P)$ *and* $\rho_Q : Q \to \boldsymbol{h}(Q)$ *are the rank functions of the well-founded posets P and Q, then the rank function* $\rho$ *on* $P \times Q$ *is given by* $\rho(x, y) = \rho_P(x) \oplus \rho_Q(y)$. *In particular,*

$$\boldsymbol{h}(P \times Q) = \sup\{\alpha \oplus \beta + 1 : \alpha < \boldsymbol{h}(P) \wedge \beta < \boldsymbol{h}(Q)\} .$$

We recall that for any two ordinals $\alpha$ and $\beta$ we have $\sup_{\alpha' < \alpha, \beta' < \beta} \alpha' \oplus \beta' + 1 < \alpha \oplus \beta$ (see e.g. [2], p. 55), thus the statement in Theorem 4.6 cannot be easily simplified.

*Remark 4.7* (*Height of products of finite ordinals*) The very nice general proof of [1, Lemma 1.8] can be done in an even more visual way in the case of finite ordinals. Let $P = n_1 \times \cdots \times n_k$ for some finite $n_1, \ldots, n_k \in \omega$; then $\boldsymbol{h}(P) = n_1 + \cdots + n_k + 1 - k$.

Indeed, we observe that any chain $\mathbf{a}_1 <_P \cdots <_P \mathbf{a}_\ell$ in $P$ leads to a strictly increasing $|\mathbf{a}_1| < \cdots < |\mathbf{a}_\ell|$, where by $|\mathbf{a}|$ we denote the sum of the numbers in $\mathbf{a}$. Since $|\mathbf{a}_\ell|$ is at most $\sum_i (n_i - 1) = (\sum_i n_i) - k$ and since $|\mathbf{a}_1|$ is at least 0, the longest chain has length $1 + \sum_i n_i - k$. Furthermore it is easy to build a witness for this length. We conclude by invoking Theorem 3.4 which states that for any WPO $P$, $\boldsymbol{h}(P)$ is the length of the longest chain in $P$. ★4.7

Having dealt with $\boldsymbol{h}$ and $\boldsymbol{o}$, we are left with $\boldsymbol{w}$. Here we cannot hope to have a uniform formula expressing $\boldsymbol{w}(P \times Q)$ as a function of $\boldsymbol{w}(P)$ and $\boldsymbol{w}(Q)$. Indeed, already in the case of ordinals one always has $\boldsymbol{w}(\alpha) = \boldsymbol{w}(\beta) = 1$, while $\boldsymbol{w}(\alpha \times \beta)$ has quite a complex form, as we are going to see next.

### 4.4.1 Products of Ordinals

Probably the simplest example of WQO which is not actually an ordinal, is provided by the product of two ordinals. Thanks to Theorem 3.6, we can translate results of [1], Sect. 3 to give a recursive formula which completely characterises $\boldsymbol{w}(\alpha \times \beta)$ for $\alpha, \beta$ ordinals. We shall sketch how this is done.

First note that if one of $\alpha, \beta$ is a finite ordinal $n$, say $\alpha = n$, then we have $\boldsymbol{w}(n \times \beta) = \min\{n, \beta\}$. The next case to consider is that of successor ordinals, which is taken care by the following Theorem 4.8. Abraham proved this theorem using the method of residuals [11] and induction, we offer an alternative proof using the rank of the tree Inc.

**Theorem 4.8** (Abraham) *For any ordinals* $\alpha, \beta$ *with* $\alpha$ *infinite, we have* $\boldsymbol{w}(\alpha \times (\beta + 1)) = \boldsymbol{w}(\alpha \times \beta) + 1$.

The proof is provided by the next two Lemmas.

**Lemma 4.9** $\boldsymbol{w}(\alpha \times (\beta + 1)) \leq \boldsymbol{w}(\alpha \times \beta) + 1$ *for any ordinals* $\alpha, \beta$.

*Proof* Write $I$ for $\mathrm{Inc}(\alpha \times (\beta + 1))$ and $I'$ for $\mathrm{Inc}(\alpha \times \beta)$. Any sequence $s = \langle p_1, \ldots, p_\ell \rangle$ which is in $I$, is either in $I'$ or contains a single pair of the form $p_i = (a, \beta)$, with $a < \alpha$. In the latter case we write $s'$ for $s$ with $p_i$ removed. Note that $s'$ is in $I'$ (except when $s$ has length 1). Let $\rho' : I' \to \mathrm{rank}(I') = \boldsymbol{w}(\alpha \times \beta)$ be a ranking function for $I'$ and define $\rho : I \to ON$ via

$$\rho(s) \stackrel{\text{def}}{=} \begin{cases} \rho'(s) + 1 & \text{if } s \in I', \\ \rho'(s') & \text{if } s \notin I' \text{ and } |s| > 1, \\ \mathrm{rank}(I') & \text{otherwise.} \end{cases}$$

One easily checks that $\rho$ is anti-monotone. For this assume $s \lhd t$: (1) if both $s$ and $t$ are in $I'$, monotonicity is inherited from $\rho'$; (2) if none are in $I'$ then $s' \lhd t'$ (or $s'$ is empty) and again monotonicity is inherited (or $\rho(s) = \mathrm{rank}(I') > \rho'(t') = \rho(t)$); (3) if $s$ is in $I'$ and $t$ is not then $s \unlhd t'$, entailing $\rho'(s) \geq \rho'(t')$ so that $\rho(s) = \rho'(s) + 1 > \rho'(t') = \rho(t)$.

In conclusion $\rho$, having values in $\boldsymbol{w}(\alpha \times \beta) + 1$, witnesses the assertion of the lemma. ★3.9

**Lemma 4.10** *If* $\alpha$ *is infinite then* $\boldsymbol{w}(\alpha \times (\beta + 1)) \geq \boldsymbol{w}(\alpha \times \beta) + 1$ *for any* $\beta$.

*Proof* Write $I$ for $\mathrm{Inc}(\alpha \times \beta)$. Any $s \in I$ has the form $s = \langle (a_1, b_1), \ldots, (a_\ell, b_\ell) \rangle$. We write $s_+$ for the sequence $\langle (a_1 + 1, b_1), \ldots, (a_\ell + 1, b_\ell) \rangle$ and observe that it is still a sequence over $\alpha \times \beta$ since $\alpha$ is infinite, and that its elements form an antichain (since the elements of $s$ did). Let now $s'_+$ be $r \frown s_+$ where $r = \langle (0, \beta) \rangle$: the prepended element is not comparable with any element of $s_+$ so that $s'_+$ is an antichain and $s'_+ \unlhd t'_+$ iff $s_+ \unlhd t_+$ iff $s \unlhd t$. Write $I'_+$ for $\{s'_+ \mid s \in I\} \cup \{r\}$. This is a tree made of a root glued below a tree isomorphic to $I$. Hence $\mathrm{rank}(I'_+) = \mathrm{rank}(I) + 1$. On the other hand, $I'_+$ is a substructure of $\mathrm{Inc}(\alpha \times (\beta + 1))$ hence $\boldsymbol{w}(\alpha \times (\beta + 1)) \geq \mathrm{rank}(I'_+)$. ★4.10

With Theorem 4.8 in hand, the remaining case is to compute $\boldsymbol{w}(\alpha \times \beta)$ when $\alpha, \beta$ are limit ordinals. This translates into saying that $\alpha = \omega\alpha'$ and $\beta = \omega\beta'$ for some $\alpha', \beta' > 0$. A recursive formula describing the weight of this product is the main theorem of Sect. 3 of [1], which we now quote. It is proved using a complex application of the method of residuals and induction.

**Theorem 4.11** (Abraham) *Suppose that $\alpha$ and $\beta$ are given in their Cantor normal forms $\alpha = \omega^{\alpha_0} \cdot m_0 + \rho$, $\beta = \omega^{\beta_0} \cdot n_0 + \sigma$, where $\omega^{\alpha_0} \cdot m_0$ and $\omega^{\beta_0} \cdot n_0$ are the leading terms and $\rho$ and $\sigma$ are the remaining terms of the Cantor normal forms of $\alpha$ and $\beta$ respectively. Then if $\alpha = 1$, we have $\boldsymbol{w}(\omega \times \omega\beta) = \omega\beta$, and in general*

$$\boldsymbol{w}(\omega\alpha \times \omega\beta) = \omega\omega^{\alpha_0 \oplus \beta_0} \cdot (m_0 + n_0 - 1) \oplus \boldsymbol{w}(\omega\omega^{\alpha_0} \times \omega\sigma) \oplus \boldsymbol{w}(\omega\omega^{\beta_0} \times \omega\rho).$$

It would be interesting to have a closed rather than a recursive formula for the width of the product of two ordinals. However, the formula does give us a closed form of values of the weight of the product of two ordinals with only one term in the Cantor normal form, as we now remark. Here $m, n$ are finite ordinals $\geq 1$.

1. If $k, \ell < \omega$ then we have
$$\boldsymbol{w}(\omega^{1+k} \cdot m \times \omega^{1+\ell} \cdot n) = \boldsymbol{w}\big(\omega(\omega^k \cdot m) \times \omega(\omega^\ell \cdot n)\big) = \omega^{k+\ell-1} \cdot (m+n-1)\,.$$
2. (example 3.4 (3) from [1]) If $\alpha, \beta \geq \omega$ then $1 + \alpha = \alpha$ and $1 + \beta = \beta$, so
$$\boldsymbol{w}(\omega^\alpha \cdot m \times \omega^\beta \cdot n) = \boldsymbol{w}\big(\omega(\omega^\alpha \cdot m) \times \omega(\omega^\beta \cdot n)\big) = \omega^{\alpha \oplus \beta} \cdot (m+n-1)\,.$$
3. If $\alpha \geq \omega$ and $k < \omega$ then $\boldsymbol{w}(\omega^\alpha \cdot m \times \omega^{1+k} \cdot n) = \omega^{\alpha+k} \cdot (m+n-1)$.

Let us mention one more result derivable from Theorem 4.11.

**Lemma 4.12** (Abraham) $\boldsymbol{w}(\omega \times \alpha) = \alpha$ *for any ordinal* $\alpha$.

*Proof* By induction on $\alpha$. If $\alpha$ is a limit, we write it $\alpha = \omega\alpha' = \omega(\omega^{\alpha_0} \cdot m_0 + \cdots + \omega^{\alpha_\ell} \cdot m_\ell)$. Now Theorem 4.11 yields $\boldsymbol{w}(\omega \times \omega\alpha') = \omega\omega^{\alpha_0} \cdot m_0 \oplus \cdots \oplus \omega\omega^{\alpha_\ell} \cdot m_\ell = \alpha$. If $\alpha$ is a successor, we use Lemmas 4.9 and 4.10. ★4.12

### 4.4.2 Finite Products and Transferable Orders

Since the width of the product of two ordinals is understood, we can approach the general question of the width of products of two or a finite number of WQO posets $P_i$ by reducing it to the width of some product of ordinals. Using that strategy, we give a lower bound to $\boldsymbol{w}(\Pi_{i \leq n} P_i)$.

**Theorem 4.13** *For any WQO posets $P_0, P_1 \dots P_n$, $\boldsymbol{w}(\prod_{i \leq n} P_i) \geq \boldsymbol{w}(\prod_{i \leq n} \boldsymbol{h}(P_i))$.*

The proof follows directly from a simple lemma, which is of independent interest:

**Lemma 4.14** *Suppose that $P_0$, $P_1 \dots P_n$ are WQO posets. Then $\prod_{i \le n} \boldsymbol{h}(P_i)$ embeds into $\prod_{i \le n} P_i$ as a substructure.*

*Proof* We use Theorem 3.2 and pick, in each $P_i$, a chain $C_i$ in $P_i$ that has order type $\boldsymbol{h}(C_i) = \boldsymbol{h}(P_i)$. Then $\prod_{i \le n} C_i$ is an induced suborder of $\prod_{i \le n} \boldsymbol{h}(P_i)$ which is isomorphic to $\prod_{i \le n} \boldsymbol{h}(P_i)$. ★4.14

Now we shall isolate a special class of orders for which it will be possible to calculate certain widths of products. Let us write $\downarrow x$ for the downwards-closure of an element $x$, i.e., for $\{y : x \le y\}$.

**Definition 4.15** A FAC partial order $P$ belongs to the class $\mathscr{T}$ of *transferable orders* if $\boldsymbol{w}(P \setminus (\downarrow x_1 \cup \dots \cup \downarrow x_n)) = \boldsymbol{w}(P)$ for any (finitely many) elements $x_1, \dots, x_n \in P$.

**Theorem 4.16** *Suppose that $P$ is a WQO transferable poset and $\delta$ is an ordinal. Then $\boldsymbol{w}(P \times \delta) \ge \boldsymbol{w}(P) \cdot \delta$.*

*Proof* Write $\gamma$ for $\boldsymbol{w}(P)$: we prove that Player 2 has a winning strategy, denoted $\sigma_{P' \times \delta, \alpha}$, for each game $G_{P' \times \delta, \alpha}$ where $P'$ is some $P \setminus (\downarrow y_1 \cup \dots \cup \downarrow y_n)$ and $\alpha \le \gamma \cdot \delta$.

The proof is by induction on $\delta$.

If $\delta = 0$ then $\alpha = 0$ and Player 1 loses immediately.

If $\delta = \lambda$ is a limit, the strategy for Player 2 depends on Player 1's first move. Say it is $\alpha' < \alpha \le \gamma \cdot \delta$. Then $\alpha' < \gamma \cdot \delta$ means that $\alpha' < \gamma \cdot \delta'$ for some $\delta' < \delta$. Player 2 chooses one such $\delta'$ and now applies $\sigma_{P' \times \delta', \alpha' + 1}$ (which exists and is winning by the induction hypothesis) for the whole game. Note that a strategy for a substructure $P' \times \delta'$ of the original $P' \times \delta$ will lead to moves that are legal in the original game. Also note that $\alpha' + 1$ is $\le \gamma \cdot \delta'$.

If $\delta = \varepsilon + 1$ is a successor then Player 2 answers each move $\alpha_1, \dots, \alpha_m$ played by Player 1 by writing it in the form $\alpha_i = \gamma \cdot \delta_i + \beta_i$ with $\beta_i < \gamma$. Note that $\delta_i < \delta$. If $\delta_1 = \dots = \delta_m = \varepsilon$, note that $\beta_1 > \beta_2 > \dots \beta_m$. Let Player 2 play $(x_m, \varepsilon)$ where $x_m$ is $\sigma_{P', \gamma}$ applied on $\beta_1, \dots, \beta_m$ (that strategy exists and is winning since $P$ is transferable and has width $\gamma$). If $\delta_m < \varepsilon$ then Player 2 switches strategy and now uses $\sigma_{P'' \times \varepsilon, \gamma \cdot \varepsilon}$ as if a new game was starting with $\alpha_m$ as Player 1's first more, and for $P'' = P' \setminus (\downarrow x_1 \cup \dots \cup \downarrow x_{m-1})$. By the induction hypothesis , Player 2 will win by producing a sequence $S''$ in $P'' \times \varepsilon$. These moves are legal since $(x_1, \varepsilon) \cdots (x_{m-1}, \varepsilon) \frown S''$ is an antichain in $P' \times (\varepsilon + 1)$. ★4.16

In order to use Theorem 4.16, we need actual instances of transferable orders.

**Lemma 4.17** *For any $1 \le \alpha_1, \dots, \alpha_n$, the order $P = \omega^{\alpha_1} \times \dots \times \omega^{\alpha_n}$ is transferable.*

*Proof* Since each $\omega^{\alpha_i}$ is additive principal, $P \setminus (\downarrow x_1 \cup \dots \cup \downarrow x_m)$ contains an isomorphic copy of $P$ for any finite sequence $x_1, \dots, x_m$ of elements of $P$. ★4.17

**Theorem 4.18** *Let $P$ be a transferable WPO poset.*

1. *Suppose that $1 \le m < \omega$. Then $\boldsymbol{w}(P) \cdot m \le \boldsymbol{w}(P \times m) \le \boldsymbol{w}(P) \otimes m$.*
2. *If $\boldsymbol{w}(P) = \omega^\gamma$ for some $\gamma$, then $\boldsymbol{w}(P \times m) = \boldsymbol{w}(P) \cdot m$ (Note that this applies to any $P$ which is the product of the form $\omega^\alpha \times \omega^\beta$, see the examples after Theorem 4.11).*
3. *$\boldsymbol{w}(\omega \times \omega \times \omega) = \omega^2$.*

An easy way to provide an upper bound needed in the proof of Theorem 4.18 is given by the following observation:

**Lemma 4.19** *For any FAC poset $P$ and $1 \le m < \omega$, $\boldsymbol{w}(P \times m) \le \boldsymbol{w}(P) \otimes m$.*

*Proof* We just need to remark that $P \times m$ is an augmentation of the perpendicular sum $\bigsqcup_{i<m} P$ and then apply Lemma 4.2. ★4.19

*Proof* (*of Theorem* 4.18) (1) We get $\boldsymbol{w}(P \times m) \ge \boldsymbol{w}(P) \cdot m$ from Theorem 4.16. We get $\boldsymbol{w}(P \times m) \le \boldsymbol{w}(P) \otimes m$ from Lemma 4.19.
(2) This follows because $\omega^\gamma \otimes m = \omega^\gamma \cdot m$.
(3) Let $P = \omega \times \omega$, hence we know that $\boldsymbol{w}(P) = \omega$. Since any $P \times m$ is a substructure of $P \times \omega$, we clearly have that $\boldsymbol{w}(P \times \omega) \ge \sup_{m<\omega} \boldsymbol{w}(P \times m) = \sup_{m<\omega} \omega \cdot m = \omega^2$. Let us now give a proof using games that $\boldsymbol{w}(P \times \omega) \le \omega^2$. It suffices to give a winning strategy to Player 1 in the game $G_{P\times\omega,\gamma}$ for any ordinal $\gamma > \omega^2$.

So, given such a $\gamma$, Player 1 starts the game by choosing as his first move the ordinal $\omega^2$. Player 2 has to answer by choosing an element $x$ in $P \times \omega$, say an element $(p, m)$ with $p = (k, \ell)$. Now notice that any element of $P \times \omega$ that is incompatible with $(p, m)$ is either an element of $P \times m$ or of the form $(q, n)$ for some $q \le p$ in $\omega \times \omega$, or is of the form $(r, i)$ for some $r$ which is incompatible with $p$ in $\omega \times \omega$. Therefore, any next step of Player 2 has to be in an order $P'$ which is isomorphic to an augmentation of a substructure of the disjoint union of the form

$$P \times m \sqcup [(k+1) \times (\ell+1)] \times \omega \sqcup [(k+1) \times \omega] \times \omega \sqcup [(\ell+1) \times \omega] \times \omega. \quad (7)$$

It now suffices for Player 1 to find an ordinal $o < \omega^2$ satisfying $o > \boldsymbol{w}(P')$ as the game will then be transferred to $G_{P',o}$, where Player 1 has a winning strategy. As $\omega^2$ is closed under $\oplus$, it suffices to show that each of the orders appearing in Eq. (7) has weight $<\omega^2$. This is the case for $P \times m$ by (2). We have that $\boldsymbol{w}\big([(k+1) \times (\ell+1)] \times \omega\big) = \boldsymbol{w}\big((k+1) \times [(\ell+1) \times \omega]\big)$, which by applying Lemma 4.19 is $\le (\ell+1) \cdot (k+1)$. For $[(k+1) \times \omega] \times \omega$, we apply Lemma 4.19 to $\omega \times \omega$, to obtain $\boldsymbol{w}\big([(k+1) \times \omega] \times \omega\big) \le \omega \cdot (k+1)$ and similarly $\boldsymbol{w}\big([(\ell+1) \times \omega] \times \omega\big) \le \omega \cdot (\ell+1)$. ★4.18

## *4.5 Finite Multisets, Sequences, and Trees*

Well-quasi-orders are also preserved by building multisets, sequences, and trees with WQO labels, together with suitable embedding relations.

*Finite sequences* in $Q^{<\omega}$ are compared by the *subsequence embedding* ordering defined by $s = \langle x_0, \ldots, x_{n-1}\rangle \leq_* s' = \langle x'_0, \ldots, x'_{p-1}\rangle$ if there exists $f : n \to p$ strictly monotone such that $x_i \leq x'_{f(i)}$ in $Q$ for all $i \in n$. The fact that $(Q^{<\omega}, \leq_*)$ is WQO when $Q$ is WQO was first shown by Higman [7].

Given a WQO $(Q, \leq)$, a *finite multiset* over $Q$ is a function $m$ from $Q \to \mathbb{N}$ with finite support, i.e. $m(x) > 0$ for finitely many $x \in Q$. Equivalently, a finite multiset is a finite sequence $m$ in $Q^{<\omega}$ where the order is irrelevant, and can be noted as a 'set with repetitions' $m = \{x_1, \ldots, x_n\}$; we denote by $M(Q)$ the set of finite multisets over $Q$. The *multiset embedding* ordering is then defined by $m = \{x_0, \ldots, x_{n-1}\} \leq_\diamond m' = \{x'_0, \ldots, x'_{p-1}\}$ if there exists an injective function $f : n \to p$ with $x_i \leq x'_{f(i)}$ in $Q$ for all $i \in n$. As a consequence of $(Q^{<\omega}, \leq_*)$ being WQO, $(M(Q), \leq_\diamond)$ is also WQO when $Q$ is.

Finally, a (rooted, ordered) *finite tree* $t$ over $Q$ is either a leaf $x()$ for some $x \in Q$, or a term $x(t_1, \ldots, t_n)$ for some $n > 0$, $x \in Q$, and $t_1, \ldots, t_n$ trees over $Q$. A tree has arity $b$ if we bound $n$ by $b$ in this definition. We let $T(Q)$ denote the set of finite trees over $Q$. The *homeomorphic tree embedding* ordering is defined by $t = x(t_1, \ldots, t_n) \leq_T t' = x'(t'_1, \ldots, t'_p)$ (where $n, p \geq 0$) if at least one the following cases occurs:

- $t \leq_T t'_j$ for some $1 \leq j \leq p$, or
- $x \leq x'$ in $Q$ and $t_1 \cdots t_n \leq_* t'_1 \cdots t'_p$ for the subsequence embedding relation on $T(Q)$.

The fact that $(T(Q), \leq_T)$ is WQO when $Q$ is WQO was first shown by [7] for trees of bounded arity, before [10] proved it in the general case. Note that it implies $(Q^{<\omega}, \leq_*)$ being WQO for the special case of trees of arity 1.

### 4.5.1 Maximal Order Types

The maximal order types of $M(Q)$, $Q^{<\omega}$, and $T(Q)$ have been studied by Weiermann [26] and Schmidt [23]; see also [13, Sect. 1.2] for a nice exposition of these results.

For finite multisets with embedding, we need some additional notations. For an ordinal $\alpha$ with Cantor normal form $\omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$ where $\boldsymbol{o}(P) \geq \alpha_1 \geq \ldots \geq \alpha_n$, we let

$$\widehat{\alpha} \stackrel{\text{def}}{=} \omega^{\alpha_1{}'} + \cdots + \omega^{\alpha_n{}'} \tag{8}$$

where $\alpha'$ is $\alpha + 1$ when $\alpha$ is an epsilon number, i.e. when $\omega^\alpha = \alpha$, and is just $\alpha$ otherwise.

The following is [26, Theorem 2], with a corrected proof due to [14, Theorem 5].

**Theorem 4.20** ([26]) *Let $Q$ be a WQO. Then $\boldsymbol{o}(M(Q)) = \omega^{\widehat{\boldsymbol{o}(Q)}}$.*

Thus, for $\boldsymbol{o}(Q) < \varepsilon_0$, one has simply $\boldsymbol{o}(M(Q)) = \omega^{\boldsymbol{o}(Q)}$.

For finite sequences with subsequence embedding, we recall the following result by Schmidt [23].

**Theorem 4.21** ([23]) *Let $Q$ be a WQO. Then*

$$\boldsymbol{o}(Q^{<\omega}) = \begin{cases} \omega^{\omega^{\boldsymbol{o}(Q)-1}} & \text{if } \boldsymbol{o}(Q) \text{ is finite,} \\ \omega^{\omega^{\boldsymbol{o}(Q)+1}} & \text{if } \boldsymbol{o}(Q) = \varepsilon + n \text{ for } \varepsilon \text{ an epsilon number and } n \text{ finite,} \\ \omega^{\omega^{\boldsymbol{o}(Q)}} & \text{otherwise.} \end{cases}$$

The case of finite trees is actually a particular case of the results of [23] on embeddings in structured trees. Her results were originally stated using Schütte's Klammer symbols, but can be translated in terms of the $\vartheta$ functions of [21]. Defining such ordinal notation systems is beyond the scope of this chapter; it suffices to say for our results that the ordinals at hand are going to be principal multiplicative.

**Theorem 4.22** ([23]) *Let $Q$ be a WQO. Then $o(T(Q)) = \vartheta(\Omega^\omega \cdot \boldsymbol{o}(Q))$.*

### 4.5.2 Heights

For a WQO $Q$ we define $\boldsymbol{h}^*(Q)$ as

$$\boldsymbol{h}^*(Q) \stackrel{\text{def}}{=} \begin{cases} \boldsymbol{h}(Q) & \text{if } \boldsymbol{h}(Q) \text{ is additive principal } \geq \omega, \\ \boldsymbol{h}(Q) \cdot \omega & \text{otherwise.} \end{cases} \tag{9}$$

We are going to show that the heights of finite multisets, finite sequences, and finite trees over $Q$ is the same, namely $\boldsymbol{h}^*(Q)$.

**Theorem 4.23** *Let $Q$ be a WF poset. Then $\boldsymbol{h}(M(Q)) = \boldsymbol{h}(Q^{<\omega}) = \boldsymbol{h}(T(Q)) = \boldsymbol{h}^*(Q)$.*

Since obviously $\boldsymbol{h}(M(Q)) \leq \boldsymbol{h}(Q^{<\omega}) \leq \boldsymbol{h}(T(Q))$, the claim is a consequence of Lemmas 4.24 and 4.26 below.

**Lemma 4.24** $\boldsymbol{h}(T(Q)) \leq \boldsymbol{h}^*(Q)$.

*Proof* Consider a strictly decreasing sequence $x_0 >_T x_1 >_T \ldots$ in $T(Q)$, where each $x_i$ is a finite tree over $Q$. Necessarily these finite trees have a nonincreasing number of nodes: $|x_0| \geq |x_1| \geq \ldots$. If we add a new minimal element $\perp$ below $Q$, we can transform any $x_i$ by padding it with some $\perp$'s so that now the resulting $x_i'$ has the same shape and size as $x_0$. Let us use $1 + Q$ instead of $\{\perp\} + Q$ so that the new trees belong to $T(1 + Q)$, have all the same shape, and form a strictly decreasing sequence. This construction is in fact an order-reflection from $\text{Dec}(T(Q))$ to $\text{Dec}\big(\bigsqcup_{n<\omega}(1 + Q)^n\big)$, from which we get

$$\boldsymbol{h}(T(Q)) \leq \boldsymbol{h}\Big(\bigsqcup_{n<\omega}(1 + Q)^n\Big) = \sup_{n<\omega} \boldsymbol{h}([1 + Q]^n)\ , \tag{10}$$

using Lemma 4.2.(2) for the last equality. For $n < \omega$, one has

$$\boldsymbol{h}([1+Q]^n) = \sup\{(\alpha \otimes n) + 1 \;:\; \alpha < 1 + \boldsymbol{h}(Q)\}\,, \tag{11}$$

using Lemmas 4.1.(2) and 4.6.

If $\boldsymbol{h}(Q) \leq 1$, $\boldsymbol{h}(T(Q)) = \boldsymbol{h}(Q) \cdot \omega = \boldsymbol{h}^*(Q)$ obviously.

For $\boldsymbol{h}(Q) > 1$, and thanks to (10) and (11), it is sufficient to show that $\alpha \otimes n + 1 \leq \boldsymbol{h}^*(Q)$ for all $n < \omega$ and all $\alpha < 1 + \boldsymbol{h}(Q)$. We consider two cases:

1. If $\boldsymbol{h}(Q) \geq \omega$ is additive principal, $\alpha < 1 + \boldsymbol{h}(Q) = \boldsymbol{h}(Q)$ entails $\alpha \otimes n < \boldsymbol{h}(Q)$ thus $\alpha \otimes n + 1 < \boldsymbol{h}(Q) = \boldsymbol{h}^*(Q)$.
2. Otherwise the CNF for $\boldsymbol{h}(Q)$ is $\sum_{i=1}^m \omega^{\alpha_i}$ with $m > 1$. Then $\alpha < 1 + \boldsymbol{h}(Q)$ implies $\alpha \leq \omega^{\alpha_1} \cdot m$, thus $\alpha \otimes n + 1 \leq \omega^{\alpha_1} \cdot m \cdot n + 1 \leq \omega^{\alpha_1 + 1} = \boldsymbol{h}(Q) \cdot \omega = \boldsymbol{h}^*(Q)$. ★4.18

Let us write $M_n(Q)$ for the restriction of $M(Q)$ to multisets of size $n$.

**Lemma 4.25** $\boldsymbol{h}(M_n(Q)) \geq \boldsymbol{h}(Q^n)$.

*Proof* With $\boldsymbol{x} = \langle x_1, \ldots, x_n \rangle \in Q^n$ we associate the multiset $M_{\boldsymbol{x}} = \{x_1, \ldots, x_n\}$. Obviously $\boldsymbol{x} <_\times \boldsymbol{y}$ implies $M_{\boldsymbol{x}} \leq_\diamond M_{\boldsymbol{y}}$. We further claim that $M_{\boldsymbol{y}} \not\leq_\diamond M_{\boldsymbol{x}}$. Indeed, assume by way of contradiction that $M_{\boldsymbol{y}} \leq_\diamond M_{\boldsymbol{x}}$. Then there is a permutation $f$ of $\{1, \ldots, n\}$ such that $y_i \leq_Q x_{f(i)}$ for all $i = 1, \ldots, n$. From $\boldsymbol{x} \leq_\times \boldsymbol{y}$, we get

$$x_i \leq_Q y_i \leq_Q x_{f(i)} \leq_Q y_{f(i)} \leq x_{f(f(i))} \leq y_{f(f(i))} \leq_Q \cdots \leq_Q x_{f^k(i)} \leq_Q y_{f^k(i)} \leq_Q \cdots$$

So that for all $j$ in the $f$-orbit of $i$, $x_j \equiv_Q x_i \equiv_Q y_j$, entailing $\boldsymbol{y} \equiv_\times \boldsymbol{x}$ which contradicts the assumption $\boldsymbol{x} <_\times \boldsymbol{y}$.

We have thus exhibited a mapping from $Q^n$ to $M_n(Q)$ that will map chains to chains. Hence $\boldsymbol{h}(Q^n) \leq \boldsymbol{h}(M_n(Q))$. ★4.24

**Lemma 4.26** $\boldsymbol{h}(M(Q)) \geq \boldsymbol{h}^*(Q)$.

*Proof* The result is clear in cases where $\boldsymbol{h}^*(Q) = \boldsymbol{h}(Q)$ and when $\boldsymbol{h}(Q) = 1$ entailing $\boldsymbol{h}(M(Q)) = \omega = \boldsymbol{h}^*(Q)$. So let us assume that $\boldsymbol{h}(Q)$ is not additive principal and has a CNF $\sum_{i=1}^m \omega^{\alpha_i}$ with $m > 1$. Thus $\boldsymbol{h}^*(Q) = \boldsymbol{h}(Q) \cdot \omega = \omega^{\alpha_1 + 1}$. Since by Lemma 4.6, for $0 < n < \omega$, $\boldsymbol{h}(Q^n) = \sup\{\alpha \otimes n + 1 \;:\; \alpha < \boldsymbol{h}(Q)\}$, we deduce $\boldsymbol{h}(Q^n) \geq \omega^{\alpha_1} \cdot n + 1$. Since $M_n(Q)$ is a substructure of $M(Q)$, and using Lemma 4.25, we deduce

$$\boldsymbol{h}(M(Q)) \geq \boldsymbol{h}(M_n(Q)) \geq \boldsymbol{h}(Q^n) \geq \omega^{\alpha_1} \cdot n + 1$$

for all $0 < n < \omega$, hence

$$\boldsymbol{h}(M(Q)) \geq \sup_{n<\omega} \omega^{\alpha_1} \cdot n + 1 = \omega^{\alpha_1} \cdot \omega = \boldsymbol{h}^*(Q)\,.$$

★4.26

### 4.5.3 Widths

The previous analyses of the maximal order types and heights of $M(Q)$, $Q^{<\omega}$, and $T(Q)$ allow us to apply the correspondence between $\boldsymbol{o}$, $\boldsymbol{h}$, and $\boldsymbol{w}$ shown by Kříž and Thomas [9, Theorem 4.13], in particular its consequence spelled out in Corollary 3.9.

**Theorem 4.27** *Let $Q$ be a WQO. Then $\boldsymbol{w}(Q^\dagger) = \boldsymbol{o}(Q^\dagger)$ where $Q^\dagger$ can be $T(Q)$, or $Q^{<\omega}$ when $\boldsymbol{o}(Q) > 1$, or $M(Q)$ when $\boldsymbol{o}(Q) > 1$ is a principal additive ordinal.*

*Proof* First observe that $\boldsymbol{h}^*(Q) \leq \boldsymbol{h}(Q) \cdot \omega \leq \boldsymbol{o}(Q) \cdot \omega < \boldsymbol{o}(Q^\dagger)$ when $Q^\dagger$ is $T(Q)$ (by Theorem 4.22), $Q^{<\omega}$ with $\boldsymbol{o}(Q) > 1$ (by Theorem 4.21), or $M(Q)$ with $\boldsymbol{o}(Q) > 1$ (by Theorem 4.20). Furthermore, when $Q^\dagger$ is $T(Q)$ or $Q^{<\omega}$, and when it is $M(Q)$ with $\boldsymbol{o}(Q)$ a principal additive ordinal, $\boldsymbol{o}(Q^\dagger)$ is a principal multiplicative ordinal. Thus Corollary 3.9 shows that $\boldsymbol{w}(Q^\dagger) = \boldsymbol{o}(Q^\dagger)$. ★4.27

The assumptions in Theorem 4.27 seem necessary. For instance, if $Q = 1$, then $M(1)$ is isomorphic to $1^{<\omega}$ and $\omega$, with height $\omega$ and width 1. If $A_3 = 1 \sqcup 1 \sqcup 1$ is an antichain with three elements, then $M(A_3)$ is isomorphic with $\omega \times \omega \times \omega$, $\boldsymbol{h}(M(A_3)) = \omega$ by Lemma 4.6 or Theorem 4.23, $\boldsymbol{o}(M(A_3)) = \omega^3$ by Lemma 4.5, and $\boldsymbol{w}(M(A_3)) = \omega^2$ by Theorem 4.18.(3).

## 4.6 *Infinite Products and Rado's Structure*

One may wonder what happens in the case of infinite products. We remind the reader that the property of being WQO is in general not preserved by infinite products. The classical example for this was provided by Rado [20], who defined what we call the *Rado structure*, denoted $(R, \leq)$[1]: Rado's order is given as a structure on $\omega \times \omega$ where we define

$$(a, b) \leq (a', b') \text{ if } [a = a' \text{ and } b \leq b'] \text{ or } b < a'.$$

The definition of BQOs was motivated by trying to find a property stronger than WQO which is preserved by infinite products, so in particular Rado's example is not a BQO (see [15], Theorems 1.11 and 2.22).

We can use the method of residuals and other tools described in previous sections to compute.

$$\boldsymbol{o}(R) = \omega^2, \qquad \boldsymbol{h}(R) = \omega, \qquad \boldsymbol{w}(R) = \omega, \tag{12}$$

which gives the same ordinal invariants as those of the product $\omega \times \omega$, even though they are not isomorphic, and moreover $\omega \times \omega$ is a BQO (since the notion of BQO is

[1] We adopted the definition from Laver [12].

preserved under products) while Rado's order is not. Therefore one cannot characterise BQOs by the ordinal invariants considered here. Moreover, the two orders do not even embed into each other. To see this, assume by way of contradiction that $f$ injects $\omega \times \omega$ into $R$. Write $(a_i, b_i)$ and $(c_i, d_i)$ for $f(0, i)$ and, resp., $f(i, 0)$ when $i \in \omega$. Necessarily the $b_i$'s and the $d_i$'s are unbounded. If the $a_i$'s are unbounded, one has the contradictory $f(1, 0) <_R f(0, i) = (a_i, b_i)$ for some $i$, and there is a similar contradiction if the $c_i$'s are unbounded, so assume the $a_i$'s and the $c_i$'s are bounded by some $k$. By the pigeonhole principle, we can find a pair $0 < i, j$ with $a_i = c_j$ so that $f(0, i) \not\perp_R f(j, 0)$, another contradiction. Hence $(\omega \times \omega) \not\leq R$. In the other direction $R \not\leq (\omega \times \omega)$, is obvious since $\omega \times \omega$ is BQO while $R$ is not.

## 5 Concluding Remarks

We provide in Table 1 a summary of our findings regarding ordinal invariants of WQOs. Mostly, the new results concern the width $\boldsymbol{w}(P)$ of WQOs. We note that the width $\boldsymbol{w}(P \times Q)$ of Cartesian products is far from elucidated, the first difficulty being that—unlike other constructs—it cannot be expressed as a function of the widths $\boldsymbol{w}(P)$ and $\boldsymbol{w}(Q)$. For Cartesian products, Sect. 4.4 only provide definite values for a few special cases: for the rest, one can only provide upper and lower bounds for the moment.

**Table 1** Ordinal invariants of the main WQOs

| $P$ | $\boldsymbol{o}(P)$ | $\boldsymbol{h}(P)$ | $\boldsymbol{w}(P)$ |
|---|---|---|---|
| $\alpha \in ON$ | $\alpha$ | $\alpha$ | 1 (or 0) |
| $A_n$ (size $n$ antichain) | $n$ | 1 | $n$ |
| Rado's $R$ | $\omega^2$ | $\omega$ | $\omega$ |
| $\sum_{i\in\alpha} P_i$ | $\sum_{i\in\alpha} \boldsymbol{o}(P_i)$ | $\sum_{i\in\alpha} \boldsymbol{h}(P_i)$ | $\sup_{i\in\alpha} \boldsymbol{w}(P_i)$ |
| $P \sqcup Q$ | $\boldsymbol{o}(P) \oplus \boldsymbol{o}(Q)$ | $\max(\boldsymbol{h}(P), \boldsymbol{h}(Q))$ | $\boldsymbol{w}(P) \oplus \boldsymbol{w}(Q)$ |
| $P \cdot Q$ | $\boldsymbol{o}(P) \cdot \boldsymbol{o}(Q)$ | $\boldsymbol{h}(P) \cdot \boldsymbol{h}(Q)$ | $\boldsymbol{w}(P) \odot \boldsymbol{w}(Q)$ |
| $P \times Q$ | $\boldsymbol{o}(P) \otimes \boldsymbol{o}(Q)$ | $\sup_{\substack{\alpha<\boldsymbol{h}(P)\\ \beta<\boldsymbol{h}(Q)}} \alpha \oplus \beta + 1$ | see Sect. 4.4 |
| $M(P)$ | $\omega^{\widehat{\boldsymbol{o}(P)}}$ | $\boldsymbol{h}^*(P)$, see Sect. 4.5.2 | see Theorem 4.27 |
| $P^{<\omega}$ | $\omega^{\omega^{\boldsymbol{o}(P)\pm 1}}$, see Theorem 4.21 | $\boldsymbol{h}^*(P)$ | $\boldsymbol{o}(P^{<\omega})$ |
| $T(P)$ | $\vartheta(\Omega^\omega \cdot \boldsymbol{o}(P))$ | $\boldsymbol{h}^*(P)$ | $\boldsymbol{o}(T(P))$ |

## References


1. Abraham, U. (1987). A note on Dilworth's theorem in the infinite case. *Order*, *4*(2), 107–125. https://doi.org/10.1007/BF00337691.
2. Abraham, U., & Bonnet, R. (1999). Hausdorff's theorem for posets that satisfy the finite antichain property. *Fundamenta Mathematicae*, *159*(1), 51–69.
3. Abraham, U., Bonnet, R., Cummings, J., Džamonja, M., & Thompson, K. (2012). A scattering of orders. *Transactions of the American Mathematical Society*, *364*(12), 6259–6278. https://doi.org/10.1090/S0002-9947-2012-05466-3.
4. Blass, A., & Gurevich, Y. (2008). Program termination and well partial orderings. *ACM Transactions on Computational Logic*, *9*(3), 18. https://doi.org/10.1145/1352582.1352586.
5. Dilworth, R. P. (1950). A decomposition theorem for partially ordered sets. *Annals of Mathematics*, *51*(1), 161–166. https://doi.org/10.2307/1969503.
6. Dushnik, B., & Miller, E. W. (1941). Partially ordered sets. *American Journal of Mathematics*, *63*(3), 600–610. https://doi.org/10.2307/2371374.
7. Higman, G. (1952). Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society*, *3*(2), 326–336. https://doi.org/10.1112/plms/s3-2.1.326.
8. de Jongh, D. H. J., & Parikh, R. (1977). Well-partial orderings and hierarchies. *Indagationes Mathematicae (Proceedings),* *39*(3):195–207. https://doi.org/10.1016/1385-7258(77)90067-1.
9. Kříž, I., Thomas, R. (1990). Ordinal types in Ramsey theory and well-partial-ordering theory. In J. Nešetřil & V. Rödl (Eds.), *Algorithms and Combinatorics: Mathematics of Ramsey Theory* (Vol. 5, pp 57–95). Springer. https://doi.org/10.1007/978-3-642-72905-8_7.
10. Kruskal, J. (1960). Well-quasi-ordering, the Tree theorem, and Vazsonyi's conjecture. *Transactions of the American Mathematical Society*, *95*(2), 210–225. https://doi.org/10.2307/1993287.
11. Kruskal, J. B. (1972). The theory of well-quasi-ordering: A frequently discovered concept. *Journal of Combinatorial Theory, Series A*, *13*(3), 297–305. https://doi.org/10.1016/0097-3165(72)90063-5.
12. Laver, R. (1976). Well-quasi-orderings and sets of finite sequences. In *Mathematical Proceedings of the Cambridge Philosophical Society* (Vol. 79, No. 1, pp. 1–10). https://doi.org/10.1017/S030500410005204X.
13. Van der Meeren, J. (2015). Connecting the two worlds: Well-partial-orders and ordinal notation systems. Ph.D. thesis, Universiteit Gent.
14. Van der Meeren, J., Rathjen, M., & Weiermann, A. (2015). Well-partial-orderings and the big Veblen number. *Archive for Mathematical Logic*, *54*(1–2), 193–230. https://doi.org/10.1007/s00153-014-0408-5.
15. Milner, E. C. (1985). Basic wqo- and bqo-theory. In *Graphs and order (Banff, Alta., 1984), NATO Advanced Science Institutes Series C: Mathematical and Physical Sciences* (Vol. 147, pp. 487–502). Dordrecht: Reidel. https://doi.org/10.1007/978-94-009-5315-4_14.
16. Milner, E. C., & Sauer, N. (1981). On chains and antichains in well founded partially ordered sets. *Journal of the London Mathematical Society*, *2*(1), 15–33. https://doi.org/10.1112/jlms/s2-24.1.15.
17. Mirsky, L. (1971). A dual of Dilworth's decomposition theorem. *American Mathematical Monthly*, *78*(8), 876–877. https://doi.org/10.2307/2316481.
18. Perles, M. A. (1963). On Dilworth's theorem in the infinite case. *Israel Journal of Mathematics*, *1*(2), 108–109. https://doi.org/10.1007/BF02759806.
19. Pouzet, M. (1979). Sur les chaînes d'un ensemble partiellement bien ordonné. *Publ Dép Math Lyon, 16*(1), 21–26. http://www.numdam.org/article/PDML_1979__16_1_21_0.pdf.
20. Rado, R. (1954). Partial well-ordering of sets of vectors. *Mathematika*, *1*(2), 89–95. https://doi.org/10.1112/S0025579300000565.
21. Rathjen, M., & Weiermann, A. (1993). Proof-theoretic investigations on Kruskal's theorem. *Annals of Pure and Applied Logic*, *60*(1), 49–88. https://doi.org/10.1016/0168-0072(93)90192-G.

22. Schmerl, J. H. (2002). Obstacles to extending Mirsky's theorem. *Order*, *19*(2), 209–211. https://doi.org/10.1023/A:1016541101728.
23. Schmidt, D. (1979). *Well-partial orderings and their maximal order types*. Heidelberg: Habilitationsschrift.
24. Schmidt, D. (1981). The relation between the height of a well-founded partial ordering and the order types of its chains and antichains. *Journal of Combinatorial Theory, Series B*, *31*(2), 183–189. https://doi.org/10.1016/S0095-8956(81)80023-8.
25. Schmitz, S., Schnoebelen, Ph. (2011). Multiply-recursive upper bounds with Higman's lemma. In *Proceeding of the ICALP 2011. Lecture Notes in Computer Science* (Vol. 6756, pp. 441–452). Springer. https://doi.org/10.1007/978-3-642-22012-8_35.
26. Weiermann, A. (2009). A computation of the maximal order type of the term ordering on finite multisets. In Proceedings of CiE-2009. Lecture Notes in Computer Science (Vol. 5635, pp. 488–498). Springer. https://doi.org/10.1007/978-3-642-03073-4_50
27. West, D. B. (1982). Extremal problems in partially ordered sets. In Rival, I. (Ed.), *Ordered Sets, NATO Advanced Study Institutes Series* (Vol. 83, pp. 473–521). Springer. https://doi.org/10.1007/978-94-009-7798-3_16.
28. Wolk, E. S. (1967). Partially well ordered sets and partial ordinals. *Fundamenta Mathematicae*, *60*(2), 175–186.

# The Ideal Approach to Computing Closed Subsets in Well-Quasi-orderings

**Jean Goubault-Larrecq, Simon Halfon, Prateek Karandikar, K. Narayan Kumar and Philippe Schnoebelen**

**Abstract** Elegant and general algorithms for handling upwards-closed and downwards-closed subsets of WQOs can be developed using the filter-based and ideal-based representation for these sets. These algorithms can be built in a generic or parameterized way, in parallel with the way complex WQOs are obtained by combining or modifying simpler WQOs.

## 1 Introduction

The theory of well-quasi-orderings (WQOs for short) has proved useful in many areas of mathematics, logic, combinatorics, and computer science. In computer science, it appears prominently in termination proofs [13], in formal languages [12], in graph algorithms (e.g., via the Graph Minor Theorem [46]), in program verification (e.g., with well-structured systems [2, 20, 58]), automated deduction, distributed computing, but also in machine learning [4], program transformation [45], etc. We refer to [37] for "four [main] reasons to be interested in WQO theory".

In computer science, tools from WQO theory were commonly seen as lacking algorithmic contents. This situation is changing. For example, tight complexity bounds for WQO-based algorithms have recently been established and are now used when comparing logics or computational models [30, 54–56]. As another example,

Supported by ANR project BRAVAS (grant ANR-17-CE40-0028).

J. Goubault-Larrecq (✉) · S. Halfon · P. Karandikar
LSV, ENS Paris-Saclay, Université Paris-Saclay, Cachan, France
e-mail: goubault@lsv.fr

P. Karandikar · K. Narayan Kumar
Chennai Mathematical Institute, Chennai, India

P. Schnoebelen
LSV, CNRS, ENS Paris-Saclay, Université Paris-Saclay, Cachan, France

P. Karandikar · K. Narayan Kumar · P. Schnoebelen
CNRS UMI ReLaX, Chennai, India



P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_3

the field of well-structured systems grows not just by the identification of new families of models, but also by the development of new generic algorithms based on WQO theory, see, e.g., [6, 9].

In this chapter we are concerned with the issue of reasoning about, and computing with, downwards-closed and upwards-closed subsets of a WQO. These sets appear in program verification (prominently in model-checking of well structured systems [6], in verification of Petri nets [29], in separability problems [27, 61], but also as an effective abstraction tool [5, 60]). The question of how to handle downwards-closed subsets of WQOs *in a generic way* was first raised by Geeraerts et al.: in [22] the authors postulated the existence of an *adequate domain of limits* satisfying some representation conditions. It turns out that the *ideals* of WQOs always satisfy these conditions, and usually enjoy further algorithmic properties.

*Outline of this chapter*. We start by recalling, as a motivating example, the algorithmic techniques that have been successfully used to handle upwards-closed and downwards-closed subsets in two different WQOs: the tuples of natural numbers with component-wise ordering, and the set of finite words with subword ordering. We then describe the fundamental structures that underlie these algorithms and propose in Sect. 3 a generic set of effectiveness assumptions on which the algorithms can be based.

The second part of the chapter, Sects. 4 and 5, shows how many examples of WQOs used in applications fulfill the required effectiveness assumptions. Since in practice complex WQOs are most often obtained by composing or modifying simpler WQOs, our strategy for showing their effectiveness involves proving that WQO constructors preserve effectiveness.

A final section discusses our choices—of effectiveness assumptions and of algorithms—and lists some of the first questions raised by our approach.

*Genesis of this chapter*. This text grew from [26] (unpublished) where Goubault-Larrecq proposed a notion of effective WQOs, and where Theorem 6.2 was first proven. There, Goubault-Larrecq also shows that products, sequence extensions, and tree extensions of effective WQOs are effective. Then, in 2016 and 2017, Karandikar, Narayan Kumar and Schnoebelen developed the framework and handled WQOs obtained by extensions, by quotients, and by substructures. Finally, in 2017 and 2018, Halfon joined the project and contributed most of the results on powersets and multisets. He also studied variant sets of axioms for effective WQOs as reported in Sect. 6.1. In the meantime, the constructions initiated by [26] have been used in several papers, starting from [17, 18], and including [8, 9, 16, 27, 41–44].

## 2 Well-Quasi-orderings, Ideals, and Some Motivations

A *quasi-ordering* (a QO) $(X, \leq)$ is a set $X$ equipped with a reflexive and transitive relation. We write $x < y$ when $x \leq y$ and $y \not\leq x$, and $x \equiv y$ when $x \leq y$ and $y \leq x$. For $S \subseteq X$, we let $\uparrow S$ and $\downarrow S$ denote the upward and downward closures, respectively, of $S$ in $X$. Formally, $\uparrow S \stackrel{\text{def}}{=} \{x \in X \mid \exists y \in S : y \leq x\}$ and $\downarrow S \stackrel{\text{def}}{=} \{x \in$

$X \mid \exists y \in S : x \leq y\}$. We will also use $\downarrow_< S$ and $\uparrow_< S$ to collect elements that are strictly above, or below, elements of $S$, i.e., $\downarrow_< S \stackrel{\text{def}}{=} \{x \in X \mid \exists y \in S : x < y\}$ and similarly for $\uparrow_< S$.

When $S = \{x\}$ is a singleton, we may simply write $\uparrow x$ or $\downarrow_< x$. A subset of $X$ of the form $\uparrow x$ is called a *principal filter* while a subset of the form $\downarrow x$ is a *principal ideal*. A subset $S \subseteq X$ is *upwards-closed* when $S = \uparrow S$, and *downwards-closed* when $S = \downarrow S$. Note that arbitrary unions and intersections of upwards-closed (resp. downwards-closed) sets are upwards-closed (resp. downwards-closed). Observe also that the complement of an upwards-closed set is downwards-closed, and conversely. We write $Up(X)$ for the set of upwards-closed subsets of $X$, with typical elements $U, U', V, \ldots$ Similarly, $Down(X)$ denotes the set of its downwards-closed subsets, with typical elements $D, D', E, \ldots$

## 2.1 *Two Motivating Examples*

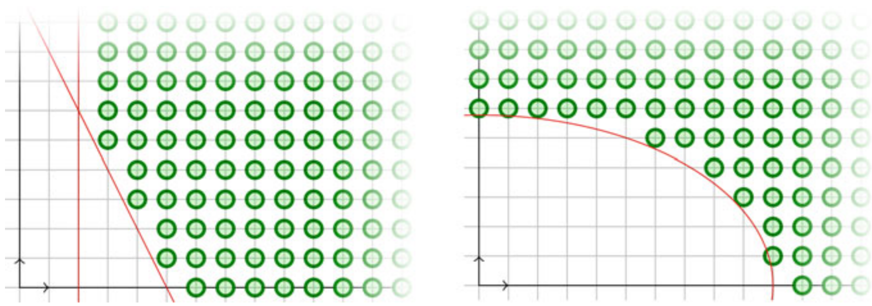
Consider the set $X = \mathbb{N}^2$ of pairs of natural numbers. These are the points with integral coordinates in the upper-right quadrant. We order these points with the coordinate-wise ordering, also called *product ordering*:

$$\langle a, b\rangle \leq \langle a', b\rangle \stackrel{\text{def}}{\Leftrightarrow} a \leq a' \wedge b \leq b' .$$

Note that this is only a partial ordering: $\langle 1, 2\rangle$ and $\langle 3, 0\rangle$ are incomparable.

### 2.1.1 $\mathbb{N}^2$ and Its Upwards-Closed Subsets

In many applications, we need to consider upwards-closed subsets $U, U', \ldots$, of $\mathbb{N}^2$. These may be defined by simple, or not so simple, constraints such as $U_{\text{ex}_1}$ and $V_{\text{ex}_1}$ in Fig. 1.



$U_{\text{ex}_1}: \quad a > 2 \wedge 2a + b > 10$ $\qquad V_{\text{ex}_1}: \quad a^2 + 3b^2 > 100$

**Fig. 1** Two upwards-closed subsets of $\mathbb{N}^2$

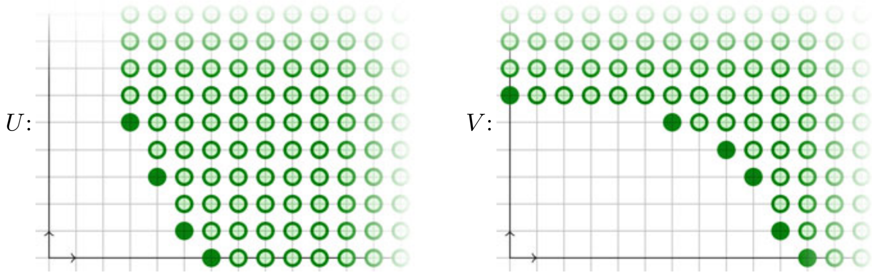



**Fig. 2** Finite bases for $U_{\mathrm{ex}_1}$ and $V_{\mathrm{ex}_1}$

A striking aspect of these depictions of $U_{\mathrm{ex}_1}$ and $V_{\mathrm{ex}_1}$—see also Fig. 2—is that both can be seen as unions of a few principal filters:

$$\begin{aligned} U_{\mathrm{ex}_1} &= \uparrow\langle 3,5\rangle \cup \uparrow\langle 4,3\rangle \cup \uparrow\langle 5,1\rangle \cup \uparrow\langle 6,0\rangle\ , \\ V_{\mathrm{ex}_1} &= \uparrow\langle 0,6\rangle \cup \uparrow\langle 6,5\rangle \cup \uparrow\langle 8,4\rangle \cup \uparrow\langle 9,3\rangle \cup \uparrow\langle 10,1\rangle \cup \uparrow\langle 11,0\rangle\ . \end{aligned}$$

We write $U = \bigcup_{i<n} \uparrow x_i$ to say that the upwards-closed subset $U$ of $X$ is the union of $\uparrow x_0, \ldots, \uparrow x_{n-1}$. The elements $x_i$ are the *generators*, and the finite set $\{x_0, \cdots, x_{n-1}\}$ is a *finite basis* of $U$. We also say that $\bigcup_{i<n} \uparrow x_i$ is a *finite basis representation* of $U$. By removing elements that are not minimal, we obtain a *minimal finite basis* of $U$.

We shall see later that all upwards-closed subsets of $(\mathbb{N}^2, \leq)$ admit such a representation. For the time being we want to stress how this representation of upwards-closed subsets is convenient *from an algorithmic viewpoint*. To begin with, it provides us with a finite data structure for subsets that are infinite and thus cannot be represented in extension on a computer. Interestingly, some important set-theoretical operations are very easy to perform on this representation: testing whether some point $\langle a, b\rangle$ is in $U$ or $V$ just amounts to comparing $\langle a, b\rangle$ with the points forming the basis of $U$ or $V$ respectively. Testing whether $U \subseteq V$ reduces to checking whether all points in the basis of $U$ belong to $V$. We see that $U_{\mathrm{ex}_1} \not\subseteq V_{\mathrm{ex}_1}$ since there is a point in $U_{\mathrm{ex}_1}$'s basis that is not in $V_{\mathrm{ex}_1}$, i.e., not larger than (or equal to) any of the points in $V_{\mathrm{ex}_1}$'s basis: for instance $\langle 3,5\rangle \notin V_{\mathrm{ex}_1}$. Similarly, $\langle 0,6\rangle \notin U_{\mathrm{ex}_1}$ hence $V_{\mathrm{ex}_1} \not\subseteq U_{\mathrm{ex}_1}$.

Two further operations that are easily performed are computing $W = U \cup V$ and $W' = U \cap V$ for upwards-closed $U$ and $V$ (see Fig. 3; recall that such unions and intersections are upwards-closed as observed earlier). For $U \cup V$, we just join the two finite bases and (optionally) remove any element that is not minimal. For example

$$\begin{aligned} U_{\mathrm{ex}_1} \cup V_{\mathrm{ex}_1} &= \big(\uparrow\langle 3,5\rangle \cup \uparrow\langle 4,3\rangle \cup \uparrow\langle 5,1\rangle \cup \uparrow\langle 6,0\rangle\big) \\ &\quad \cup \big(\uparrow\langle 0,6\rangle \cup \cancel{\uparrow\langle 6,5\rangle} \cup \cancel{\uparrow\langle 8,4\rangle} \cup \cancel{\uparrow\langle 9,3\rangle} \cup \cancel{\uparrow\langle 10,1\rangle} \cup \cancel{\uparrow\langle 11,0\rangle}\big) \\ &= \uparrow\langle 0,6\rangle \cup \uparrow\langle 3,5\rangle \cup \uparrow\langle 4,3\rangle \cup \uparrow\langle 5,1\rangle \cup \uparrow\langle 6,0\rangle\ . \end{aligned}$$

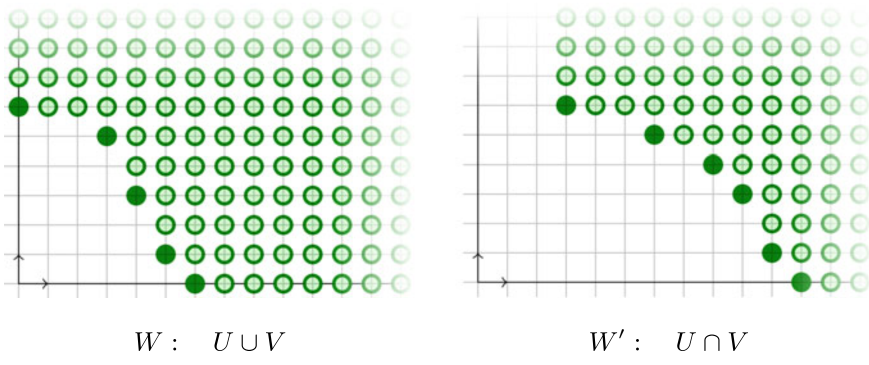



**Fig. 3** Computing intersections and unions via finite bases

For $U \cap V$, we first observe that principal filters can be intersected with

$$\uparrow\langle a, b\rangle \cap \uparrow\langle a', b'\rangle = \uparrow\langle \max(a, a'), \max(b, b')\rangle \tag{1}$$

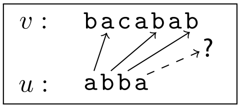
and then use the distributivity law $(\bigcup_{i<n} \uparrow x_i) \cap (\bigcup_{j<m} \uparrow y_j) = \bigcup_{i,j} (\uparrow x_i \cap \uparrow y_j)$ to handle the general case. This gives, for example,

$$\begin{aligned} U_{\text{ex}_1} \cap V_{\text{ex}_1} = & \big[\uparrow\langle 3,5\rangle \cap \uparrow\langle 0,6\rangle\big] \cup \big[\uparrow\langle 3,5\rangle \cap \uparrow\langle 6,5\rangle\big] \cup \big[\uparrow\langle 3,5\rangle \cap \uparrow\langle 8,4\rangle\big] \\ & \cup \big[\uparrow\langle 4,3\rangle \cap \uparrow\langle 9,3\rangle\big] \cup \big[\uparrow\langle 5,1\rangle \cap \uparrow\langle 10,1\rangle\big] \cup \big[\uparrow\langle 6,0\rangle \cap \uparrow\langle 11,0\rangle\big] \\ & \cup \cdots \, more\ filters\ on\ elements\ that\ are\ not\ minimal \, \cdots \\ = & \uparrow\langle 3,6\rangle \cup \uparrow\langle 6,5\rangle \cup \uparrow\langle 8,4\rangle \cup \uparrow\langle 9,3\rangle \cup \uparrow\langle 10,1\rangle \cup \uparrow\langle 11,0\rangle\,. \end{aligned}$$

Finally, a last feature of the finite basis representation for upwards-closed subsets of $\mathbb{N}^2$ is that, if we only consider minimal bases, namely bases of incomparable elements—in essence, if we systematically remove unnecessary generators that are subsumed by smaller generators,—then the representation is *canonical*: there is a unique way of representing any $U \in Up(\mathbb{N}^2)$. Algorithmically, this allows one to implement the required structures using *hash-consing* [15, 24], where structures with the same contents are allocated at the same address, with the help of auxiliary hash-tables. In particular, finite *sets* can be implemented this way, efficiently [25]. Equality tests can then be performed in constant time, notably.

### 2.1.2 Words and Their Subwords

Our second example comes from formal languages and combinatorics [53]. Let us fix a three-letter alphabet $A = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$ and write $A^* = \{u, v, \cdots\}$ for the set of all finite words over $A$. Standardly, the empty word is denoted by $\epsilon$, concatenation is denoted multiplicatively, and $|u|$ is the length of $u$. We write $u \preccurlyeq v$ when $u$ is a *subword* of $v$, i.e., a subsequence: $u$ can be obtained from $v$ by erasing some (occurrences of) letters. It is easy to check whether $u \preccurlyeq v$ by attempting to construct a leftmost embedding of $u$ into $v$: this only requires at most one traversal of $u$ and $v$ and takes time linear in $|u| + |v|$. For example, the box below shows that $u = \mathtt{abba}$ is not a subword of $v = \mathtt{bacabab}$.




With the subword ordering comes the notion of upwards-closed and downwards-closed languages (i.e., sets of words). For example the language $U_{\mathrm{ex}_2} \subseteq A^*$ of words with at least one a and at least two bs is upwards-closed, as is $V_{\mathrm{ex}_2}$, the language of words with length at least 2. These upwards-closed languages occur in many applications and one would like to know good data structures and algorithms for manipulating them. It turns out that any such upwards-closed language can be represented as a finite union of principal filters.[1] For example, $U_{\mathrm{ex}_2}$ and $V_{\mathrm{ex}_2}$ can be written

$$U_{\mathrm{ex}_2} = \uparrow \mathtt{abb} \cup \uparrow \mathtt{bab} \cup \uparrow \mathtt{bba}\ , \quad V_{\mathrm{ex}_2} = \uparrow \mathtt{aa} \cup \uparrow \mathtt{ab} \cup \cdots \cup \uparrow \mathtt{cc} = \bigcup_{|u|=2} \uparrow u\ .$$

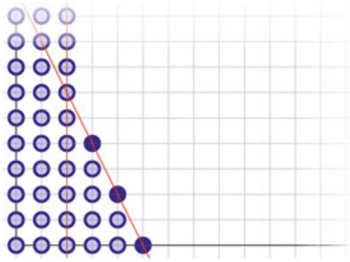
In the subword setting, a principal filter is always a regular language. Indeed, for any $u \in A^*$, of the form $u = a_1 a_2 \cdots a_\ell$, one has $\uparrow u = A^* a_1 A^* a_2 A^* \cdots A^* a_\ell A^*$, which is a language at level $\frac{1}{2}$ in the Straubing-Thérien hierarchy [49]. Being simple star-free regular languages, the upwards-closed subsets can be handled with well-known automata-theoretic techniques. However, one can also use the same simple ideas we used for $\mathbb{N}^2$: testing $U \subseteq V$ reduces to comparing the generators, computing unions is trivial, and bases made of incomparable words provide a canonical representation. Finally, computing intersections reduces to intersecting principal filters, exactly as in $\mathbb{N}^2$. For this, we observe that $\uparrow u \cap \uparrow v$ is generated by the minimal words that contain both $u$ and $v$ as subwords. This set of minimal words, written $u \sqcap v$, is called the *infiltration product* of $u$ and $v$ [11]. For example $\mathtt{ab} \sqcap \mathtt{ca} = \{\mathtt{abca}, \mathtt{acba}, \mathtt{acab}, \mathtt{cab}\}$. Infiltrations are a generalization of shuffles and we shall describe a simple algorithm for a generalized infiltration product in Sect. 4.4.

[1]This result is known as Haines' Theorem [31], and is also a consequence of Higman's Lemma: see Sect. 4.4.

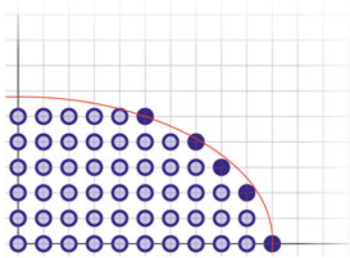
## 2.2 *What About Downwards-Closed Subsets?*

With the previous two examples, we showed how it is natural and easy to work with upwards-closed subsets of a quasi-ordered set when these subsets are represented as a finite union $\bigcup_{i<n} \uparrow x_i$ of principal filters.

Let us now return to our previous setting, $X = \mathbb{N}^2$, and look at the downwards-closed subsets $D, E, \ldots \in Down(\mathbb{N}^2)$. As an example, consider $D_{\mathrm{ex}_1} \stackrel{\mathrm{def}}{=} \mathbb{N}^2 \setminus U_{\mathrm{ex}_1}$ and $E_{\mathrm{ex}_1} \stackrel{\mathrm{def}}{=} \mathbb{N}^2 \setminus V_{\mathrm{ex}_1}$. We shall sometimes write $D_{\mathrm{ex}_1} = \neg U_{\mathrm{ex}_1}$ and $E_{\mathrm{ex}_1} = \neg V_{\mathrm{ex}_1}$.



$D_{\mathrm{ex}_1}$ : $a \leq 2 \vee 2a + b \leq 10$



$E_{\mathrm{ex}_1}$ : $a^2 + 3b^2 \leq 100$

Here, $E_{\mathrm{ex}_1}$ can be represented using its maximal points as generators:

$$E_{\mathrm{ex}_1} = \downarrow\langle 5, 5\rangle \cup \downarrow\langle 7, 4\rangle \cup \downarrow\langle 8, 3\rangle \cup \downarrow\langle 9, 2\rangle \cup \downarrow\langle 10, 0\rangle \ .$$

Representing downwards-closed sets via a finite "basis", i.e., as a finite union of principal ideals, of the form $\bigcup_{i<n} \downarrow x_i$, allows for simple and efficient algorithms, exactly as for upwards-closed subsets: one tests inclusion by comparing the generators of the ideals, and computes unions by gathering all generators and (optionally) removing non-maximal ones. For intersections one uses

$$\downarrow\langle a, b\rangle \cap \downarrow\langle a', b'\rangle = \downarrow\langle \min(a, a'), \min(b, b')\rangle \tag{2}$$

and the distribution law $(\bigcup_i \downarrow x_i) \bigcap (\bigcup_j \downarrow y_j) = \bigcup_i \bigcup_j (\downarrow x_i \cap \downarrow y_j)$, valid in every QO.

However, there is an important limitation here that we did not have with upwards-closed subsets: not all downwards-closed subsets in $\mathbb{N}^2$ can be generated from finitely many elements. Indeed, for any $x \in \mathbb{N}^2$, the ideal $\downarrow x$ is finite and thus only the finite downwards-closed subsets of $\mathbb{N}^2$ can be represented via principal ideals. Hence $D_{\mathrm{ex}_1}$ in the previous figure, or even $\mathbb{N}^2$ itself, while perfectly downwards-closed, cannot be represented in this way.

A possible solution is to represent a downwards-closed subset $D \in Down(X)$ via the finite basis of its upwards-closed complement, writing $D = X \setminus \bigcup_{i<n} \uparrow x_i$, or also $D = X(\setminus \uparrow x_i)_{i<n}$. Continuing our example, $D_{\mathrm{ex}_1} = \neg U_{\mathrm{ex}_1}$ can be written $D_{\mathrm{ex}_1} = \mathbb{N}^2 \setminus \uparrow\langle 3, 5\rangle \setminus \uparrow\langle 4, 3\rangle \setminus \uparrow\langle 5, 1\rangle \setminus \uparrow\langle 6, 0\rangle$. This representation *by excluded minors* is contrapositive and thus counter-intuitive. Computing intersections becomes

easier while unions become harder, which is usually not what we want in applications. More annoyingly, constructing a representation of $\downarrow x$ from $x$ involves actually computing complements, a task that can be difficult in general as we shall see later. Even in the easy $\mathbb{N}^2$ case, it is not transparent how from, e.g., $x = \langle 2, 3\rangle$, one gets to $\downarrow x = \mathbb{N}^2 \setminus \uparrow\langle 0, 4\rangle \setminus \uparrow\langle 3, 0\rangle$.

### 2.2.1 Downwards-Closed Subsets with $\omega$'s

In the case of $\mathbb{N}^2$, there exists an elegant solution to the representation problem for downwards-closed sets: one uses pairs $\langle a, b\rangle \in \mathbb{N}_\omega^2$ where $\mathbb{N}_\omega$ extends $\mathbb{N}$ with an extra value $\omega$ that is larger than all natural numbers. We can now denote $D_{\mathrm{ex}_1} = \neg U_{\mathrm{ex}_1}$ (see last figure) as $\downarrow\langle 2, \omega\rangle \cup \downarrow\langle 3, 4\rangle \cup \downarrow\langle 4, 2\rangle \cup \downarrow\langle 5, 0\rangle$. We note that $\downarrow\langle 2, \omega\rangle$ should probably be written more explicitly as $(\downarrow 2) \times \mathbb{N}$ since it denotes $\{\langle c, d\rangle \mid a \leq 2 \wedge b \in \mathbb{N}\}$, a subset of $\mathbb{N}^2$, not of $\mathbb{N}_\omega^2$, however the $\omega$-notation inherited from vector addition systems [36] is now well-entrenched and we retain it here.

The sets of the form $\downarrow\langle a, b\rangle$ where $a, b \in \mathbb{N}_\omega$ are the *ideals*[2] of $\mathbb{N}^2$, and we see that they comprise the principal ideals as a special case. They also comprise infinite subsets and, for example, $\mathbb{N}^2 = \downarrow\langle \omega, \omega\rangle$ is one of them.

Using such ideals, all the downwards-closed subsets of $\mathbb{N}^2$ can be represented, and the algorithms for membership, inclusion, union and intersection are just minor extensions of what we showed for finite downwards-closed sets, when all generators were proper elements of $\mathbb{N}^2$. The only difference is that we have to handle $\omega$'s in the obvious way when comparing generators (e.g., in inclusion tests) and when computing min's, e.g., in (2). Additionally, and like for upwards-closed subsets of $\mathbb{N}^2$, the representation of downwards-closed sets by the downward closure of incomparable elements is canonical, which here too brings in important algorithmic benefits.

Now that we have finite representations for both upwards-closed and downwards-closed subsets of $\mathbb{N}^2$, it is natural to ask whether we can compute complements.

It turns out that, for $\mathbb{N}^2$, this is an easy task. For complementing filters, one uses

$$\neg\uparrow\langle a, b\rangle = \begin{Bmatrix} \downarrow\langle a-1, \omega\rangle & \text{if } a > 0 \\ \emptyset & \text{otherwise} \end{Bmatrix} \bigcup \begin{Bmatrix} \downarrow\langle \omega, b-1\rangle & \text{if } b > 0 \\ \emptyset & \text{otherwise} \end{Bmatrix}. \tag{3}$$

We see where the $\omega$'s are needed. In fact, only $\neg\uparrow\langle 0, 0\rangle = \emptyset$ does not involve $\omega$'s. We note that $\emptyset$, a downwards-closed subset, is indeed a finite union of ideals: it is the empty union.

Complementing an ideal is also easy:

$$\neg\downarrow\langle a, b\rangle = \begin{Bmatrix} \uparrow\langle a+1, 0\rangle & \text{if } a < \omega \\ \emptyset & \text{otherwise} \end{Bmatrix} \bigcup \begin{Bmatrix} \uparrow\langle 0, b+1\rangle & \text{if } b < \omega \\ \emptyset & \text{otherwise} \end{Bmatrix}. \tag{4}$$

[2] We shall soon give the general definition. For now, the reader has to accept the $\mathbb{N}^2$ case.

We see here that complementing an ideal in $\mathbb{N}^2$ always returns a union of principal filters, with no $\omega$'s.

Complementing an arbitrary upwards-closed subset $U$ is easy if $U = \bigcup_{i<n} \uparrow x_i$ is given as a finite union of filters: we compute $\bigcap_{i<n}(X \setminus \uparrow x_i)$. This needs complementing filters and intersecting downwards-closed sets, two operations we know how to perform on $\mathbb{N}^2$. Complementing an arbitrary downwards-closed subset $D = \bigcup_{i<n} \downarrow x_i$ is done similarly, even with $x_i \in \mathbb{N}^2_\omega$: we complement each ideal and intersect the resulting upwards-closed sets.

Finally, let us observe that, since any upwards-closed set is a finite union of filters, the proof that the complement $\neg \uparrow \langle a, b \rangle$ of any filter, and the intersection of any two ideals of $\mathbb{N}^2$, can be expressed as a finite union of ideals, entails that any downwards-closed $D \in Down(\mathbb{N}^2)$ is a finite union of ideals, a result known as *expressive completeness*.

### 2.2.2 Downwards-Closed Sets of Subwords

What about downwards-closed sets in $(A^*, \preccurlyeq)$? As with $\mathbb{N}^2$, finite unions of principal ideals, of the form $\downarrow u_1 \cup \cdots \cup \downarrow u_\ell$, are easy to compare and combine but they can only describe the finite downwards-closed languages. The contrapositive representation by excluded minors can describe any downwards-closed set but here too it is cumbersome. For example, let us fix $A = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$ and consider the language $D_{\text{ex}_3} = \mathtt{a}^*\mathtt{b}^*$, i.e., the set of all words composed of any number of a's followed by any number of b's: it is clear that $D_{\text{ex}_3}$ is closed by taking subwords, hence $D_{\text{ex}_3} \in Down(A^*)$. Its representation by excluded minors is $D_{\text{ex}_3} = \neg(\uparrow \mathtt{ba} \cup \uparrow \mathtt{c})$. That is, "a word $w \in A^*$ is in $D_{\text{ex}_3}$ iff it does not contain any c, nor some b before an a": arguably, using $\mathtt{a}^*\mathtt{b}^*$ to denote $D_{\text{ex}_3}$ is clearer.

We do not develop this example further, and just announce that indeed the regular expression $\mathtt{a}^*\mathtt{b}^*$ denotes an ideal of $(A^*, \preccurlyeq)$, as we shall show in Sect. 4.4. Furthermore, and as with $\mathbb{N}^2_\omega$, algorithms for comparing ideals in $A^*$ are similar to algorithms that compare elements of $A^*$. For example, testing whether (the language denoted by) $\mathtt{a}^*\mathtt{b}^*$ is a subset of $\mathtt{b}^*\mathtt{c}^*\mathtt{a}^*$ is essentially like testing whether ab is a subword of bca.

## 2.3 *Well-Quasi-orders*

The previous section has made it clear that writing upwards-closed sets as a finite union of principal filters, when possible, is handy as far as computation is concerned. The quasi-orders for which it is possible to represent all upwards-closed sets as such is known: it is the class of well-quasi orders, which we introduce below.

A QO $(X, \leq)$ is *well-founded* $\overset{\text{def}}{\Leftrightarrow}$ it does not contain an infinite strictly decreasing sequence $x_0 > x_1 > x_2 > \cdots$. A subset $S \subseteq X$ is an *antichain* if for all distinct

$x, y \in S$, neither of $x \leq y$ and $y \leq x$ holds. A QO is *well* (WQO) $\overset{\text{def}}{\Leftrightarrow}$ it is well-founded and does not contain an infinite antichain. Equivalently, $(X, \leq)$ is WQO iff every infinite sequence $(x_i)_{i \in \mathbb{N}}$ contains an infinite monotonic subsequence $x_{i_0} \leq x_{i_1} \leq x_{i_2} \leq \cdots$ with $i_0 < i_1 < i_2 < \cdots$. See [39, 57] for proofs and other equivalent characterizations.

*Example 2.1 (Some well-known WQOs)*

**linear orderings**: $(\mathbb{N}, \leq)$ is a WQO, as is every ordinal or every well-founded linear-ordering.

**words and sequences**: $(\Sigma^*, \preccurlyeq)$, the set of words over a finite alphabet with the (scattered) subword ordering is a WQO. Variants and extensions abound [12, 28, 61]. By Higman's Lemma, for any WQO $(X, \leq)$, its sequence extension ordered by embedding, $(X^*, \leq_*)$, is a WQO too.

**powersets**: $(\mathcal{P}_f(X), \sqsubseteq_H)$, the set of all *finite* subsets of $(X, \leq)$ with Hoare's subset embedding is a WQO when $X$ is. The full powerset $\mathcal{P}(X)$ is a WQO if $X$ is an $\omega^2$-WQO, a slightly stronger requirement than just being WQO, see [48].

**trees**: Labeled finite trees ordered by embedding form a WQO (Kruskal's Tree Theorem [38]).

**graphs**: Finite graphs ordered by the minor relation constitute a WQO (Robertson and Seymour's Graph Minor Theorem [51]). □

Coming back to our motivation, here is the result claimed at the beginning of this section:

**Lemma 2.2** (Finite basis property) *If $(X, \leq)$ is WQO then every upwards-closed $U \in Up(X)$ contains a* finite basis $B \subseteq U$ *such that $U = \bigcup_{x \in B} \uparrow x$.*

It is easy to see that the converse holds: if every upwards-closed set has a finite basis, then $(X, \leq)$ is WQO.

Lemma 2.2 validates our choice of representing sets via a finite set of generators, as we did in our two motivating examples. It also entails that, when $X$ is a countable WQO, $Up(X)$ is countable too, as is $Down(X)$ since complementation bijectively relates upwards-closed and downwards-closed subsets (see [10] for a more general statement).

We conclude this section by mentioning another useful characterization of WQOs, see [39].

**Lemma 2.3** (Ascending/Descending chain condition) *If $(X, \leq)$ is WQO then there exists no infinite strictly increasing sequence $U_0 \subsetneq U_1 \subsetneq U_2 \subsetneq \cdots$ of upwards-closed subsets. Dually, there exists no infinite strictly decreasing sequence $D_0 \supsetneq D_1 \supsetneq D_2 \supsetneq \cdots$ of downwards-closed subsets.*

*In other words, $(Up(X), \supseteq)$ and $(Down(X), \subseteq)$ are well-founded posets.*

### 2.4 Canonical Prime Decompositions of Closed Subsets

We now recall some basic facts about the canonical decompositions of upwards-closed and downwards-closed subsets in prime components.

Let $(X, \leq)$ be a WQO. We use *Up* and *Down* as abbreviations for $Up(X)$ and $Down(X)$.

**Definition 2.4** (*Prime subsets*) 1. A non-empty $U \in Up$ is *(up) prime* if for any $U_1, U_2 \in Up$, $U \subseteq (U_1 \cup U_2)$ implies $U \subseteq U_1$ or $U \subseteq U_2$.
2. Similarly, a non-empty $D \in Down$ is *(down) prime* if $D \subseteq (D_1 \cup D_2)$ implies $D \subseteq D_1$ or $D \subseteq D_2$.

Observe that all principal filters are up prime and all principal ideals are down prime. Note also that, by definition, the empty subset is not prime.

**Lemma 2.5** (Irreducibility) *1. $U \in Up$ is prime if, and only if, for all $U_1, \ldots, U_n \in Up$, $U = U_1 \cup \cdots \cup U_n$ implies $U = U_i$ for some $i$.*
*2. $D \in Down$ is prime if, and only if, for all $D_1, \ldots, D_n \in Down$, $D = D_1 \cup \cdots \cup D_n$ implies $D = D_i$ for some $i$.*

The following lemma highlights the importance of prime subsets.

**Lemma 2.6** *1. Every upwards-closed set $U \in Up$ is a finite union of up primes.*
*2. Every downwards-closed set $D \in Down$ is a finite union of down primes.*

*Proof* 1. is trivial: the finite basis property of WQOs (Lemma 2.2) shows that any upwards-closed set is a finite union of filters.

2. is a classical result, going back to Noether, see [7, Chap. VIII, Corollary, p. 181]. We include a proof for the reader's convenience. That proceeds by well-founded induction on $D$ in the well-founded poset $(Down, \subseteq)$ (Lemma 2.3). If $D$ is empty, then it is an empty (hence finite) union of primes. If $D$ is prime, the claim holds trivially. Finally, if $D \neq \emptyset$ is not prime, then by Lemma 2.5 it can be written as $D = D_1 \cup \cdots \cup D_n$ where each $D_i$ is properly contained in $D$. By induction hypothesis each $D_i$ is a finite union of primes. Hence $D$ is too. □

We say that a finite collection $\{P_1, \cdots, P_n\}$ of up (resp. down) primes is a *decomposition* of $U \in Up$ (resp., of $D \in Down$) if $U = P_1 \cup \cdots \cup P_n$ (resp., $D = P_1 \cup \cdots \cup P_n$). The decomposition is *minimal* if $P_i \subseteq P_j$ implies $i = j$.

**Theorem 2.7** (Canonical decomposition) *Any upwards-closed $U$ (resp. downwards-closed $D$) has a finite minimal decomposition. Furthermore this minimal decomposition is unique. We call it the canonical decomposition of $U$ (resp. $D$).*

*Proof* By Lemma 2.6, any $U$ (or $D$) has a finite decomposition: $U$ (or $D$) $= \bigcup_{i=1}^{n} P_i$. The decomposition can be made minimal by removing any $P_i$ that is strictly included in some $P_j$. To prove uniqueness we assume that $\bigcup_{i=1}^{n} P_i = \bigcup_{j=1}^{m} P'_j$ are two minimal decompositions. From $P_i \subseteq \bigcup_j P'_j$, and since $P_i$ is prime, we deduce that

$P_i \subseteq P'_{k_i}$ for some $k_i$. Similarly, for each $P'_j$ there is $\ell_j$ such that $P'_j \subseteq P_{\ell_j}$. The inclusions $P_i \subseteq P'_{k_i} \subseteq P_{\ell_{k_i}}$ require $i = \ell_{k_i}$ by minimality of the decomposition, hence are equalities $P_i = P'_{k_i}$. Similarly $j = k_{\ell_j}$ and $P'_j = P_{\ell_j}$ for any $j$. This one-to-one correspondence shows $\{P_1, \cdots, P_n\} = \{P'_1, \cdots, P'_m\}$. □

## 2.5 Filter Decompositions and Ideal Decompositions

**Definition 2.8** (*Ideals)* A subset $S$ of $X$ is an *ideal* it if is non-empty, downwards-closed, and directed. We write $Idl(X) = \{I, J, \cdots\}$ for the set of all ideals of $X$.

Recall that $S$ is *directed* if for all $x_1, x_2 \in S$, there exists $x \in S$ such that $x_1 \le x$ and $x_2 \le x$.

A *filter* is a non-empty, upwards-closed, and filtered set $S$, where *filtered* means that for all $x_1, x_2 \in S$, there exists $x \in S$ such that $x \le x_1, x_2$. In a WQO, the filters are exactly the principal filters, hence there is no need to introduce a new notion. We write $Fil(X)$ for the set of all (principal) filters of $X$.

Every principal ideal $\downarrow x$ is directed hence is an ideal. However not all ideals are principal. For example, in $(\mathbb{N}, \le)$, the set $\mathbb{N}$ itself is an ideal (it is directed) and not of the form $\downarrow n$ for any $n \in \mathbb{N}$.

*Remark 2.9* The above example illustrates a general phenomenon: the limit of an monotonic sequence of ideals (more generally, of a directed family of ideals) is an ideal. In particular, if $x_0 < x_1 < x_2 < \cdots$ is an infinite increasing sequence, $\bigcup_{i=0,1,2,\ldots} \downarrow x_i$ is an ideal. It can be seen as the downward closure of a limit point, e.g. when one writes things like "$\bigcup_{n\in\mathbb{N}} \downarrow n = \downarrow \omega$". It turns out that $(Idl(X), \subseteq)$, the domain-theoretical *ideal completion* of $X$, is isomorphic to the *sobrification* $(\widehat{X}, \le)$ —a topological completion—of $(X, \le)$, see [17] for definitions and details. □

The following appears for example as Lemma 1.1 in [35].

**Proposition 2.10** *1. The up primes are exactly the filters.*
*2. The down primes are exactly the ideals.*

*Proof* 1. is clear and we focus on 2.
($\Longrightarrow$): We only have to check that a down prime $P$ is directed. Assume it is not. Then it contains two elements $x_1, x_2$ such that $\uparrow x_1 \cap \uparrow x_2 \cap P = \emptyset$. In other words, $P \subseteq (P \setminus \uparrow x_1) \cup (P \setminus \uparrow x_2)$. But $P \setminus \uparrow x_i$ is downwards-closed for both $i = 1, 2$, so $P$ being prime is included in one $P \setminus \uparrow x_i$. This contradicts $x_i \in P$.
($\Longleftarrow$): Consider an ideal $I \subseteq X$. Aiming for a contradiction, we assume that $I \subseteq D_1 \cup D_2$ but $I \not\subseteq D_1$, $I \not\subseteq D_2$. Pick a point $x_1$ from $I \setminus D_1$, and a point $x_2$ from $I \setminus D_2$. Since $I$ is directed, there is a point $x \in I$ such that $x_1, x_2 \le x$. Since $D_1$ is downwards-closed, $x$ is not in $D_1$, and similarly $x$ is not in $D_2$, so $x$ is not in $D_1 \cup D_2$, contradiction. □

Proposition 2.10 explains why ideals appeared in our representation of downwards-closed sets of $\mathbb{N}^2$ in Sect. 2.2. There is a general reason: ideals are the down primes used in canonical decompositions, just like filters do for upwards-closed sets. Primality explains why the representation is canonical, and why comparing downwards-closed sets reduces to comparing generators. Meanwhile, the view of ideals as sets of the form $\downarrow x$ where $x$ is either a normal point in $X$ or, possibly, a limit point in $\widehat{X}$—recall Remark 2.9—explains why comparing ideals is often very similar to comparing points—recall testing whether $\downarrow\langle 3, 4\rangle \subseteq \downarrow\langle \omega, 1\rangle$ or whether $\mathtt{a}^*\mathtt{b}^* \subseteq \mathtt{b}^*\mathtt{c}^*\mathtt{a}^*$.

## 3 Ideally Effective WQOs

When describing generic algorithms for WQOs, one needs to make some basic computational assumptions on the WQOs at hand. Such assumptions are often summarized informally along the line of "*the* WQO $(X, \leq)$ *is effective*" and their precise meaning is often defined at a later stage, when one gives sufficient conditions based on the algorithm one is describing, a classic example being [20]. Sometimes the effectiveness assumptions are not even spelled out formally, e.g., when one has in mind applications where the WQO is $(\mathbb{N}^k, \leq_\times)$ or $(A^*, \preccurlyeq)$ which are obviously "effective" under all expected understandings.

The situation is different in this chapter since our goal is to provide a formal notion of effectiveness that is *preserved* by the main WQO constructions (and that supports the computation on closed subsets illustrated in Sect. 2.1). As a consequence, we cannot avoid giving a formal definition, even if this mostly amounts to administrative technicalities.

To simplify this task, we start by fixing the representation for closed subsets: these will be represented as finite unions of prime subsets as explained in Sect. 2. This provides a robust, generic, and convenient data structure for $Up(X)$ and $Down(X)$ based on data structures (to be defined) for $Fil(X)$ and $Idl(X)$. We do not require the decomposition to be canonical and leave this as an implementation choice (the underlying complexity trade-offs depend on the WQO and the application at hand). Moreover, and since all filters are principal in WQOs, any data structure for $X$ can be reused for representing $Fil(X)$, so we will only need to assume that $X$ and $Idl(X)$ have an effective presentation.

This leads to the following definition. Note that, rather than being completely formal and talk of recursive languages or Gödel numberings, we will allow considering more versatile data structures like terms, tuples, graphs, etc., that are closer to actual implementations. All data structures considered in this paper will be recursive sets, and in particular one can enumerate their elements.

**Definition 3.1** (*Ideally Effective WQOs*) A WQO $(X, \leq)$ further equipped with data structures for representing $X$ and $Idl(X)$ is *ideally effective* if:

(OD) the ordering $\leq$ is decidable on (the representation of) $X$;
(ID) similarly, $\subseteq$ is decidable on $Idl(X)$;
(PI) principal ideals are computable, that is,
$x \mapsto \downarrow x$is computable;
(CF) complementation of filters, denoted
$\neg : Fil(X) \to Down(X)$, is computable;
(IF) intersection of filters, denoted
$\cap : Fil(X) \times Fil(X) \to Up(X)$, is computable;
(CI) complementation of ideals, denoted
$\neg : Idl(X) \to Up(X)$, is computable;
(II) intersection of ideals, denoted
$\cap : Idl(X) \times Idl(X) \to Down(X)$, is computable.

Some immediate remarks are in order:

- As mentioned earlier, elements of $Up(X)$ and $Down(X)$ are represented as collections (via lists, or sets, or ...) of elements of $X$ and of $Idl(X)$ respectively. The computability of unions is thus trivial and therefore was not required in the formal definition.
- Similarly, checking membership $x \in D$ for downwards-closed sets reduces to deciding $\downarrow x \subseteq D$, hence was not required either.
- We said earlier that operations on *Up* and *Down* boil down to operations on filters and ideals. Note that there are some subtleties. For example, deciding inclusions over $Up(X)$ or $Down(X)$ is made possible because the decompositions only use prime subsets. Explicitly, in order to check whether $D \subseteq D'$ for example, where $D = \bigcup_{i<m} I_i$ and $D' = \bigcup_{j<n} I'_j$, we check whether every $I_i$ is included in some $I'_j$—this is correct because every ideal $I_i$ is down prime.
- There is some asymmetry in the definition between upwards-closed and downwards-closed sets. This should be expected since WQOs are well-founded but the reverse orderings need not be.
- The astute reader may have noticed that the definition contains some hidden redundancies. Our proposal is justified by algorithmic efficiency concerns, see discussion in Sect. 6.1.

## 3.1 Some First Ideally Effective WQOs

We quickly show that the simplest WQOs are ideally effective. They will be used later as building blocks for more complex WQOs.

### 3.1.1 Finite Orderings

A frequently occurring quasi-ordering in computer science is the *finite alphabet with $n$ symbols*. It consists of a set with $n$ elements, usually denoted $A$, ordered by equality. This is a WQO since $A$ is finite. The name "alphabet" comes from its applications in language theory but this very basic WQO appears in many other situations, e.g., as colorings of some other objects, as the set of control states in formal models of computations such as Turing machines, communicating automata, etc.

Let us spell out, as a warming-up exercise, why this WQO $(A, =)$ is ideally effective. One can for instance represent elements of $A$ using natural numbers up to $|A| - 1$. The ordering is trivially decidable. All ideals of $(A, =)$ are principal, that is of the form $\downarrow x = \{x\}$ for $x \in A$. We thus represent ideals as elements, exactly as we do for filters. Therefore, ideal inclusion coincides with equality, and (PI) is given by the identity function. All other operations are trivial: intersection of filters (*resp.* ideals) is always empty except if the two filters (*resp.* ideals) are equal, and $\neg \uparrow x = \neg \downarrow x = A \setminus \{x\}$.

We could of course have dispensed with these explanations since, more generally, any finite QO is a WQO and is ideally effective. In particular, all operations required by Definition 3.1 are always computable, being operations on a finite set. Let us note that all ideals are principal in this setting, which is no surprise since $(X, \geq)$ is also a WQO, and its filters are the ideals of $(X, \leq)$.

### 3.1.2 Natural Numbers

Apart from finite orders, the simplest WQO is $(\mathbb{N}, \leq)$. We now restate our observations from Sect. 2.1 in the more formal framework of Definition 3.1.

Observe that since $\leq$ is linear, any downwards-closed set is actually an ideal, except for $\emptyset$. The ideals that are bounded from above have the form $\downarrow n$ for some $n \in \mathbb{N}$, and the only unbounded ideal is the whole set $\mathbb{N}$ itself, often denoted $\downarrow \omega$ as we did in Sect. 2.2. Ideal inclusion is thus decidable: principal ideals are compared as elements, and $\downarrow \omega$ is larger than all the others. Thus $(Idl(\mathbb{N}), \subseteq)$ is linearly ordered, which makes intersections trivial: one has $\uparrow n \cap \uparrow m = \uparrow \max(n, m)$ and $\downarrow n \cap \downarrow m = \downarrow \min(n, m)$. Finally, complements are computed as follows:

$$\neg \uparrow (n+1) = \downarrow n \,, \qquad \neg \downarrow n = \uparrow (n+1) \,,$$
$$\neg \uparrow 0 = \emptyset \,, \qquad \neg \downarrow \omega = \emptyset \,.$$

### 3.1.3 Ordinals

The above analysis extends to any recursive linear WQO, i.e., any recursive ordinal (see [52] for definitions). Given an ordinal $\alpha$, we write $\boldsymbol{\alpha}$ (in bold font) for the set of ordinals $\{\beta \mid \beta < \alpha\}$—the classical set-theoretic construction of ordinals equates $\alpha$ with $\boldsymbol{\alpha}$.

Let $(X, \leq) = (\boldsymbol{\alpha}, \leq)$. Once again, $X$ being linearly ordered, its ideals are its downwards-closed sets (except $\emptyset$). Therefore, there are three types of ideals:

1. $I = X$,
2. $I$ has a maximal element $\beta \in X$, in which case $I = \downarrow\beta$,
3. Or $I$ has a supremum $\beta \in X \setminus I$, in which case $I = \downarrow_< \beta = \boldsymbol{\beta}$.

Note that in the second case, $I = \downarrow\beta = \downarrow_<(\beta + 1) = \boldsymbol{\beta} + \mathbf{1}$. Thus every ideal of $(X, \leq)$ is a $\boldsymbol{\beta}$ for some $\beta \in \boldsymbol{\alpha} + \mathbf{1} \setminus 0$, and ideal inclusion coincides with the natural ordering on $\boldsymbol{\alpha} + \mathbf{1}$.

Now, assuming that we can represent elements of $X$ in a way that makes $\leq$ decidable, then $(X, \leq)$ is ideally effective. Indeed, the representation is easily extended to $(\boldsymbol{\alpha} + \mathbf{1}, \leq)$ and one can thus decide ideal inclusion. Intersections are computable as the maximum for filters, as the minimum for ideals. Finally, complements of filters and ideals are computed as follows:

$$\begin{aligned} \neg\uparrow 0 &= \emptyset\,, \\ \neg\boldsymbol{\alpha} &= \emptyset\,, \end{aligned} \qquad \left.\begin{aligned} \neg\uparrow\beta &= \boldsymbol{\beta} \\ \neg\boldsymbol{\beta} &= \uparrow\beta \end{aligned}\right\} \text{for } \beta \in \boldsymbol{\alpha}, \beta \neq 0\,.$$

While the above applies to any recursive ordinal, the applications that we are aware of usually only need ordinals below $\epsilon_0$, for which the *Cantor Normal Form* is well known and understood, and leads to natural data structures [47]. One can push this at least to all ordinals below the larger ordinal $\Gamma_0$ [21].

Note that, when $\alpha = \omega$, the representation of ideals differs from the representation for $Idl(\mathbb{N})$ proposed in Sect. 3.1.2: in one case we use $\downarrow_< n$ while in the other we use $\downarrow n$. Both options are equivalent, leading to very similar algorithms. In Sect. 3.1.2 we adopted the representation that has long been common in Petri net tools.

# 4 Constructing Ideally Effective WQOs

We now look at more complex WQOs. In practice these are obtained by combining simpler WQOs via well-known operations such as Cartesian product, sequence extension, etc. Our strategy is thus to show that these operations produce ideally effective WQOs when they are applied to ideally effective WQOs.

## *4.1 Ideally Effective WQO Constructors*

We shall provide generic (i.e., uniform) algorithms that manage filters and ideals of compound WQOs by invoking the algorithms for the filters and ideals of their components. This is made precise in Definition 4.2, and to this end, we have to introduce the following notion:

**Definition 4.1** A *presentation* of an ideally effective WQO $(X, \leq)$, is a list of:

- data structures for $X$ and $Idl(X)$,
- algorithms for the seven computable functions required by Definition 3.1,

(XI) – the ideal decomposition $X = \bigcup_{i<n} I_i$ of $X$ as a downwards-closed set,

(XF) – as well as its filter decomposition $X = \bigcup_{i<n'} F_i$.

Obviously, a WQO is ideally effective if and only if it has a presentation as defined in Definition 4.1.

The notion of presentations as actual objects is needed because they are the actual inputs of our WQO constructions. This explains why we added (XI) and (XF) in the requirements. For a given $(X, \leq)$, the ideal and filter decompositions of $X$ always exist and requiring them in Definition 3.1 would make no sense. However, these decompositions are needed by algorithms that work uniformly on WQOs given via their presentations.

Let us informally call *order-theoretic constructor* (constructor for short) any operation $C$ that produces a quasi-ordering $C[(X_1, \leq_1), \ldots, (X_n, \leq_n)]$ from given quasi-orderings $(X_1, \leq_1), \ldots, (X_n, \leq_n)$. In subsequent sections, $C$ will be instantiated with very well-known constructions, such as Cartesian product with component-wise ordering, finite sequences with Higman's ordering, finite sets with the Hoare quasi-ordering, and so on. In practice, we will always have $n = 1$ or 2. We also say that an order-theoretic constructor preserves WQO if $C[(X_1, \leq_1), \ldots, (X_n, \leq_n)]$ is a WQO whenever $(X_1, \leq_1), \ldots, (X_n, \leq_n)$ are. The constructors we just mentioned are well-known to be WQO-preserving. We extend this concept to ideally effective WQOs:

**Definition 4.2** An order-theoretic WQO-preserving constructor $C$ is said to be *ideally effective* if:

- It preserves ideal effectiveness, that is $C[(X_1, \leq_1), \ldots, (X_n, \leq_n)]$ is ideally effective when each $(X_i, \leq_i)$ is.
- A presentation of $C[(X_1, \leq_1), \ldots, (X_n, \leq_n)]$ is uniformly computable from presentations of the WQOs $(X_i, \leq_i)$ $(i = 1, \ldots, n)$.

In the following sections, we proceed to prove that some of the most prominent WQO-preserving constructors are also ideally effective.

## 4.2 Sums of WQOs

We start with two simple constructions, disjoint sums and lexicographic sums of WQOs. They will be our first examples of ideally effective constructors and will set the template for later constructions.

### 4.2.1 Disjoint Sum

The *disjoint sum* $X_1 \sqcup X_2$ of two QOs $(X_1, \leq_1)$ and $(X_2, \leq_2)$ is the set $\{1\} \times X_1 \cup \{2\} \times X_2$, quasi-ordered by:

$$\langle i, x\rangle \leq_\sqcup \langle j, y\rangle \text{ iff } i = j \text{ and } x \leq_i y\ .$$

We use $X_\sqcup$ to denote $X_1 \sqcup X_2$ and generally use the $\sqcup$ subscript to identify operations associated with the structure $(X_\sqcup, \leq_\sqcup)$. This structure is obviously well quasi-ordered when $(X_1, \leq_1)$ and $(X_2, \leq_2)$ are.

We let the reader check the following characterization.

**Proposition 4.3** (Ideals of $X_1 \sqcup X_2$) *Given* $(X_1, \leq_1)$ *and* $(X_2, \leq_2)$ *two WQOs, the ideals of* $(X_1 \sqcup X_2, \leq_\sqcup)$ *are exactly the sets of the form* $I = \{i\} \times J$ *with* $i \in \{1, 2\}$ *and* $J$ *an ideal of* $X_i$.

Thus $(\mathit{Idl}(X_1 \sqcup X_2), \subseteq)$ is isomorphic to $(\mathit{Idl}(X_1), \subseteq) \sqcup (\mathit{Idl}(X_2), \subseteq)$.

Given data structures for $X_1$ and $X_2$, we use the natural data structure for $X_1 \sqcup X_2$. Moreover, Proposition 4.3 shows that ideals of the WQO $(X_1 \sqcup X_2, \leq_\sqcup)$ can similarly be represented using data structures for $\mathit{Idl}(X_1)$ and $\mathit{Idl}(X_2)$.

**Theorem 4.4** *With the above representations of elements and ideals, disjoint union is an ideally effective constructor.*

*Proof (Sketch)* Let $(X_1, \leq_1)$ and $(X_2, \leq_2)$ be two ideally effective WQOs.

In the following, we write $\bar{\imath}$ for $3 - i$ when $i \in \{1, 2\}$, so that $\{i, \bar{\imath}\} = \{1, 2\}$. We also abuse notation and, for a downwards-closed subset $D = \bigcup_a I_a$ of $X_i$, we write $\langle i, D\rangle$ to denote $\bigcup_a \langle i, I_a\rangle$, a downwards-closed subset of $X_\sqcup$ represented via ideals. Similarly, for an upwards-closed subset $U = \bigcup_a \uparrow_i x_a$ of $X_i$, we let $\langle i, U\rangle$ denote $\bigcup_a \uparrow_\sqcup \langle i, x_a\rangle$.

**(OD)**: the definition of $\leq_\sqcup$ is already an implementation.
**(ID)**: we use $\langle i, J\rangle \subseteq \langle i', J'\rangle \iff i = i' \wedge J \subseteq J'$.
**(PI)**: we use $\downarrow_\sqcup \langle i, x\rangle = \langle i, \downarrow_i x\rangle$ for $i \in \{1, 2\}$.
**(CF)**: we use $X_\sqcup \setminus \uparrow_\sqcup \langle i, x\rangle = \langle i, X_i \setminus \uparrow_i x\rangle \cup \langle \bar{\imath}, X_{\bar{\imath}}\rangle$. Note that this relies on (CF) for $X_i$ (to express $X_i \setminus \uparrow_i x$ as a union of ideals) and on (XI) for $X_{\bar{\imath}}$.
**(II)**: we rely on (II) for $X_1$ and $X_2$, using

$$\langle i, I\rangle \cap \langle j, J\rangle = \begin{cases} \langle i, I \cap J\rangle & \text{if } i = j, \\ \emptyset & \text{otherwise.} \end{cases}$$

Operations (CI) to complement ideals and (IF) to intersect filters are analogous.

Observe that the presentation of $(X_1 \sqcup X_2, \leq_\sqcup)$ described above is clearly computable from presentations for $(X_i, \leq_i)$ $(i = 1, 2)$. Notably, a filter (resp. ideal) decomposition of $X_1 \sqcup X_2$ is easily obtained by taking the union of filter (resp. ideal) decompositions of $X_1$ and $X_2$, thus establishing (XF) (resp. (XI)). □

### 4.2.2 Lexicographic Sums

The *lexicographic sum* $X_1 \oplus X_2$[3] of two QOs $(X_1, \leq_1)$, $(X_2, \leq_2)$ is the QO $(X_\oplus, \leq_\oplus)$ given by $X_\oplus = \{1\} \times X_1 \cup \{2\} \times X_2$ and

$$\langle i, x\rangle \leq_\oplus \langle j, y\rangle \text{ iff } i < j \text{ or } (i = j \text{ and } x \leq_i y) .$$

Therefore $X_1 \oplus X_2$ and $X_1 \sqcup X_2$ share the same underlying set. The ordering $\leq_\oplus$ is an extension of $\leq_\sqcup$ hence is a WQO too, when $(X_1, \leq_1)$ and $(X_2, \leq_2)$ are.

Again, the following characterization is easy to obtain.

**Proposition 4.5** (Ideals of $X_1 \oplus X_2$) *Given two WQOs* $(X_1, \leq_1)$ *and* $(X_2, \leq_2)$*, the ideals of* $X_1 \oplus X_2$ *are exactly the sets of the form* $\{1\} \times J_1$ *with* $J_1 \in \mathit{Idl}(X_1)$*, or of the form* $\{1\} \times X_1 \cup \{2\} \times J_2$ *with* $J_2 \in \mathit{Idl}(X_2)$.

Thus $(\mathit{Idl}(X_1 \oplus X_2), \subseteq)$ is isomorphic to $(\mathit{Idl}(X_1), \subseteq) \oplus (\mathit{Idl}(X_2), \subseteq)$, which leads to a simple data structure for the set of ideals[4] when $X_1$ and $X_2$ are effective.

**Theorem 4.6** *With the above representations, lexicographic union is an ideally effective constructor.*

*Proof (Sketch)* We reuse the abbreviations $\langle i, U\rangle$, $\langle i, D\rangle$, $\bar{\imath}$, ..., introduced for disjoint sums. Also, we only consider the case where both $X_1$ and $X_2$ are non-empty (the claim is trivial otherwise).

- **(OD)**: follows from the definition.
- **(ID)**: ideal inclusion can be tested as for the disjoint union of $\mathit{Idl}(X_1)$ and $\mathit{Idl}(X_2)$.
- **(PI)**: $\downarrow_\oplus \langle i, x\rangle$ is (represented by) $\langle i, \downarrow_i x\rangle$.
- **(CF)**: the complement $X_\oplus \setminus \uparrow_\oplus \langle i, x\rangle$ is (represented by) $\langle i, X_i \setminus \uparrow_i x\rangle$ except when $i = 2$ and $\uparrow_i x = X_2$, in which case $X_\oplus \setminus \uparrow_\oplus \langle 2, x\rangle$ is $\langle 1, X_1\rangle$.
- **(II)**: intersection of two ideals splits into two cases. First $\langle 1, I\rangle \cap \langle 2, J\rangle$ is (represented by) $\langle 1, I\rangle$ for ideals issued from different components in $X_\oplus$. For $\langle i, I\rangle \cap \langle i, J\rangle$, i.e., ideals issued from the same component, we use $\langle i, I \cap J\rangle$ except when $i = 2$ and $I \cap J = \emptyset$, in which case $\langle 2, I\rangle \cap \langle 2, J\rangle$ is $\langle 1, X_1\rangle$.

[3] A warning about notation: the lexicographic sum should not be confused with the natural sum of ordinals even if they are both denoted with ⊕. In particular, the lexicographic sum of ordinals is their usual addition.

[4] Note that with this representation, a pair $\langle i, J\rangle$ where $J \in \mathit{Idl}(X_i)$ denotes $\{1\} \times J$ when $i = 1$, and $\{1\} \times X_1 \cup \{2\} \times J$ —and not $\{2\} \times J$— when $i = 2$.

Procedures for the dual operations (CI) and (IF) are similar. Moreover, the presentation above is obviously computable from presentations for $(X_1, \leq_1)$ and $(X_2, \leq_2)$. Regarding (XI) and (XF), the ideal decomposition of $X_1 \oplus X_2$ is the ideal decomposition of $X_2$ and the filter decomposition of $X_1 \oplus X_2$ is the filter decomposition of $X_1$. □

## 4.3 Products of WQOs and Dickson's Lemma

Given two QOs $(X_1, \leq_1)$ and $(X_2, \leq_2)$, we define the *componentwise quasi-ordering* $\leq_\times$ on the Cartesian product $X_1 \times X_2$ by $\langle x_1, x_2\rangle \leq_\times \langle y_1, y_2\rangle \overset{\text{def}}{\Leftrightarrow} x_1 \leq_1 y_1 \wedge x_2 \leq_2 y_2$.

**Lemma 4.7** (Dickson's Lemma) *If $X_1$ and $X_2$ are WQOs, so is $X_1 \times X_2$.*

*Proof (Idea)* Given an infinite sequence in $X_1 \times X_2$, we extract an infinite sequence which is monotonic in the first component, and from that, an infinite sequence that is monotonic in the second component. □

The ideals of $(X_1 \times X_2, \leq_\times)$ are well known.

**Proposition 4.8** (Ideals of $X_1 \times X_2$) *Let $(X_1, \leq_1)$ and $(X_2, \leq_2)$ be two WQOs. A subset $I$ is an ideal of $X_1 \times X_2$ if, and only if, $I = I_1 \times I_2$ for some ideals $I_1, I_2$ of $X_1$ and $X_2$ respectively.*

*Proof* ($\Longleftarrow$): One checks that $I = I_1 \times I_2$ is non-empty, downwards-closed, and directed, when $I_1$ and $I_2$ are. For directedness, we consider two elements $\langle x_1, x_2\rangle$, $\langle y_1, y_2\rangle \in I$. Since $I_1$ is directed and contains $x_1, y_1$, it contains some $z_1$ with $x_1 \leq_1 z_1$ and $y_1 \leq_1 z_1$. Similarly $I_2$ contains some $z_2$ above $x_2$ and $y_2$ (wrt. $\leq_2$). Finally, $\langle z_1, z_2\rangle$ is in $I$, and above both $\langle x_1, y_1\rangle$ and $\langle x_2, y_2\rangle$.
($\Longrightarrow$): Consider $I \in \mathit{Idl}(X_1 \times X_2)$ and write $I_1$ and $I_2$ for its projections on $X_1$ and $X_2$. These projections are downwards-closed (since $I$ is), non-empty (since $I$ is) and directed (since $I$ is), hence they are ideals (in $X_1$ and $X_2$). We now show that $I_1 \times I_2 \subseteq I$. Consider an arbitrary $x_1 \in I_1$: since $I_1$ is the projection of $I$, there is some $y_2 \in X_2$ such that $\langle x_1, y_2\rangle \in I$. Similarly, for any $x_2 \in I_2$, there is some $y_1 \in X_1$ such that $\langle y_1, x_2\rangle \in I$. Since $I$ is directed, there is some $\langle z_1, z_2\rangle \in I$ with $\langle x_1, y_2\rangle \leq_\times \langle z_1, z_2\rangle$ and $\langle y_1, x_2\rangle \leq_\times \langle z_1, z_2\rangle$. But then $x_1 \leq_1 z_1$ and $x_2 \leq_2 z_2$. Thus $\langle x_1, x_2\rangle \in I$ since $I$ contains $\langle z_1, z_2\rangle$ and is downwards-closed. Hence $I = I_1 \times I_2$ and $I$ is a product of ideals. □

Thus $\mathit{Idl}(X_1 \times X_2, \subseteq)$ is isomorphic to $(\mathit{Idl}(X_1), \subseteq) \times (\mathit{Idl}(X_2), \subseteq)$. If $(X_1, \leq_1)$ and $(X_2, \leq_2)$ are ideally effective, we naturally represent elements of $X_1 \times X_2$ as pairs of elements of $X_1$ and $X_2$, and similarly ideals of $(X_1 \times X_2, \leq_\times)$ as pairs of ideals of $X_1$ and $X_2$. This is notably how we handled $\mathit{Idl}(\mathbb{N}^2)$ in Sect. 2.2.

**Theorem 4.9** *With the above representations, Cartesian product is an ideally effective constructor.*

*Proof* Let $D_1$ and $D_2$ be downwards-closed sets of $(X_1, \leq_1)$ and $(X_2, \leq_2)$ respectively, given by some ideal decompositions $D_1 = \bigcup_i I_{1,i}$ and $D_2 = \bigcup_j I_{2,j}$. Then $D_1 \times D_2$ is downwards-closed in $X_1 \times X_2$, and it decomposes as $\bigcup_i \bigcup_j I_{1,i} \times I_{2,j}$ since products distribute over unions. The same reasoning holds for upwards-closed sets and their filter decompositions and we rely on these properties in the following explanations.

**(OD)**: the ordering $\leq_\times$ is obviously decidable.
**(ID)**: $I_1 \times I_2 \subseteq J_1 \times J_2$ iff $I_1 \subseteq J_1$ and $I_2 \subseteq J_2$ (exercise: the nonemptiness of ideals is required here).
**(PI)**: $\downarrow\langle x_1, x_2\rangle = \downarrow x_1 \times \downarrow x_2$.
**(II)**: to compute intersections, use $(I_1 \times I_2) \cap (I_1' \times I_2') = (I_1 \cap I_1') \times (I_2 \cap I_2')$, and build the product of downwards-closed sets as explained above.
**(CF)**: to complement filters, use $(X_1 \times X_2) \setminus \uparrow_\times\langle x_1, x_2\rangle = \big[(X_1 \setminus \uparrow x_1) \times X_2\big] \cup \big[X_1 \times (X_2 \setminus \uparrow x_2)\big]$ and build products of downwards-closed sets.

Procedures for the remaining operations are obtained similarly. Note that here too, the presentation above is computable from presentations for $(X_1, \leq_1)$ and $(X_2, \leq_2)$. Notably, a filter and ideal decomposition of $X_1 \times X_2$ is easily obtained from decompositions of $X_1$ and $X_2$, by distributing products over unions. □

## 4.4 Sequence Extensions of WQOs and Higman's Lemma

Given a QO $(X, \leq)$, we denote by $X^*$ the *sequence extension* of $X$, i.e., the set of all finite sequences over $X$, often called *words* when $X$ is an alphabet. We write $\epsilon$ for the empty (zero-length) sequence, and denote multiplicatively the concatenation of sequences, as $\boldsymbol{uv}$ or $\boldsymbol{u} \cdot \boldsymbol{v}$. Elements of $X^*$ will be denoted in bold font, such as $\boldsymbol{u}, \boldsymbol{v}, \ldots$, while elements of $X$ are denoted $x, y, \ldots$. In particular, if $x \in X$, then $\boldsymbol{x} \in X^*$ denotes the sequence of length one containing only the symbol $x$.

The set $X^*$ is often quasi-ordered with *Higman's quasi-ordering* $\leq_*$, also known as the *sequence embedding* quasi-ordering, defined by $\boldsymbol{u} = x_1 \cdots x_n \leq_* \boldsymbol{v} = y_1 \cdots y_m$ $\stackrel{\text{def}}{\Leftrightarrow}$ there are $n$ indices $1 \leq p_1 < p_2 < \cdots < p_n \leq m$ such that $x_i \leq y_{p_i}$ for each $i = 1, \ldots, n$. In other words, and writing $[n]$ for the set $\{1, \cdots, n\}$, there is a strictly increasing mapping $p$ from $[n]$ to $[m]$ such that $x_i \leq y_{p(i)}$. Such a mapping will be called a *witness* of $\boldsymbol{u} \leq_* \boldsymbol{v}$. Equivalently, $\boldsymbol{u} \leq_* \boldsymbol{v}$ if $\boldsymbol{v}$ contains a length $n$ subsequence $\boldsymbol{v}' = y_{p_1} \cdots y_{p_n}$ such that $\boldsymbol{u} \leq_\times \boldsymbol{v}'$ using the product ordering from Sect. 4.3.

The structure $(X^*, \leq_*)$ is sometimes called the *Higman extension* of $(X, \leq)$. This constructor preserves WQO: this is Higman's Lemma [33].

Showing that this constructor is ideally effective requires some work and Sect. 4.4 is one of the longest in this chapter. This is justified by the importance of this construction. Being generic, our algorithms apply to non-trivial instances such as $(\mathbb{N}^k)^*$—used in Timed-arc nets [30], in data nets [40], for runs of Vector Addition Systems [44]—, or $(\Sigma^*)^k \times (\Sigma^*)^*$—used in Dynamic Lossy Channel

Systems [1]—, or to even richer settings such as the Priority Channel Systems and the Higher-Order Channel Systems of [28]. The algorithms for $(X^*, \leq_*)$ are also invoked when showing ideal effectiveness of many WQOs derived from $X^*$.

Before we study the ideals of $X^*$, let us first lift the concatenation of sequences to sets of sequences: the product (for concatenation) of two sets of sequences $\boldsymbol{U}, \boldsymbol{V} \subseteq X^*$ is denoted $\boldsymbol{U} \cdot \boldsymbol{V} \stackrel{\text{def}}{=} \{\boldsymbol{u} \cdot \boldsymbol{v} \mid \boldsymbol{u} \in \boldsymbol{U}, \boldsymbol{v} \in \boldsymbol{V}\}$. A useful property of $(X^*, \leq_*)$ is that the concatenation of downwards-closed sets distributes over intersection:

**Lemma 4.10** *Let* $\boldsymbol{D}_1, \boldsymbol{D}_2, \boldsymbol{D} \in \mathit{Down}(X^*)$. *Then* $(\boldsymbol{D}_1 \cap \boldsymbol{D}_2) \cdot \boldsymbol{D} = (\boldsymbol{D}_1 \cdot \boldsymbol{D}) \cap (\boldsymbol{D}_2 \cdot \boldsymbol{D})$.

*Proof* The left-to-right inclusion is obvious. For the right-to-left inclusion, let $\boldsymbol{w} \in \boldsymbol{D}_1 \cdot \boldsymbol{D} \cap \boldsymbol{D}_2 \cdot \boldsymbol{D}$. Then $\boldsymbol{w} = \boldsymbol{u}_1\boldsymbol{v}_1$ for some $\boldsymbol{u}_1 \in \boldsymbol{D}_1$ and $\boldsymbol{v}_1 \in \boldsymbol{D}$. Also, $\boldsymbol{w} = \boldsymbol{u}_2\boldsymbol{v}_2$ for some $\boldsymbol{u}_2 \in \boldsymbol{D}_2$ and $\boldsymbol{v}_2 \in \boldsymbol{D}$. One of $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ is a prefix of the other. Assume $\boldsymbol{u}_1$ is a prefix of $\boldsymbol{u}_2$ (the other case is analogous). Since $\boldsymbol{D}_2$ is downwards-closed and $\boldsymbol{u}_1 \leq_* \boldsymbol{u}_2$, $\boldsymbol{u}_1 \in \boldsymbol{D}_2$. Thus, $\boldsymbol{u}_1 \in \boldsymbol{D}_1 \cap \boldsymbol{D}_2$ and $\boldsymbol{w} = \boldsymbol{u}_1\boldsymbol{v}_1 \in (\boldsymbol{D}_1 \cap \boldsymbol{D}_2) \cdot \boldsymbol{D}$. □

The structure of ideals of $(X^*, \leq_*)$ is given in [35] where the following theorem is proved. An alternate proof is presented in Sect. 4.4.2.

**Theorem 4.11** (Ideals of $\boldsymbol{X^*}$) *Given a WQO* $(X, \leq)$, *the ideals of* $(X^*, \leq_*)$ *are exactly the finite products of atoms, of the form* $\boldsymbol{P} = \boldsymbol{A}_1 \cdot \boldsymbol{A}_2 \cdots \boldsymbol{A}_n$ *where* atoms *are:*

– *any set of the form* $\boldsymbol{A} = D^*$, *for* $D \in \mathit{Down}(X)$,
– *any set of the form* $\boldsymbol{A} = I + \epsilon \stackrel{\mathit{def}}{=} \{\boldsymbol{x} \mid x \in I\} \cup \{\epsilon\}$, *for* $I \in \mathit{Idl}(X)$.

### 4.4.1 Ideal Effectiveness

The elements of $X^*$ will be represented in the natural way, e.g., via lists of elements of $X$ (assuming a data structure for $X$). When $(X, \leq)$ is ideally effective, Theorem 4.11 leads to a natural data structure for ideals of $X^*$, as lists of atoms, where the representation of atoms is directly inherited from those for $\mathit{Idl}(X)$ and $\mathit{Down}(X)$.

**Theorem 4.12** *With the above representations, the sequence extension is an ideally effective constructor.*

*Proof* Let $(X, \leq)$ be an ideally effective WQO.

**(OD)**: deciding $\leq_*$ over $X^*$ reduces to comparing elements of $X$, e.g. by looking for a *leftmost embedding*.
**(PI)**: given a finite sequence $\boldsymbol{u} = x_1 \cdots x_n$, the principal ideal $\downarrow \boldsymbol{u}$ is represented by the product $(\downarrow x_1 + \epsilon) \cdots (\downarrow x_n + \epsilon)$.

Procedures for the remaining operations required by Definition 3.1 are more elaborate, and we therefore introduce a lemma for each one. This series of lemmas concludes the proof since the fact that a presentation of $(X^*, \leq_*)$ can be uniformly computed from a presentation of $(X, \leq)$ will be clear. As for (XF), the filter decomposition of $X^* = \uparrow \boldsymbol{\epsilon}$ is given by the empty sequence (and does not depend on $X$), while for (XI) we note that $X^*$ is already an ideal made of a single atom. □

Subsequently, $(X, \leq)$ denotes an ideally effective WQO. We begin with ideal inclusion. A similar procedure was already obtained by Abdulla et al. in the case where $X$ is a finite alphabet with equality [3].

**Lemma 4.13** (ID) *Inclusion between ideals of* $(X^*, \leq_*)$ *can be tested using a linear number of inclusion tests between downwards-closed sets of X, using a version of* leftmost embedding *search. The following equations implicitly describe an algorithm deciding inclusion by induction on the length, or number of atoms, of ideals:*

1. *Atoms are compared as follows:*

$$(I_1 + \epsilon) \subseteq (I_2 + \epsilon) \iff I_1 \subseteq I_2, \tag{5.1}$$
$$(I + \epsilon) \subseteq D^* \iff I \subseteq D, \tag{5.2}$$
$$D_1^* \subseteq D_2^* \iff D_1 \subseteq D_2, \tag{5.3}$$
$$D^* \subseteq (I + \epsilon) \iff D = \emptyset. \tag{5.4}$$

2. *For any ideal* $\boldsymbol{P}$*:* $\boldsymbol{\epsilon} \subseteq \boldsymbol{P}$.
3. *For any ideal* $\boldsymbol{P}$ *and atom* $\boldsymbol{A}$*:* $\boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{\epsilon} \iff \boldsymbol{A} = \emptyset^* \wedge \boldsymbol{P} \subseteq \boldsymbol{\epsilon}$.
4. *Finally, for all atoms* $\boldsymbol{A}$ *and* $\boldsymbol{B}$*, and ideals* $\boldsymbol{P}$ *and* $\boldsymbol{Q}$*:*
   (a) *if* $\boldsymbol{A} \not\subseteq \boldsymbol{B}$ *then:*
   $$\boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{B} \cdot \boldsymbol{Q} \iff \boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{Q}\,;$$
   (b) *if* $\boldsymbol{A} \subseteq \boldsymbol{B}$ *as in* (5.1)*, i.e.,* $\boldsymbol{A} = (I_1 + \epsilon)$*,* $\boldsymbol{B} = (I_2 + \epsilon)$ *for some* $I_1, I_2 \in Idl(X)$*, then:*
   $$\boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{B} \cdot \boldsymbol{Q} \iff \boldsymbol{P} \subseteq \boldsymbol{Q}\,;$$
   (c) *if* $\boldsymbol{A} \subseteq \boldsymbol{B}$ *as in any of Eqs.* (5.2) *to* (5.4)*, then:*
   $$\boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{B} \cdot \boldsymbol{Q} \iff \boldsymbol{P} \subseteq \boldsymbol{B} \cdot \boldsymbol{Q}\,.$$

*Proof* The first three cases are trivial. We concentrate on the fourth one.

4a Since $\boldsymbol{B}$ contains $\epsilon$, $\boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{Q}$ implies $\boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{B} \cdot \boldsymbol{Q}$. Conversely, let $\boldsymbol{u} \in \boldsymbol{A}$ and $\boldsymbol{v} \in \boldsymbol{P}$, so that $\boldsymbol{uv} \in \boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{B} \cdot \boldsymbol{Q}$. Assuming $\boldsymbol{A} \not\subseteq \boldsymbol{B}$, there exists $\boldsymbol{w}' \in \boldsymbol{A} \setminus \boldsymbol{B}$ and by directedness, there exists a word $\boldsymbol{w} \in \boldsymbol{A}$ such that $\boldsymbol{w} \geq_* \boldsymbol{w}', \boldsymbol{u}$. In particular, $\boldsymbol{w}$ is in $\boldsymbol{A} \setminus \boldsymbol{B}$ and $\boldsymbol{w} \geq_* \boldsymbol{u}$.
If $\boldsymbol{A} = I + \epsilon$ for some $I \in Idl(X)$, then $\boldsymbol{w}$ is of length at most one. Since $\boldsymbol{w}' \notin \boldsymbol{B}$, in particular $\boldsymbol{w}' \neq \epsilon$, so $\boldsymbol{w}$ is of length exactly one. Also, since $\boldsymbol{w} \notin \boldsymbol{B}$, the word

$\boldsymbol{wv}$, which is in $\boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{B} \cdot \boldsymbol{Q}$ has to actually be in $\boldsymbol{Q}$. Since $\boldsymbol{Q}$ is downwards-closed, $\boldsymbol{uv} \in \boldsymbol{Q}$.
Otherwise, $\boldsymbol{A} = D^*$ for some $D \in Down(X)$. In this case, $\boldsymbol{ww} \in \boldsymbol{A}$ and thus $\boldsymbol{wwv} \in \boldsymbol{B} \cdot \boldsymbol{Q}$. We factor $\boldsymbol{wwv}$ as $\boldsymbol{v}_1\boldsymbol{v}_2$ with $\boldsymbol{v}_1 \in \boldsymbol{B}$ and $\boldsymbol{v}_2 \in \boldsymbol{Q}$. Since $\boldsymbol{w}$ is not in $\boldsymbol{B}$, no word of which $\boldsymbol{w}$ is a prefix is in $\boldsymbol{B}$ either, and that implies that $\boldsymbol{v}_1$ is a proper prefix of $\boldsymbol{w}$, and that $\boldsymbol{v}_2$ has $\boldsymbol{wv}$ as a suffix. In particular, $\boldsymbol{v}_2 \geq_* \boldsymbol{wv}$. Recalling that $\boldsymbol{wv} \geq_* \boldsymbol{uv}$ and that $\boldsymbol{Q}$ is downwards-closed, $\boldsymbol{uv}$ is in $\boldsymbol{Q}$.

4b Here also, the right-to-left implication is trivial. Conversely, assume $\boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{B} \cdot \boldsymbol{Q}$ and $\boldsymbol{A} = (I_1 + \epsilon)$ and $\boldsymbol{B} = (I_2 + \epsilon)$ for some $I_1 \subseteq I_2 \in Idl(X)$. Let $\boldsymbol{u} \in \boldsymbol{P}$. Pick $x \in I_1$: $\boldsymbol{xu} \in \boldsymbol{A} \cdot \boldsymbol{P}$, thus $\boldsymbol{xu} \in \boldsymbol{B} \cdot \boldsymbol{Q}$. Therefore, $\boldsymbol{u} \in \boldsymbol{Q}$ since sequences of $\boldsymbol{B}$ have length at most one.

4c The left-to-right implication is trivial, since $\epsilon \in \boldsymbol{A}$. For the other implication, we consider some $\boldsymbol{u} \in \boldsymbol{A}$ and $\boldsymbol{v} \in \boldsymbol{P}$, and we have to show that $\boldsymbol{uv} \in \boldsymbol{B} \cdot \boldsymbol{Q}$. Since $\boldsymbol{P} \subseteq \boldsymbol{B} \cdot \boldsymbol{Q}$, we can factor $\boldsymbol{v}$ as $\boldsymbol{v}_1\boldsymbol{v}_2$ with $\boldsymbol{v}_1 \in \boldsymbol{B}$ and $\boldsymbol{v}_2 \in \boldsymbol{Q}$. We claim that $\boldsymbol{uv}_1 \in \boldsymbol{B}$: if $\boldsymbol{B} = D^*$ is an atom of the second kind, the claim follows from $\boldsymbol{A} \subseteq \boldsymbol{B}$; if $\boldsymbol{B} = I + \epsilon$ is an atom of the first kind, then we are in case (5.4), $\boldsymbol{A}$ is $\emptyset^*$, and $\boldsymbol{u} = \epsilon$. With $\boldsymbol{uv}_1 \in \boldsymbol{B}$ we have $\boldsymbol{uv} \in \boldsymbol{B} \cdot \boldsymbol{Q}$ as needed. □

The next lemma deals with the complementation of filters:

**Lemma 4.14** (CF) *Given* $\boldsymbol{w} \in X^*$*, the downwards-closed set* $X^* \setminus \uparrow \boldsymbol{w}$ *can be computed inductively using the following equations:*

$$X^* \setminus \uparrow \epsilon = \emptyset \text{ (empty union)}, \tag{6}$$

$$X^* \setminus \uparrow x\boldsymbol{v} = \begin{cases} (X \setminus \uparrow x)^* & \text{if } \boldsymbol{v} = \epsilon, \\ (X \setminus \uparrow x)^* \cdot (X + \epsilon) \cdot (X^* \setminus \uparrow \boldsymbol{v}) & \text{otherwise.} \end{cases} \tag{7}$$

Note that $X$ might not be an ideal, in which case $X + \epsilon$ is not an atom in Eq. (7). In this case, one has to first get the ideal decomposition $X = \bigcup_i I_i$ from a presentation of $(X, \leq)$ and use distributivity of concatenation over unions.

In the commonly encountered case where $X$ is a finite alphabet, ordered by equality, there is no need to distribute, and indeed, the complement of a filter is always an ideal. More precisely, if $X = \{a_1, \ldots, a_n\}$ is a finite alphabet under equality, then one checks easily that $(X \setminus \uparrow a_i)^* \cdot (X + \epsilon) = \{a_j \mid j \neq i\}^* \cdot (a_i + \epsilon)$. It follows that complement of filters are ideals in this case.

*Remark 4.15* Kabil and Pouzet [35] use the following (equivalent) expression to complement filters:

$$X^* \setminus \uparrow xy\boldsymbol{w} = (X \setminus \uparrow x)^* \cdot [\downarrow(\uparrow x \cap \uparrow y) + \epsilon] \cdot (X^* \setminus \uparrow y\boldsymbol{w})\,. \tag{8}$$

We used a different formula because, in general, our setting does not guarantee that the expression $\downarrow U$ is computable for $U \in Up(X)$, even in the particular case where $U = \uparrow x \cap \uparrow y$. It is fair to mention that Kabil and Pouzet make no claim on computability.

Still, Eq. (8) is interesting when $X$ is a finite alphabet since then the expression $\uparrow x \cap \uparrow y$ either denotes the empty set or $(x + \epsilon)$, depending on whether $x$ and $y$ coincide. Therefore, using Eq. (8), one directly obtains an ideal written in canonical form (a notion defined below, in Sect. 4.4.3). □

*Proof* (*of Lemma* 4.14) We only prove the second case of Eq. (7) since the other equalities are obvious.
($\supseteq$): Let $\boldsymbol{w}' = \boldsymbol{u}\boldsymbol{y}\boldsymbol{w}$ with $\boldsymbol{u} \in (X \setminus \uparrow x)^*$, $\boldsymbol{y} \in X + \epsilon$ and $\boldsymbol{w} \in (X^* \setminus \uparrow \boldsymbol{v})$. Thus $\boldsymbol{v} \not\leq_* \boldsymbol{w}$. Since $\boldsymbol{y}$ has length at most 1, we deduce $x\boldsymbol{v} \not\leq_* \boldsymbol{y}\boldsymbol{w}$. Since all elements in $\boldsymbol{u}$ are taken from $X \setminus \uparrow x$, we further have $x\boldsymbol{v} \not\leq_* \boldsymbol{u}\boldsymbol{y}\boldsymbol{v}$. Therefore $\boldsymbol{w'} \in X^* \setminus \uparrow x\boldsymbol{v}$.
($\subseteq$): Let $\boldsymbol{w}' \notin \uparrow x\boldsymbol{v}$. Then either $\boldsymbol{w}' \in (X \setminus \uparrow x)^*$, or we can write $\boldsymbol{w}' = \boldsymbol{u}y\boldsymbol{w}$ with $\boldsymbol{u} \in (X \setminus \uparrow x)^*$ and $y \geq x$. Moreover, $\boldsymbol{w} \notin \uparrow \boldsymbol{v}$, since otherwise $x\boldsymbol{v} \leq_* y\boldsymbol{w} \leq_* \boldsymbol{u}y\boldsymbol{w} = \boldsymbol{w}'$. Therefore, $\boldsymbol{w}' \in (X \setminus \uparrow x)^* \cdot X \cdot (X^* \setminus \uparrow \boldsymbol{v})$. Joining the two cases, and since $\epsilon \in (X^* \setminus \uparrow \boldsymbol{v})$, we obtain the required $\boldsymbol{w}' \in (X \setminus \uparrow x)^* \cdot (X + \epsilon) \cdot (X^* \setminus \uparrow \boldsymbol{v})$. □

We now show how to intersect ideals:

**Lemma 4.16** (II) *The intersection of two ideals of* $(X^*, \leq_*)$ *can be computed inductively using the following equations:*

$$\boldsymbol{\epsilon} \cap \boldsymbol{Q} = \boldsymbol{P} \cap \boldsymbol{\epsilon} = \boldsymbol{\epsilon}\ , \tag{9}$$

$$D_1^* \cdot \boldsymbol{P} \cap D_2^* \cdot \boldsymbol{Q} = (D_1 \cap D_2)^* \cdot \begin{bmatrix} (D_1^* \cdot \boldsymbol{P}) \cap \boldsymbol{Q} \\ \cup\ \boldsymbol{P} \cap (D_2^* \cdot \boldsymbol{Q}) \end{bmatrix}, \tag{10}$$

$$(I_1 + \epsilon) \cdot \boldsymbol{P} \cap (I_2 + \epsilon) \cdot \boldsymbol{Q} = \begin{bmatrix} ((I_1 + \epsilon) \cdot \boldsymbol{P}) \cap \boldsymbol{Q} \\ \cup\ \boldsymbol{P} \cap ((I_2 + \epsilon) \cdot \boldsymbol{Q}) \\ \cup\ \big((I_1 \cap I_2) + \epsilon\big) \cdot (\boldsymbol{P} \cap \boldsymbol{Q}) \end{bmatrix}, \tag{11}$$

$$D^* \cdot \boldsymbol{P} \cap (I + \epsilon) \cdot \boldsymbol{Q} = \begin{bmatrix} \boldsymbol{P} \cap ((I + \epsilon) \cdot \boldsymbol{Q}) \\ \cup\ \big((D \cap I) + \epsilon\big) \cdot ((D^* \cdot \boldsymbol{P}) \cap \boldsymbol{Q}) \end{bmatrix}. \tag{12}$$

Here also, some shortcuts are used. For instance, the intersection of two ideals need not be an ideal. Therefore, $(I_1 \cap I_2) + \epsilon$ in Eq. (11) might not be an ideal. As before, by decomposing downwards-closed sets as union of ideals, and distributing concatenations over unions, one can compute the actual ideal decomposition of the intersection of two ideals of $(X^*, \leq_*)$.

*Proof* (*of Lemma* 4.16) Equation (9) is obviously correct. The other right-to-left inclusions are easily checked using Lemma 4.13. For the left-to-right inclusions:

**Equation** (10): Let $\boldsymbol{u} \in D_1^* \cdot \boldsymbol{P} \cap D_2^* \cdot \boldsymbol{Q}$. Let $\boldsymbol{v}$ be the longest prefix of $\boldsymbol{u}$ which is in $D_1^*$. Without loss of generality, we assume that the longest prefix of $\boldsymbol{u}$ which is in $D_2^*$ is longer than $|\boldsymbol{v}|$, and thus can be written $\boldsymbol{v}\boldsymbol{w}$ for some $\boldsymbol{w} \in D_2^*$. Moreover, there exists $\boldsymbol{t} \in X^*$ so that $\boldsymbol{u} = \boldsymbol{v}\boldsymbol{w}\boldsymbol{t}$. We have $\boldsymbol{v} \in (D_1 \cap D_2)^*$, $\boldsymbol{w}\boldsymbol{t} \in \boldsymbol{P}$ and $\boldsymbol{t} \in \boldsymbol{Q}$. Therefore, $\boldsymbol{w}\boldsymbol{t} \in \boldsymbol{P} \cap D_2^* \cdot \boldsymbol{Q}$.
**Equation** (11): Consider any word in $(I_1 + \epsilon) \cdot \boldsymbol{P} \cap (I_2 + \epsilon) \cdot \boldsymbol{Q}$. If it is empty, it is also in the right-hand side of Eq. (11), so we assume that it of the form $x\boldsymbol{u}$. Depending on whether $x \in I_1 \setminus I_2$, $x \in I_2 \setminus I_1$ or $x \in I_1 \cap I_2$, $\boldsymbol{u}$ is easily proved to

be in $((I_1 + \epsilon) \cdot \boldsymbol{P}) \cap \boldsymbol{Q}$, in $\boldsymbol{P} \cap ((I_2 + \epsilon) \cdot \boldsymbol{Q})$, or in $((I_1 \cap I_2) + \epsilon) \cdot (\boldsymbol{P} \cap \boldsymbol{Q})$. If $x$ is neither in $I_1$ nor $I_2$, then $x\boldsymbol{u}$ belongs to all three sets.

**Equation** (12): This is similar, combining arguments from the previous two cases. □

We now turn to intersecting filters:

**Lemma 4.17** (IF) *The intersection of two filters can be computed inductively using the following equations:*

$$\uparrow \boldsymbol{v} \cap \uparrow \epsilon = \uparrow \epsilon \cap \uparrow \boldsymbol{v} = \uparrow \boldsymbol{v}\,, \tag{13}$$

$$\uparrow x\boldsymbol{v} \cap \uparrow y\boldsymbol{w} = \begin{bmatrix} (\uparrow \boldsymbol{x}) \cdot (\uparrow \boldsymbol{v} \cap \uparrow y\boldsymbol{w}) \ \cup \ (\uparrow \boldsymbol{y}) \cdot (\uparrow x\boldsymbol{v} \cap \uparrow \boldsymbol{w}) \\ \cup \, (\uparrow_X x \cap \uparrow_X y) \cdot (\uparrow \boldsymbol{v} \cap \uparrow \boldsymbol{w}) \end{bmatrix}, \tag{14}$$

*where $\boldsymbol{v}, \boldsymbol{w} \in X^*$ and $x, y \in X$. The actual filter decomposition in the last equation is obtained using $(\uparrow \boldsymbol{u}) \cdot (\uparrow \boldsymbol{u}') = \uparrow(\boldsymbol{u}\boldsymbol{u}')$ and distributivity over unions.*

*Proof* Equation (13) and the "$\supseteq$" half of Eq. (14) are obvious. For the remaining "$\subseteq$" half, we consider $\boldsymbol{u} \in \uparrow x\boldsymbol{v} \cap \uparrow y\boldsymbol{w}$. Let us write $\boldsymbol{u}$ as $\boldsymbol{u} = \boldsymbol{u}_1 z \boldsymbol{u}_2$ where $z\boldsymbol{u}_2$ is the shortest suffix of $\boldsymbol{u}$ in $\uparrow x\boldsymbol{v} \cap \uparrow y\boldsymbol{w}$—this suffix cannot be empty since it contains $x\boldsymbol{v}$ and $y\boldsymbol{w}$ as embedded sequences. Note that $z$ must be above $x$ or $y$ in $X$, otherwise $\boldsymbol{u}_2$ would be a shorter suffix of $\boldsymbol{u}$ in $\uparrow x\boldsymbol{v} \cap \uparrow y\boldsymbol{w}$. One now considers whether $z$ is above $x$, $y$, or both, and picks the corresponding summand in the right of Eq. (14). □

Finally, we focus on the complementation of ideals. This operation requires more work, and is decomposed in several lemmas. We first show how to complement atoms, and then how to complement products of atoms.

- If $D \subseteq X$ is downwards-closed, then $X^* \setminus D^*$, also written $\neg D^*$, consists of all sequences having at least one element not in $D$. One first computes $X \setminus D = \uparrow a_1 \cup \cdots \cup \uparrow a_n$, using (CI) for $X$. Then $\neg D^* = \uparrow_{X^*} \boldsymbol{a}_1 \cup \cdots \cup \uparrow_{X^*} \boldsymbol{a}_n$.
- If $I \subseteq X$ is an ideal, $\neg(I + \epsilon)$ consists of all sequences of length at least 2, as well as all sequences having an element not in $I$. The latter is obtained as in the previous case, by computing $X \setminus I = \uparrow b_1 \cup \cdots \cup \uparrow b_m$ in $X$. The former is $\uparrow_{X^*}(X \cdot X)$, easily computed in a similar way using (XF) for $X$.

We now consider products $\boldsymbol{A}_1 \cdots \boldsymbol{A}_n$ of atoms. We know how to compute $\boldsymbol{U}_i = \neg \boldsymbol{A}_i$. One has $\neg(\boldsymbol{A}_1 \cdots \boldsymbol{A}_n) = \neg(\neg \boldsymbol{U}_1 \cdots \neg \boldsymbol{U}_n)$, and this motivates the following definition:

**Definition 4.18** Define the operator $\odot$: $Up(X^*) \times Up(X^*) \to Up(X^*)$ as $\boldsymbol{U} \odot \boldsymbol{V} := \neg(\neg \boldsymbol{U} \cdot \neg \boldsymbol{V})$.

Note that $\boldsymbol{U} \odot \boldsymbol{V}$ is upwards-closed when $\boldsymbol{U}$ and $\boldsymbol{V}$ are. The operation $\odot$ is easily shown associative using the associativity of the product, thus $\boldsymbol{U}_1 \odot \cdots \odot \boldsymbol{U}_n = \neg(\neg \boldsymbol{U}_1 \cdot \cdots \cdot \neg \boldsymbol{U}_n)$. The previous relation becomes $\neg(\boldsymbol{A}_1 \cdots \boldsymbol{A}_n) = \boldsymbol{U}_1 \odot$

$\cdots \odot \boldsymbol{U}_n$, and it only remains to show that $\odot$ is computable on upwards-closed sets. In what follows, we will often use the following obvious characterization: $\boldsymbol{w} \in \boldsymbol{S} \odot \boldsymbol{T}$ if and only if for all factorizations $\boldsymbol{w} = \boldsymbol{w}_1\boldsymbol{w}_2$, $\boldsymbol{w}_1 \in \boldsymbol{S}$ or $\boldsymbol{w}_2 \in \boldsymbol{T}$.

We first show that $\odot$ is computable on principal filters, then we show how to complement ideals.

**Lemma 4.19** *On principal filters, $\odot$ can be computed using the following equations:*

$$\uparrow \boldsymbol{v} \odot \uparrow \epsilon = \uparrow \epsilon \odot \uparrow \boldsymbol{v} = X^* \,, \tag{15}$$

$$\uparrow \boldsymbol{v}a \odot \uparrow b\boldsymbol{w} = \uparrow(\boldsymbol{v}ab\boldsymbol{w}) \cup (\uparrow \boldsymbol{v}) \cdot (\uparrow_X a \cap \uparrow_X b) \cdot (\uparrow \boldsymbol{w}) \,, \tag{16}$$

*where $\boldsymbol{v}, \boldsymbol{w} \in X^*$ and $a, b \in X$.*

*Proof* Equation (15) is clear. We concentrate on Eq. (16):

($\supseteq$) If $\boldsymbol{u} \geq_* \boldsymbol{v}ab\boldsymbol{w}$, then for every factorization of $\boldsymbol{u} = \boldsymbol{u}_1\boldsymbol{u}_2$, the left factor $\boldsymbol{u}_1$ is above $\boldsymbol{v}a$, or the right factor $\boldsymbol{u}_2$ is above $b\boldsymbol{w}$, and thus $\boldsymbol{u} \in \uparrow \boldsymbol{v}a \odot \uparrow b\boldsymbol{w}$. If $\boldsymbol{u} \geq_* \boldsymbol{v}c\boldsymbol{w}$, where $c \in X$ is such that $c \geq a$ and $c \geq b$, then in every factorization of $\boldsymbol{u}$ as $\boldsymbol{u}_1\boldsymbol{u}_2$, $c$ appears either in the left factor $\boldsymbol{u}_1$ or in the right factor $\boldsymbol{u}_2$, and this suffices to show that either $\boldsymbol{u} \geq_* \boldsymbol{v}a$ or $\boldsymbol{u} \geq_* b\boldsymbol{w}$.

($\subseteq$) Let $\boldsymbol{u} \in (\uparrow \boldsymbol{v}a) \odot (\uparrow b\boldsymbol{w})$. From the factorizations $\boldsymbol{u} = \boldsymbol{u} \cdot \epsilon$ and $\boldsymbol{u} = \epsilon \cdot \boldsymbol{u}$ we get $\boldsymbol{v}a \leq_* \boldsymbol{u}$ and $b\boldsymbol{w} \leq_* \boldsymbol{u}$. Consider the shortest prefix of $\boldsymbol{u}$ above $\boldsymbol{v}a$ and the shortest suffix above $b\boldsymbol{w}$. These factors cannot have an overlap of length $\geq 2$, otherwise splitting $\boldsymbol{u}$ in the middle of the overlap would provide a shorter factor above $\boldsymbol{v}a$ or one above $b\boldsymbol{w}$, contradicting our assumption. If the factors do not overlap, we get $\boldsymbol{u} \geq_* \boldsymbol{v}ab\boldsymbol{w}$. If they overlap, necessarily over a single letter $c \in X$, we write $\boldsymbol{u} = \boldsymbol{u}_1 c\boldsymbol{u}_2$. Then $\boldsymbol{u}_1 \geq_* \boldsymbol{v}$, $c \geq a$, $c \geq b$, and $\boldsymbol{u}_2 \geq_* \boldsymbol{w}$, which proves the statement. □

**Lemma 4.20** (CI) *Complementing ideals of $(X^*, \leq_*)$ is computable.*

*Proof* Given an ideal $\boldsymbol{P} = \boldsymbol{A}_1 \cdots \boldsymbol{A}_n$, its complement is $\neg\boldsymbol{P} = \neg\boldsymbol{A}_1 \odot \cdots \odot \neg\boldsymbol{A}_n$. Using the procedure to complement downwards-closed sets of $(X, \leq)$, we can write each $\neg\boldsymbol{A}_i$ as a union of filters. Since $\odot$ distributes over unions of upwards-closed sets (from Lemma 4.10 by duality), we can write $\neg\boldsymbol{P}$ as a finite union of sets of the form $F_1 \odot F_2 \odot \cdots \odot F_n$, where the $F_i$'s are filters. Finally, Lemma 4.19 allows us to reduce these expressions to a finite union of filters. □

### 4.4.2 A Proof of Theorem 4.11

One direction of the theorem is easy to check: products of atoms are indeed ideals (downwards-closed and directed) of $(X^*, \leq_*)$. For the other direction, consider an arbitrary ideal $I$ of $(X^*, \leq_*)$. Its complement is upwards-closed, hence can be written $\neg I = \bigcup_{i<n} F_i$ for some filters $F_1, \ldots, F_n$. Therefore,

$$I = \neg \bigcup_{i<n} F_i = \bigcap_{i<n} \neg F_i .$$

Now, since any $\neg F_i$ is a finite union of products of atoms (see Lemma 4.14), by distributing the intersection over the unions, we are left with a finite union of finite intersections of products of atoms. Since these intersections can be decomposed as finite unions of products of atoms (see Lemma 4.16), we have decomposed the ideal $I$ into a finite union of products of atoms. Since products of atoms are ideals (cf. first direction), and since ideals are prime subsets (by Proposition 2.10), we obtain that $I$ is actually equal to one of those products of atoms (by Lemma 2.5).

This proof highlights a general technique for identifying the ideals of some WQO: if we have some subclass $\mathcal{J}$ of the ideals such that the complement of any filter can be written as a finite union of ideals of $\mathcal{J}$, and the intersection of any two ideals of $\mathcal{J}$ can be written as a finite union of ideals of $\mathcal{J}$, then $\mathcal{J}$ is the class of all ideals.

### 4.4.3 Uniqueness of Ideal Representation

Writing ideals as products of atoms can be done in several ways. For example $D^* \cdot D^*$ and $D^*$ coincide. They also coincide with $D^* \cdot (I + \epsilon)$ and $D^* \cdot D'^*$ if $I$, resp. $D'$, are subsets of $D^*$.

More generally, if $\boldsymbol{A}$ is an atom and $D \in Down(X)$ is such that $\boldsymbol{A} \subseteq D^*$, then $\boldsymbol{A}D^* = D^*\boldsymbol{A} = D^*$. Subsequently, we show that these are the only causes of non-uniqueness: avoiding such redundancies, every ideal has a unique representation as a product of atoms. (This was already observed for finite alphabets in [3].) This can be used to define a canonical representation for ideals of $X^*$ (assuming one has defined a canonical representation for the ideals of $X$) and then for the downwards-closed sets. This representation is easy to use (moving from an arbitrary product of atoms to the canonical representation just requires testing inclusions between atoms) and can lead to more efficient algorithms.

Below, we use letters such as $\mathtt{A}$, $\mathtt{P}$, etc., to denote sequences of atoms (syntax), and corresponding letters such as $\boldsymbol{A}$, $\boldsymbol{P}$, etc. to denote the ideals obtained by taking the product (semantics). For example if $\mathtt{P} = (\mathtt{A}_1, \mathtt{A}_2, \cdots, \mathtt{A}_n)$, then $\boldsymbol{P} = \boldsymbol{A}_1 \cdot \boldsymbol{A}_2 \cdot \cdots \cdot \boldsymbol{A}_n$. Thus it is possible to have $\mathtt{P} \neq \mathtt{Q}$ and $\boldsymbol{P} = \boldsymbol{Q}$.

**Definition 4.21** A sequence of atoms $\boldsymbol{A}_1, \cdots, \boldsymbol{A}_n$ is said to be *reduced* if for every $i$, the following hold:

- $\boldsymbol{A}_i \neq \emptyset^* = \{\epsilon\}$;
- if $i < n$ and $\boldsymbol{A}_{i+1}$ is some $D^*$, then $\boldsymbol{A}_i \nsubseteq \boldsymbol{A}_{i+1}$;
- if $i > 1$ and $\boldsymbol{A}_{i-1}$ is some $D^*$, then $\boldsymbol{A}_i \nsubseteq \boldsymbol{A}_{i-1}$.

Every ideal has a reduced decomposition into atoms, since any decomposition can be converted to a reduced one by dropping atoms which are redundant as per Definition 4.21. It remains to show that reduced representations are unique:

**Theorem 4.22** *If* $\mathtt{P}$ *and* $\mathtt{Q}$ *are reduced sequences of atoms such that* $\boldsymbol{P} = \boldsymbol{Q}$*, then* $\mathtt{P} = \mathtt{Q}$.

*Proof* Let us first observe that the claim is obvious for atoms: $\boldsymbol{A} = \boldsymbol{B}$ entails $\mathtt{A} = \mathtt{B}$. We now prove the statement in several steps: let $\mathtt{A} \cdot \mathtt{P}$ and $\mathtt{B} \cdot \mathtt{Q}$ be two reduced sequences of atoms.

*First claim:* $\boldsymbol{A} \cdot \boldsymbol{P} \neq \boldsymbol{P}$:
By induction on $\mathtt{P}$. If $\mathtt{P}$ is the empty sequence, then $\boldsymbol{P} = \{\epsilon\}$ and $\boldsymbol{A} \cdot \boldsymbol{P} = \boldsymbol{A}$. Now Definition 4.21 guarantees $\boldsymbol{A} \neq \{\epsilon\}$. Otherwise, $\mathtt{P}$ is some $\mathtt{A}' \cdot \mathtt{P}'$. If $\boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{P}$, the inclusion test described in Lemma 4.13 implies either $\boldsymbol{A} \subseteq \boldsymbol{A}'$, which contradicts reducedness, or $\boldsymbol{A} \cdot \boldsymbol{A}' \cdot \boldsymbol{P}' \subseteq \boldsymbol{P}'$, which entails $\boldsymbol{A}' \cdot \boldsymbol{P}' \subseteq \boldsymbol{P}'$ and contradicts the induction hypothesis. Therefore $\boldsymbol{A} \cdot \boldsymbol{P} \neq \boldsymbol{P}$.

*Second claim*: $\boldsymbol{A} \cdot \boldsymbol{P} = \boldsymbol{B} \cdot \boldsymbol{Q}$ implies $\boldsymbol{A} = \boldsymbol{B}$:
Since $\boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{B} \cdot \boldsymbol{Q}$, Lemma 4.13 implies either $\boldsymbol{A} \subseteq \boldsymbol{B}$, or $\boldsymbol{A} \cdot \boldsymbol{P} \subseteq \boldsymbol{Q}$. The second option, combined with $\boldsymbol{Q} \subseteq \boldsymbol{B} \cdot \boldsymbol{Q} \subseteq \boldsymbol{A} \cdot \boldsymbol{P}$, leads to $\boldsymbol{Q} = \boldsymbol{B} \cdot \boldsymbol{Q}$, which is impossible (first claim). Therefore, $\boldsymbol{A} \subseteq \boldsymbol{B}$, and the reverse inclusion is proved symmetrically.
*Third claim*: $\boldsymbol{A} \cdot \boldsymbol{P} = \boldsymbol{B} \cdot \boldsymbol{Q}$ and $\boldsymbol{A} = \boldsymbol{B}$ imply $\boldsymbol{P} = \boldsymbol{Q}$:
If $\mathtt{Q}$ is the empty sequence, then $\boldsymbol{A} \cdot \boldsymbol{P} = \boldsymbol{B} \cdot \boldsymbol{Q} = \boldsymbol{B} = \boldsymbol{A}$, thus $\boldsymbol{P} \subseteq \boldsymbol{A}$. But if $\mathtt{P}$ is some $\mathtt{A}' \cdot \mathtt{P}'$ then by Lemma 4.13 either $\boldsymbol{A}' \subseteq \boldsymbol{A}$, which is impossible by reducedness of $\mathtt{A} \cdot \mathtt{P}$, or $\boldsymbol{A}' \cdot \boldsymbol{P}' \subseteq \{\epsilon\}$, requiring $\boldsymbol{A}' \subseteq \{\epsilon\}$ which is also impossible. Thus $\mathtt{P}$ too is the empty sequence.

If $|\mathtt{P}| = 0$ the same reasoning applies so we now assume that both products are non-trivial, writing $\mathtt{P} = \mathtt{A}' \cdot \mathtt{P}'$ and $\mathtt{Q} = \mathtt{B}' \cdot \mathtt{Q}'$. If now $\mathtt{A}$ is $I + \epsilon$ for some $I$, then so is $\mathtt{B}$ and Lemma 4.13 implies $\boldsymbol{A}' \cdot \boldsymbol{P}' \subseteq \boldsymbol{B}' \cdot \boldsymbol{Q}'$. Otherwise, $\mathtt{A}$ is $D^*$ for some $D$, in which case Lemma 4.13 entails first $\boldsymbol{A}' \cdot \boldsymbol{P}' \subseteq \boldsymbol{B} \cdot \boldsymbol{B}' \cdot \boldsymbol{Q}'$, then $\boldsymbol{A}' \cdot \boldsymbol{P}' \subseteq \boldsymbol{B}' \cdot \boldsymbol{Q}'$ (since $\boldsymbol{A}' \nsubseteq \boldsymbol{A} = \boldsymbol{B}$ by reducedness of $\mathtt{A} \cdot \mathtt{P}$). In other words, we deduce $\boldsymbol{P} \subseteq \boldsymbol{Q}$ and the reverse inclusion is proved symmetrically.

*Proof of the Theorem*. By induction on $|\mathtt{P}| + |\mathtt{Q}|$. If either $\mathtt{P}$ or $\mathtt{Q}$ is the empty sequence, the property is trivially verified, otherwise we can write $\mathtt{P} = \mathtt{A} \cdot \mathtt{P}'$ and $\mathtt{Q} = \mathtt{B} \cdot \mathtt{Q}'$. From $\boldsymbol{P} = \boldsymbol{Q}$ we deduce $\boldsymbol{A} = \boldsymbol{B}$ (second claim), which in turn implies $\boldsymbol{P}' = \boldsymbol{Q}'$ (third claim), hence $\mathtt{P}' = \mathtt{Q}'$ by induction hypothesis. We already noted that $\boldsymbol{A} = \boldsymbol{B}$ implies $\mathtt{A} = \mathtt{B}$ and combining those gives $\mathtt{P} = \mathtt{Q}$. □

## 4.5 Finitary Powersets

Given a QO $(X, \leq)$, we write $\mathcal{P}(X)$ to denote its powerset, with typical elements $S, T, \ldots$ A usual way of extending the quasi-ordering between elements of $X$ into a quasi-ordering between sets of such elements is the *Hoare quasi-ordering* (also called *domination quasi-ordering*), denoted $\sqsubseteq_H$, and defined by

$$S \sqsubseteq_H T \overset{\text{def}}{\Leftrightarrow} \forall x \in S : \exists y \in T : x \leq y.$$

A convenient characterization of this ordering is the following: $S \sqsubseteq_H T$ iff $S \subseteq \downarrow_X T$. Note that $(\mathcal{P}(X), \sqsubseteq_H)$ is in general not antisymmetric even when $(X, \leq)$ is. For example, and writing $\equiv_H$ to denote $\sqsubseteq_H \cap \sqsupseteq_H$, the above characterization implies that $S \equiv_H \downarrow_X S$ for any $S \subseteq X$. In particular, this shows that the quotient $\mathcal{P}(X)/{\equiv_H}$ is isomorphic to $(Down(X), \subseteq)$. While $(Down(X), \subseteq)$ is well-founded if, and only if, $(X, \leq)$ is a WQO (cf. Lemma 2.3), this does not guarantee that $(Down(X), \subseteq)$ is a WQO, as famously shown by Rado [50].[5] In other words, powerset is not a WQO-preserving construction.

However, the *finitary powerset* construction is WQO-preserving. Let $\mathcal{P}_f(X)$, sometimes also written $[X]^{<\omega}$, denote the set of all *finite subsets* of $X$.

**Theorem 4.23** *$(\mathcal{P}_f(X), \sqsubseteq_H)$ is WQO if, and only if, $(X, \leq)$ is WQO.*

The if direction is an easy consequence of Higman's Lemma: the function that maps each word in $X^*$ to its set of letters, in $\mathcal{P}_f(X)$, is monotonic and surjective, and the image of a WQO by any monotonic map is WQO. We shall see another proof in Sect. 5.2.

**Proposition 4.24** (Ideals of $\mathcal{P}_f(X)$) *Given a WQO $(X, \leq)$, the ideals of $(\mathcal{P}_f(X), \sqsubseteq_H)$ are exactly the sets $\mathcal{J}$ of the form $\mathcal{J} = \mathcal{P}_f(D)$ for $D \in Down(X)$.*

*Proof* ($\Longleftarrow$): $\emptyset \in \mathcal{P}_f(D)$, so $\mathcal{P}_f(D)$ is non-empty. It is downwards-closed, since if $S \sqsubseteq_H T \in \mathcal{P}_f(D)$, then $S \subseteq \downarrow_X T \subseteq \downarrow_X D = D$. It is directed, since if $S, T \in \mathcal{P}_f(D)$, then $S \cup T \in \mathcal{P}_f(D)$, and $S, T \sqsubseteq_H S \cup T$.
($\Longrightarrow$): Let $\mathcal{J}$ be an ideal of $\mathcal{P}_f(X)$ and let $D = \bigcup_{S \in \mathcal{J}} S$, so that $\mathcal{J} \subseteq \mathcal{P}_f(D)$. Since $\mathcal{J}$ is downwards-closed under $\sqsubseteq_H$, $D$ is downwards-closed under $\leq$ and $\{x\} \in \mathcal{J}$ for all $x \in D$. Since $\mathcal{J}$, being an ideal, is non-empty, $\emptyset \in \mathcal{J}$. Finally, if $S, T \in \mathcal{J}$, then there is some $U \in \mathcal{J}$ such that $S, T \sqsubseteq_H U$. Thus $S \cup T \sqsubseteq_H U$, and therefore $S \cup T \in \mathcal{J}$. Therefore, $\mathcal{J}$ contains the empty set, all the singletons included in $D$, is closed under finite unions, and so is equal to $\mathcal{P}_f(D)$. □

When $(X, \leq)$ is ideally effective, finite subsets of $X$ can be represented using any of the usual data structures and Proposition 4.24 directly leads to a data structure for $Idl(\mathcal{P}_f(X))$ inherited from the representation of $X$'s ideals and downwards-closed sets.

**Theorem 4.25** *With the above representations, the finitary powerset with Hoare's ordering is an ideally effective constructor.*

*Proof* Let $(X, \leq)$ be an ideally effective WQO. In the following we use shorthand notations such as $\downarrow_H$ for $\downarrow_{\mathcal{P}_f(X)}$, etc., with the obvious meaning.

**(OD)**: The sets we consider being finite, the definition of $\sqsubseteq_H$ leads to an obvious implementation.

**(ID)**: Testing inclusion in $Idl(\mathcal{P}_f(X))$ reduces to testing inclusion in $Down(X)$ by $\mathcal{J}_1 = \mathcal{P}_f(D_1) \subseteq \mathcal{J}_2 = \mathcal{P}_f(D_2) \iff D_1 \subseteq D_2$.

[5] In fact $(\mathcal{P}(X), \sqsubseteq_H)$ is a WQO iff $X$ is an $\omega^2$-WQO [34, 48].

**(PI)**: Given $S$ a finite subset of $X$, the principal ideal $\downarrow_H S$ is $\mathcal{P}_f(\downarrow_X S)$ so we just need to compute the downwards-closed $\downarrow_X S = \bigcup_{x\in S} \downarrow x$ in $X$'s representation.

**(CF)**: Given $S \in \mathcal{P}_f(X)$, the complement of $\uparrow_H S$ can be given an ideal decomposition via

$$\mathcal{P}_f(X) \setminus \uparrow_H S = \bigcup_{x\in S} \mathcal{P}_f(X) \setminus \uparrow_H\{x\} = \bigcup_{x\in S} \mathcal{P}_f(X \setminus \uparrow x) .$$

This can now be computed using (CF) for $X$.

**(II)**: We have $\mathcal{P}_f(D_1) \cap \mathcal{P}_f(D_2) = \mathcal{P}_f(D_1 \cap D_2)$.

**(IF)**: Filters may be intersected using $\uparrow S \cap \uparrow T = \uparrow(S \cup T)$.

**(CI):** Given an ideal $\mathcal{J} = \mathcal{P}_f(D)$, $\mathcal{P}_f(X) \setminus \mathcal{J}$ consists of the sets that contain at least one element not in $D$. That is:

$$\neg\mathcal{J} = \uparrow_H\{x_1\} \cup \cdots \cup \uparrow_H\{x_n\} \text{ if } X \setminus D = \uparrow_X x_1 \cup \cdots \cup \uparrow_X x_n,$$

which is computable using (CI) for $X$.

The above proves that the finite powerset constructor is an ideally effective constructor. Once again, the computability of the presentation described above from a presentation of $(X, \leq)$ is clear. For (XI), observe that $\mathcal{P}_f(X)$ is its own ideal decomposition since $X \in Down(X)$. For (XF), use $\mathcal{P}_f(X) = \uparrow_H \emptyset$. □

## 5 More Constructions on Ideally Effective WQOs

In this section we describe more constructions that yield new ideally effective WQOs from previously defined ones. By contrast with the constructors of Sect. 4.1, these constructions take some extra parameters that are not WQOs—for example, an equivalence relation in order to build quotient WQOs (see Sect. 5.2). Showing that the quotient WQO is ideally effective will need some effectiveness assumptions on the equivalence at hand, in the spirit of what we did with the one-sorted constructors.

### *5.1 Order Extension*

Let $(X, \leq)$ be a WQO and let $\leq'$ be an extension of $\leq$ (i.e., $\leq \subseteq \leq'$). Then $(X, \leq')$ is also a WQO. In this subsection, we investigate the ideals of $(X, \leq')$ and present some sufficient condition for $(X, \leq')$ to be ideally effective, assuming $(X, \leq)$ is. In the next subsections, we will present natural applications of this framework.

**Proposition 5.1** *Given a WQO $(X, \leq)$ and an extension $\leq'$ of $\leq$, the ideals of $(X, \leq')$ are exactly the downward closures under $\leq'$ of the ideals of $(X, \leq)$. That is,*

$$Idl(X, \leq') = \{\downarrow_{\leq'} I \mid I \in Idl(X, \leq)\} .$$

*Proof* ($\supseteq$) Let $I \in Idl(X, \leq)$. Even though $I$ may not be downwards-closed in $(X, \leq')$, it is still directed. It is easy to see that $\downarrow_{\leq'} I$ is directed, non-empty, and downwards-closed for $\leq'$. Thus it is an ideal of $(X, \leq')$.
($\subseteq$) Let $J$ be an ideal of $(X, \leq')$. $J$ may not be directed in $(X, \leq)$, but it is still downwards-closed under $\leq$. As a consequence, it can be decomposed as a finite union of ideals of $(X, \leq)$: $J = I_1 \cup \cdots \cup I_n$. Then $J = \downarrow_{\leq'} J = \downarrow_{\leq'} I_1 \cup \cdots \cup \downarrow_{\leq'} I_n$. Now applying Lemma 2.5 to $(X, \leq')$, we have $J = \downarrow_{\leq'} I_i$ for some $i$. □

Assume that $(X, \leq)$ is an ideally effective WQO for which we have a presentation at hand, in particular data structures for $X$ and $Idl(X)$. Let $\leq'$ be an extension of $\leq$. To represent elements of $(X, \leq')$, it is natural to use the same data structure for $X$ as the one used for $(X, \leq)$. For ideals, Proposition 5.1 suggests to also use the same data structure as the one for ideals of $X$. That is, an ideal $J \in Idl(X, \leq')$ will actually be represented by any $I \in Idl(X)$ such that $J = \downarrow_{\leq'} I$.

Using these representations for $(X, \leq')$ does not always lend itself to algorithms that would witness ideal effectiveness, even under the assumptions that $(X, \leq)$ is ideally effective and that $\leq'$ is decidable. There is even a "natural" counter example: the lexicographic ordering over $X \times X$ (see Sect. 6.2). This fact justifies that we make further assumptions. More precisely, we show that $(X, \leq')$ is ideally effective if we can compute downward closures under $\leq'$:

**Theorem 5.2** *Let $(X, \leq)$ be an ideally effective WQO and $\leq'$ an extension of $\leq$. Then, $(X, \leq')$ is ideally effective for the aforementioned data structures of $X$ and $Idl(X, \leq')$, whenever the following functions are computable:*

$$\mathcal{C}l_{\mathrm{I}} : \begin{array}{c} Idl(X, \leq) \to Down(X, \leq) \\ I \mapsto \downarrow_{\leq'} I \end{array} \qquad \mathcal{C}l_{\mathrm{F}} : \begin{array}{c} Fil(X, \leq) \to Up(X, \leq) \\ \uparrow x \mapsto \uparrow_{\leq'}(\uparrow x) = \uparrow_{\leq'} x \end{array}$$

*Moreover, under these assumptions, a presentation of $(X, \leq')$ can be computed uniformly from a presentation of $(X, \leq)$ and algorithms realizing $\mathcal{C}l_{\mathrm{I}}$ and $\mathcal{C}l_{\mathrm{F}}$.*

Note that if $I \in Idl(X, \leq)$, then $\downarrow_{\leq'} I$ is also downwards-closed for $\leq$ and thus can be represented as a downwards-closed subset of $(X, \leq)$. This is precisely this representation that the function $\mathcal{C}l_{\mathrm{I}}$ outputs. Same goes for $\mathcal{C}l_{\mathrm{F}}$. Note that using functions $\mathcal{C}l_{\mathrm{I}}$ and $\mathcal{C}l_{\mathrm{F}}$, it is possible to compute the downward and upward closure under $\leq'$ of arbitrary downwards- and upwards-closed sets for $\leq$ using the canonical decompositions: $\downarrow_{\leq'}(I_1 \cup \cdots \cup I_n) = (\downarrow_{\leq'} I_1) \cup \cdots \cup (\downarrow_{\leq'} I_n)$ and $\uparrow_{\leq'}(\uparrow x_1 \cup \cdots \cup \uparrow x_n) = \uparrow_{\leq'} x \cup \cdots \cup \uparrow_{\leq'} x_n$.

*Proof* We proceed to show that $(X, \leq')$ is ideally effective.

**(OD)**: One can test $x \leq' y$, since this is equivalent to $y \in \mathcal{C}l_{\mathrm{F}}(\uparrow_{\leq} x)$.

**(ID)**: Ideal inclusion can be decided using $\mathcal{C}l_{\mathrm{I}}$ and the inclusion test for downwards-closed sets of $(X, \leq)$: $\downarrow_{\leq'} I_1 \subseteq \downarrow_{\leq'} I_2 \iff I_1 \subseteq \mathcal{C}l_{\mathrm{I}}(I_2)$.

**(PI)**: The principal ideal $\downarrow_{\leq'} x$ of $(X, \leq')$ is represented by $\downarrow_{\leq} x$, since $\downarrow_{\leq'}(\downarrow_{\leq} x) = \downarrow_{\leq'} x$.

**(CF)**: For $x \in X$, the filter complement $X \setminus \uparrow_{\leq'} x$ is $X \setminus \mathcal{C}l_{\mathrm{F}}(\uparrow_{\leq} x)$ which can be computed, using (CF) and (II) for $(X, \leq)$, as a downwards-closed set in $(X, \leq)$. This is represented by an ideal decomposition $D = \bigcup_{i<n} I_i$ which is canonical in $(X, \leq)$ but not necessarily in $(X, \leq')$ since one may have $\downarrow_{\leq'} I_i \subseteq \downarrow_{\leq'} I_j$ for $i \neq j$. However, extracting the canonical ideal decomposition wrt. $\leq'$ can be done using (ID) for $(X, \leq')$.

**(II)**: Intersection of ideals is computed with $\downarrow_{\leq'} I_1 \cap \downarrow_{\leq'} I_2 = \mathcal{C}l_{\mathrm{I}}(I_1) \cap \mathcal{C}l_{\mathrm{I}}(I_2)$. Here again, this results in an ideal decomposition that is canonical for $\leq$ but not for $\leq'$ until we process it as done for (CF).

**(CI), (IF)**: these operations are obtained similarly.

With algorithms for the closure functions $\mathcal{C}l_{\mathrm{I}}$ and $\mathcal{C}l_{\mathrm{F}}$, the presentation above is computable from a presentation of $(X, \leq)$. Regarding (XF) and (XI), we note that filter and ideal decompositions of $X$ for $\leq$ are also valid decompositions for $\leq'$. However, these decompositions might not be canonical for $\leq'$ even if they are for $\leq$, in which case the canonical decompositions can be obtained using (OD) and (ID), as usual. □

### 5.1.1 Sequences Under Stuttering

In this subsection, we apply Theorem 5.2 to an extension of Higman's ordering $\leq_*$ on finite sequences (from Sect. 4.4).

Given a QO $(X, \leq)$, we define the *stuttering ordering* $\leq_{\mathrm{st}}$ over $X^*$ by $\boldsymbol{x} = x_1 \cdots x_n \leq_{\mathrm{st}} \boldsymbol{y} = y_1 \cdots y_m \overset{\mathrm{def}}{\Leftrightarrow}$ there are $n$ indices $1 \leq p_1 \leq p_2 \leq \cdots \leq p_n \leq m$ such that $x_i \leq y_{p_i}$ for all $i = 1, \ldots, n$. Compared with Higman's ordering, the sequence of positions $(p_i)_{i=1,\ldots,n}$ in $\boldsymbol{y}$ need not be *strictly* increasing: repetitions are allowed. For instance, if $X = \{a, b\}$ is a finite alphabet, then $aabbaa \leq_{\mathrm{st}} aba \leq_{\mathrm{st}} aabbaa$ but $aabbaa \not\leq_{\mathrm{st}} ab$. Or with $X = \mathbb{N}$, $(1, 1, 1) \leq_{\mathrm{st}} (2)$. Note that even when $(X, \leq)$ is antisymmetric, $(X^*, \leq_{\mathrm{st}})$ need not be.

*Remark 5.3* There is another way to define the stuttering ordering: define the stuttering equivalence relation $\sim_{\mathrm{st}}$ on $X^*$ as the smallest equivalence relation such that for all $\boldsymbol{x}, \boldsymbol{y} \in X^*$ and $a \in X$, $\boldsymbol{x}a\boldsymbol{y} \sim_{\mathrm{st}} \boldsymbol{x}aa\boldsymbol{y}$. Informally, this equivalence does not distinguish between a single and several consecutive occurrences of a same element. Then, $\leq_{\mathrm{st}} = \leq_* \circ \sim_{\mathrm{st}}$, where $\circ$ denotes the composition of relations. Observe that $\sim_{\mathrm{st}}$ is *not* the same as the equivalence relation $\equiv_{\mathrm{st}} = \leq_{\mathrm{st}} \cap \geq_{\mathrm{st}}$ induced by the ordering, even if $(X, \leq)$ is a partial order $\leq$. For instance, if $a \leq b$ in $X$, then $ab \equiv_{\mathrm{st}} b$ in $X^*$, but $ab \sim_{\mathrm{st}} b$ does not hold. However the inclusion $\sim_{\mathrm{st}} \subseteq \equiv_{\mathrm{st}}$ is always valid. □

Obviously, $\leq_{\mathrm{st}}$ is an order extension of $\leq_*$, thus $(X^*, \leq_{\mathrm{st}})$ is a WQO when $(X, \leq)$ is, and we can apply Theorem 5.2.

**Theorem 5.4** *The stuttering extension of a WQO* $(X, \leq)$ *is an ideally effective constructor.*

*Proof* In the light of Theorems 4.12 and 5.2, it suffices to show that the following closure functions are computable:

$$\mathcal{C}l_{\mathrm{I}} : \begin{array}{c} Idl(X^*, \leq_*) \to Down(X^*, \leq_*) \\ I \mapsto \downarrow_{\mathrm{st}} I \end{array} \qquad \mathcal{C}l_{\mathrm{F}} : \begin{array}{c} Fil(X^*, \leq_*) \to Up(X^*, \leq_*) \\ \uparrow \boldsymbol{u} \mapsto \uparrow_{\mathrm{st}}(\uparrow \boldsymbol{u}) = \uparrow_{\mathrm{st}} \boldsymbol{u} \end{array}$$

Recall from Sect. 5.1 that the ideals of $(X^*, \leq_*)$ are the (concatenation) products of atoms, where atoms are either of the form $D^*$ for some $D \in Down(X)$ or $I + \epsilon$ for some $I \in Idl(X)$. It is quite immediate to see that $\mathcal{C}l_{\mathrm{I}}(D^*) = D^*$ and $\mathcal{C}l_{\mathrm{I}}(I + \epsilon) = I^*$, and that given two products of atoms $\boldsymbol{P}_1, \boldsymbol{P}_2, \mathcal{C}l_{\mathrm{I}}(\boldsymbol{P}_1 \cdot \boldsymbol{P}_2) = \mathcal{C}l_{\mathrm{I}}(\boldsymbol{P}_1) \cdot \mathcal{C}l_{\mathrm{I}}(\boldsymbol{P}_2)$. From these equations, it is simple to write an inductive algorithm computing $\mathcal{C}l_{\mathrm{I}}$.

Function $\mathcal{C}l_{\mathrm{F}}$ is computable as well, although less straightforward. We provide an expression for $\mathcal{C}l_{\mathrm{F}}$ in Lemma 5.5 which is clearly computable. □

**Lemma 5.5** *Given* $\boldsymbol{u} = x_1 \cdots x_n \in X^*$ *a non-empty sequence,*

$$\mathcal{C}l_{\mathrm{F}}(\boldsymbol{u}) = \uparrow_{\mathrm{st}} \boldsymbol{u} = \uparrow_* \left\{ y_1 \cdots y_k \;\middle|\; \begin{array}{l} 0 < k \leq n \\ 0 = \ell_0 < \ell_1 < \cdots < \ell_k = n \\ \forall j = 1, \ldots, k : y_j \in \min(\bigcap_{\ell_{j-1} < \ell \leq \ell_j} \uparrow_X x_\ell) \end{array} \right\},$$

*where* $\min(A)$ *denotes a finite basis of the upwards-closed subset* $A$*. The remaining case is trivial:* $\mathcal{C}l_{\mathrm{F}}(\epsilon) = \uparrow_* \epsilon$.

*(Intuitively, the set ranges over all ways to cut* $\boldsymbol{u}$ *in* $k$ *consecutive pieces, and embeds all elements of the* $j$*-th piece into the same element* $y_j$*. It has long sequences, the longest being* $\boldsymbol{u}$*, and shorter ones with potentially larger elements.)*

This is the fully generic formula to describe the function $\mathcal{C}l_{\mathrm{F}}$ for any $X$. However, in simple cases, $\mathcal{C}l_{\mathrm{F}}(\boldsymbol{w})$ takes a much simpler form. For instance, for $X = \mathbb{N}$, we have $\mathcal{C}l_{\mathrm{F}}(x_1 \cdots x_n) = \uparrow_* \max(x_1, \cdots, x_n)$, and for $X = \Sigma$ a finite alphabet, $\mathcal{C}l_{\mathrm{F}}(\boldsymbol{w}) = \uparrow_* \boldsymbol{v}$ where $\boldsymbol{v}$ is the shortest member of the equivalence class $[\boldsymbol{w}]_{\sim_{st}}$ (that is, $\boldsymbol{v}$ is obtained from $\boldsymbol{w}$ by fusing consecutive equal letters).

*Proof (Of Lemma 5.5)* The "⊇" direction is obvious.

(⊆) Given $\boldsymbol{w} \geq_{\mathrm{st}} x_1 \cdots x_n$, there exists a decomposition $\boldsymbol{w} = \boldsymbol{w}_0 y_1 \boldsymbol{w}_1 y_2 \cdots y_k \boldsymbol{w}_k$ for some $k \leq n$, $y_1, \ldots, y_k \in X$ and $\boldsymbol{w}_0, \ldots, \boldsymbol{w}_k \in X^*$, and there exists a monotonic mapping $p : [n] \to [k]$ such that $x_i \leq y_{p(i)}$. For $j \in [k]$, define $i_j$ to be the largest $i$ such that $p(i) = j$ (i.e., the index of the right-most symbol of $x_1 \cdots x_n$ to be mapped to $y_j$), and let $i_0 = 0$. It follows that $0 = i_0 < i_1 < \cdots < i_k = n$, and for all $\ell \in [n]$ and $j \in [k]$, $i_{j-1} < \ell \leq i_j \implies x_\ell \leq y_j$. Then $\boldsymbol{w} \geq_* y_1 \cdots y_k$ which is indeed an element of the set described in the proposition. □

### *5.2 Quotienting Under a Compatible Equivalence*

In this subsection, we apply the results of Sect. 5.1 to the most commonly encountered case of order-extension: quotient under an equivalence relation.

Let $(X, \leq)$ be a WQO and let $E$ be an equivalence relation on $X$ which is *compatible* with $\leq$ in the sense that $\leq \circ E = E \circ \leq$, where $\circ$ denotes the composition of relations. Define the relation $\leq_E$ on $X$ to be $\leq \circ E$. Then $\leq_E$ is clearly reflexive, and is transitive since

$$\leq_E \circ \leq_E = \ (\leq \circ E) \circ (\leq \circ E) \ = \leq \circ (E \circ \leq) \circ E = \leq \circ (\leq \circ E) \circ E = \leq \circ E = \leq_E .$$

In this subsection, we give sufficient conditions for $(X, \leq_E)$ to be ideally effective, provided $(X, \leq)$ is.

*Remark 5.6* Note that stuttering from Sect. 5.1.1 is *not* an example: Although $\leq_{\mathrm{st}} = \leq_* \circ \sim_{\mathrm{st}}$, the other condition does not hold: $\leq_{\mathrm{st}} \neq \sim_{\mathrm{st}} \circ \leq_*$. For instance, consider $X = \mathbb{N}^2$ where $\langle 1, 2\rangle\langle 2, 1\rangle \leq_{\mathrm{st}} \langle 2, 2\rangle$. However, if $X$ is a finite alphabet, the equality $\leq_{\mathrm{st}} = \sim_{\mathrm{st}} \circ \leq_*$ holds and $(X^*, \leq_{\mathrm{st}})$ can be treated as a quotient. As another example, the finitary powerset $\mathcal{P}_f(X)$ from Sect. 4.5 can be obtained as a quotient of $(X^*, \leq_*)$, and could be shown ideally effective using Theorem 5.7 below. However, because operations in $\mathcal{P}_f(X)$ are quite simple, and because powerset is a fundamental constructor, we decided to provide a direct, more concrete, construction. □

Observe that $\leq_E$ is an extension of $\leq$, and thus results on quotients can be seen as an application of Sect. 5.1. However, since quotients are of such importance in computer science (and used more often than mere extensions), we reformulate Theorem 5.2 in this specific context: functions $\mathcal{C}l_{\mathrm{I}}$ and $\mathcal{C}l_{\mathrm{F}}$ take an interesting form. As in the case of extensions, elements and ideals of $(X, \leq_E)$ will be represented using the data structures coming from a presentation of $(X, \leq)$.

**Theorem 5.7** *Let $(X, \leq)$ be an ideally effective WQO and $E$ be an equivalence relation on $X$ compatible with $\leq$. Then, $(X, \leq_E)$ is ideally effective for the aforementioned data structures of $X$ and $Idl(X, \leq_E)$, whenever the following functions are computable:*

$$\mathcal{C}l_{\mathrm{I}} : \begin{array}{c} Idl(X, \leq) \to Down(X, \leq) \\ I \mapsto \overline{I} \end{array} \qquad \mathcal{C}l_{\mathrm{F}} : \begin{array}{c} Fil(X, \leq) \to Up(X, \leq) \\ \uparrow x \mapsto \uparrow \overline{x} \end{array}$$

*where, given $S \subseteq X$, $\overline{S}$ denotes the closure under $E$ of $S$, i.e., $\overline{S} \overset{def}{=} \{y \mid \exists x \in S : x \; E \; y\}$, and $\overline{x}$ is a shortcut for $\overline{\{x\}}$ which is the equivalence class of $x$.*

*Moreover, under these assumptions, we can compute a presentation of $(X, \leq_E)$ from a presentation of $(X, \leq)$.*

*Proof* In the light of Theorem 5.2, it suffices to show $\uparrow_{\leq_E} F = \overline{F}$ and $\downarrow_{\leq_E} I = \overline{I}$ for any filter $F$ and any ideal $I$ of $(X, \leq)$. The first equality follows from $\leq_E = \leq \circ E$ while the second comes from $\leq_E = E \circ \leq$. This is why we introduced the compatibility condition $\leq \circ E = E \circ \leq$. □

In particular, we see that the ideals of $(X, \leq_E)$ are exactly the closures under $E$ of the ideals of $(X, \leq)$. That is, $Idl(X, \leq_E) = \{\overline{I} : I \in Idl(X, \leq)\}$.

We conclude this section with two results that are specific to WQOs obtained by quotienting, and which lead to simplifications in several algorithms.

**Proposition 5.8** *Let $J$ be an ideal under $\leq_E$, and let $J = I_1 \cup \cdots \cup I_k$ be the canonical ideal decomposition of $J$ under $\leq$. Then $J = \overline{I_i}$ for every $i$.*

*Proof* Recall from the proof of Proposition 5.1 that $J = \overline{I_i}$ for some $i$. Without loss of generality, we can assume $i = 1$. For the sake of contradiction, suppose that there exists some $i$ such that $J \neq \overline{I_i}$. Again without loss of generality, we can assume $i = k$. From $\overline{I_1} \neq \overline{I_k}$, we deduce that there exists $x \in I_1$ which has no $E$-equivalent in $I_k$.

We will now show that $J \subseteq I_1 \cup \cdots \cup I_{k-1}$, which will be a contradiction since we assumed that we started from a canonical ideal decomposition. Let $y \in J$. Then there exists a $y' \in I_1$ such that $y \; E \; y'$. Since $I_1$ is an ideal under $\leq$, there is a $z \in I_1$ such that $x \leq z$ and $y' \leq z$. We have $y \; E \; y' \leq z$, thus there exists $z'$ such that $y \leq z' \; E \; z$. Since $J$ is closed under $E$-equivalence, $z' \in J$, hence $z' \in I_i$ for some $i$. However, $z'$ cannot belong to $I_k$, since $x \leq_E z'$ and the $E$-equivalence class of $x$ is disjoint from $I_k$. So $z' \in I_1 \cup \cdots \cup I_{k-1}$, and hence $y \in I_1 \cup \cdots \cup I_{k-1}$. Thus $J = I_1 \cup \cdots \cup I_{k-1}$, and we have a contradiction. □

**Proposition 5.9** *For any two ideals $I_1, I_2 \in Idl(X, \leq)$, $\overline{I_1} \cap \overline{I_2} = \overline{I_1 \cap \overline{I_2}} = \overline{\overline{I_1} \cap I_2}$. For any two filters $F_1, F_2 \in Fil(X, \leq)$, $\overline{F_1} \cap \overline{F_2} = \overline{F_1 \cap \overline{F_2}} = \overline{\overline{F_1} \cap F_2}$.*

*Proof* We show $\overline{I_1} \cap \overline{I_2} = \overline{I_1 \cap \overline{I_2}}$, the other equality is symmetric. For the right-to-left inclusion, we have $I_1 \cap \overline{I_2} \subseteq \overline{I_1} \cap \overline{I_2}$, and closing both sides under $E$ gives the required result. For the left-to-right inclusion, let $x \in \overline{I_1} \cap \overline{I_2}$. Then there exist $x_1 \in I_1$ and $x_2 \in I_2$ such that $x_1 \; E \; x \; E \; x_2$. Then $x_1 \in I_1 \cap \overline{I_2}$, and thus $x \in \overline{I_1 \cap \overline{I_2}}$.

The same proof applies to filters. □

Thanks to Proposition 5.9, we can compute intersections of filters (resp., ideals) with only one invocation of $\mathcal{C}l_{\mathrm{F}}$ (resp., $\mathcal{C}l_{\mathrm{I}}$) instead of the two invocations required by the algorithm described in the proof of Theorem 5.2.

### 5.2.1 Sequences Under Conjugacy

Consider a WQO $(X, \leq)$, and define an equivalence relation $\sim_{\mathrm{cj}}$ on $X^*$ as follows: $\boldsymbol{w} \sim_{\mathrm{cj}} \boldsymbol{v}$ iff there exist $\boldsymbol{x}, \boldsymbol{y}$ such that $\boldsymbol{w} = \boldsymbol{x}\boldsymbol{y}$ and $\boldsymbol{v} = \boldsymbol{y}\boldsymbol{x}$. One can imagine an equivalence class of $\sim_{\mathrm{cj}}$ as a sequence written on an (oriented) circle instead of a line. We can now define a notion of *subwords under conjugacy* via $\leq_{\mathrm{cj}} \stackrel{\mathrm{def}}{=} \sim_{\mathrm{cj}} \circ \leq_*$, which is exactly the relation denoted $\preceq_c$ in [3, p. 49].

Since $\sim_{\mathrm{cj}}$ is compatible with $\leq_*$, that is $\leq_* \circ \sim_{\mathrm{cj}} \; = \; \sim_{\mathrm{cj}} \circ \leq_*$, our results over quotients apply to $(X^*, \leq_{\mathrm{cj}})$.

**Theorem 5.10** *Sequence extension with conjugacy is an ideally effective constructor.*

*Proof* Note that the data structures used for elements and ideals of $(X^*, \leq_{\mathrm{cj}})$ are obtained from data structures for $(X^*, \leq_*)$ as done with Theorem 5.7.

In the light of Theorem 5.7, it suffices to show that we can compute closures under $\sim_{\mathrm{cj}}$ of elements and ideals of $(X^*, \leq_*)$. Given $\boldsymbol{w} \in X^*$, the equivalence class of $\boldsymbol{w}$ under $\sim_{\mathrm{cj}}$ is equal to $\overline{\boldsymbol{w}} = \{c^{(i)}(\boldsymbol{w}) \mid 0 \leq i < \max(1, |\boldsymbol{w}|)\}$, where $c^{(i)}$ denotes the $i$-th iterate of the cycle operator $c(w_1 \cdots w_n) = w_2 \cdots w_n w_1$, which corresponds to rotating the sequence $i$ times. This expression is obviously computable.

Computing the closure under $\sim_{\mathrm{cj}}$ of ideals is quite similar. Remember that ideals of $(X^*, \leq_*)$ are products of atoms, where atoms are either of the form $D^*$ for some $D \in Down(X)$, or of the form $I + \epsilon$, for some $I \in Idl(X)$. Then, given $\boldsymbol{P} = \boldsymbol{A}_0 \cdots \boldsymbol{A}_{k-1}$ an ideal of $(X^*, \leq_*)$:

$$\overline{\boldsymbol{P}} = \bigcup_{i=0}^{k-1} c^{(i)}(\boldsymbol{P}) \cdot e(\boldsymbol{A}_i) ,$$

where $e(D^*) = D^*$ and $e(I + \epsilon) = \epsilon$. The presence of the extra $e(A_i)$ in the above expression might become clearer when considering a simple example as $\boldsymbol{P} = \{a\}^*\{b\}^*$ where $\overline{\boldsymbol{P}} = \{a\}^*\{b\}^*\{a\}^* \cup \{b\}^*\{a\}^*\{b\}^*$. Indeed, $abba \sim_{\mathrm{cj}} baab \sim_{\mathrm{cj}} aabb \in \boldsymbol{P}$. □

### 5.2.2 Multisets Under the Embedding Ordering

Given a WQO $(X, \leq)$, we consider the set $X^{\circledast}$ of finite multisets over $X$. Intuitively, multisets are sets where an element might occur multiple times. Formally, a multiset $M \in X^{\circledast}$ is a function from $X$ to $\mathbb{N}$: $M(x)$ denotes the number of occurrences of $x$ in $M$. The support of a multiset $M$ denoted $Supp(M)$ is the set $\{x \in X \mid M(x) \neq 0\}$. A multiset is said to be finite if its support is.

A natural algorithmic representation for these objects are lists of elements of $X$, keeping in mind that a permutation of a list represents the same multiset. Formally, this means that $X^{\circledast}$ is the quotient of $X^*$ by the equivalence relation $\sim$ defined by

$$\boldsymbol{u} = u_1 \cdots u_n \sim \boldsymbol{v} = v_1 \cdots v_m \overset{\text{def}}{\Leftrightarrow} n = m \wedge \exists \sigma \in S_n : u_i = v_{\sigma(i)} \text{ for all } i = 1, \ldots, n,$$

where $S_n$ denotes the group of permutations over $\{1, \cdots, n\}$.

Once again, the equivalence relation $\sim$ is compatible with $\leq_*$. We denote by $\leq_{\mathrm{emb}}$ the composition $\sim \circ \leq_* = \leq_* \circ \sim$, often called multiset embedding. (There exist other classical quasi-orderings on finite multisets, such as the domination quasi-ordering, aka the Dershowitz-Manna quasi-ordering [14]: see [32, Theorem 7.2.3] for a proof that it is an ideally effective constructor.) For this section, we focus on $(X^{\circledast}, \leq_{\mathrm{emb}})$, which is an application of our results on quotients.

**Theorem 5.11** *Finite multisets with multiset embedding is an ideally effective constructor.*

*Proof (Sketch)* Note that the data structures used for elements and ideals of $(X^{\circledast}, \leq_{\text{emb}})$ are obtained from data structures for $(X^*, \leq_*)$ as done with Theorem 5.7.

In the light of Theorem 5.7, it suffices to show that we can compute closures under $\sim$ of elements and ideals of $(X^*, \leq_*)$. Given $\boldsymbol{w} = x_1 \cdots x_n \in X^*$, the equivalence class of $\boldsymbol{w}$ under $\sim$ simply consists of all the possible permutations of the word $\boldsymbol{w}$:

$$\overline{\boldsymbol{w}} = \bigcup_{\sigma \in S_n} x_{\sigma(1)} \cdots x_{\sigma(n)} \ .$$

Closures of ideals are a little more complex. Let $\boldsymbol{P} = \boldsymbol{A}_1 \cdots \boldsymbol{A}_n$ be an ideal of $(X^*, \leq_*)$. Define:

$$D \stackrel{\text{def}}{=} \bigcup \big\{ E \in Down(X) \bigm| \exists i \in \{1, \cdots, n\} : E^* = \boldsymbol{A}_i \big\}.$$

In other words, $D \in Down(X)$ is obtained from $\boldsymbol{P}$ by picking the atoms $\boldsymbol{A}_i$ that are of the second kind, $\boldsymbol{A}_i = E^*$, and taking the union of their generators. Similarly, let $I_1, \ldots, I_p$ be the ideals of $(X, \leq)$ that appear as $I_i + \epsilon$ in $\boldsymbol{P}$, with repetitions, and in order of occurrence. Then:

$$\overline{\boldsymbol{P}} = \bigcup_{\sigma \in S_p} D^* I_{\sigma(1)} D^* \cdots D^* I_{\sigma(p)} D^* \ .$$

□

## 5.3 Induced WQOs

Let $(X, \leq)$ be a WQO. A subset $Y$ of $X$ (not necessarily finite) induces a quasi-ordering $(Y, \leq)$ which is also WQO.

Any subset $S \subseteq X$ induces a subset $Y \cap S$ in $Y$. Obviously, if $S$ is upwards-closed (or downwards-closed) in $X$, then it induces an upwards-closed (resp., downwards-closed) subset in $Y$. However an ideal $I$ or a filter $F$ in $X$ does not always induce an ideal or a filter in $Y$. In the other direction though, if $J \in Idl(Y)$, the downward closure $\downarrow_X J$ is an ideal of $X$. Therefore, to describe the ideals of $Y$, we need to identify those ideals of $X$ that are of the form $\downarrow_X J$ for some ideal $J$ of $Y$. This is captured by the following notion:

**Definition 5.12** Given a WQO $(X, \leq)$ and a subset $Y$ of $X$, we say that an ideal $I \in Idl(X)$ is *in the adherence* of $Y$ if $I = \downarrow_X (I \cap Y)$.

In particular this implies that $I \subseteq \downarrow_X Y$ (we say that $I$ is "below $Y$") and $I \cap Y \neq \emptyset$ (we say that $I$ is "crossing $Y$"). The converse implication does not hold, as witnessed by $X = \mathbb{N}$, $Y = [1, 3] \cup [5, 7]$ and $I = \downarrow 4$.

We now show that the ideals of $Y$ are exactly the subsets induced by ideals of $X$ that are in the adherence of $Y$.

**Theorem 5.13** *Let $(X, \leq)$ be a WQO and $Y$ be a subset of $X$. A subset $J$ of $Y$ is an ideal of $Y$ if and only if $J = I \cap Y$ for some $I \in Idl(X)$ in the adherence of $Y$. In this case, $I = \downarrow_X J$, and is thus uniquely determined from $J$.*

*Proof* ($\Longrightarrow$): If $J \in Idl(Y)$ then $I \stackrel{\text{def}}{=} \downarrow_X J$ is directed hence is an ideal of $X$. Clearly, $J = I \cap Y$, so $I$ is in the adherence of $Y$.
($\Longleftarrow$): If $I \in Idl(X)$ is in the adherence of $Y$ then $J \stackrel{\text{def}}{=} I \cap Y$ is non-empty (since $I$ is crossing $Y$) and it is directed since for any $x, y \in J$ there is $z \in I$ above $x$ and $y$, and $z \leq z'$ for some $z' \in J$ since $I$ is below $Y$.

Uniqueness is clear since the compatibility assumption "$I = \downarrow_X (I \cap Y)$" completely determines $I$ from the ideal $J = I \cap Y$ it induces. □

An earlier definition of *adherence* can be found in the literature: an ideal $I \in Idl(X)$ is in the adherence of $Y$ if and only if there exists a directed subset $\Delta \subseteq Y$ such that $I = \downarrow_X \Delta$ [43]. The two definitions are equivalent [27, Lemma 14], so that, notably, Lemma 5.13 extends Lemma 4.6 from [43].

*Proof (that the two notions of adherence coincide)* ($\Longrightarrow$) : Assume $I = \downarrow_X (I \cap Y)$. We show that $\Delta = I \cap Y$ is directed: let $x, y \in \Delta \subseteq I$, since $I$ is directed, there exists $z \in I$ such that $z \geq x, y$. But since $I = \downarrow_X \Delta$, there exists $z' \in \Delta$ such that $z' \geq z \geq x, y$, which proves that $\Delta$ is directed.
($\Longleftarrow$): Assume that there exists a directed subset $\Delta \subseteq Y$ such that $I = \downarrow_X \Delta$. Then $\downarrow_X (I \cap Y) = \downarrow_X (\downarrow_X \Delta \cap Y) = \downarrow_X (\Delta \cap Y) = \downarrow_X \Delta = I$. □

Similarly, we can define a notion of adherence for filters. However, in this case, the condition $F = \uparrow_X (F \cap Y)$ simplifies: writing $F$ as $\uparrow_X x$, this means that $x' \equiv_X x$ for some $x' \in Y$, in which case $F = \uparrow_X x'$. This is not surprising: $(Y, \leq)$ is a WQO, hence all its filters are principal.

Assuming that $(X, \leq)$ is an ideally effective WQO, and given $Y \subseteq X$, we can simply represent elements of $Y$ by restricting the data structure for $X$ to $Y$. This requires that $Y$ be a recursive set. Alternatively, Theorem 5.13 suggests that we represent ideals of $Y$ as ideals of $X$ that are in the adherence of $Y$. This requires that we can decide membership in the adherence of $Y$. As in the case of extensions, the ideal effectiveness of $(Y, \leq)$ does not always follow from the ideal effectiveness of $(X, \leq)$ (see [32, Sect. 8.4] for an example). We therefore have to introduce extra assumptions.

**Theorem 5.14** *Let $(X, \leq)$ be a WQO and $Y \subseteq X$. Then $(Y, \leq)$ is ideally effective (for the aforementioned representations) provided:*

- *membership in $Y$ is decidable over (the representation for) $X$,*
- *the following functions are computable:*

$$\mathcal{S}_{\text{I}} : \begin{array}{r} Idl(X, \leq) \to Down(X, \leq) \\ I \mapsto \downarrow_X (I \cap Y) \end{array} \qquad \mathcal{S}_{\text{F}} : \begin{array}{r} Fil(X, \leq) \to Up(X, \leq) \\ F \mapsto \uparrow_X (F \cap Y) \end{array}$$

*Moreover, in this case, a presentation of* $(Y, \leq)$ *can be computed from a presentation of* $(X, \leq)$.

The rest of this subsection is dedicated to the proof of this theorem.

First, let us mention that our first assumption implies that we have a data structure for elements of $Y$ and that thanks to function $\mathcal{S}_{\mathrm{I}}$, we can decide whether an ideal $I$ of $X$ is in the adherence of $Y$: it suffices to check that $\mathcal{S}_{\mathrm{I}}(I) = I$.

Let us prove that $(Y, \leq)$ is ideally effective.

**(OD)**: since $\leq$ is decidable on $X$, its restriction to $Y$ is still decidable.

**(ID)**: Given two ideals $I_1$, $I_2$ that are in the adherence of $Y$, $I_1 \cap Y \subseteq I_2 \cap Y \iff I_1 \subseteq I_2$. The left-to-right implication uses that $I_i = \downarrow_X(I_i \cap Y)$. Therefore, inclusion for ideals of $Y$ can be implemented by relying on (ID) for $X$.

**(PI)**: if $y \in Y$, then $\downarrow_X y$ is adherent to $Y$ and one relies on $\downarrow_Y y = \downarrow_X y \cap Y$.

For the four remaining operations, we need to be able to compute a representation of $D \cap Y$ and $U \cap Y$ for $D \in Down(X)$ and $U \in Up(X)$.

**Lemma 5.15** *Let* $D \in Down(X)$. *The canonical representation of* $D \cap Y$ *(as a downwards-closed set of* $Y$*) is exactly the canonical representation of* $\downarrow_X(D \cap Y)$ *(as a downwards-closed set of* $X$*).*

*Proof* Let $\bigcup_i I_i$ be the *canonical* decomposition of $\downarrow_X(D \cap Y)$. Remember that an ideal $J$ of $Y$ is represented by the unique ideal $I$ of $X$ which is in the adherence of $Y$ such that $J = I \cap Y$. Thus, stating that $\bigcup_i I_i$ is the canonical representation of $D \cap Y$ means that:

1. $D \cap Y = \bigcup_i (I_i \cap Y)$;
2. for every $i$, $I_i \cap Y$ is an ideal of $Y$;
3. $I_i \cap Y$ and $I_j \cap Y$ are incomparable for inclusion, for $i \neq j$.

For the first point, $\bigcup_i (I_i \cap Y) = (\bigcup_i I_i) \cap Y = (\downarrow_X(D \cap Y)) \cap Y = D \cap Y$.

We now argue that each $I_i \cap Y$ is indeed an ideal of $Y$, i.e., all $I_i$'s are in the adherence of $Y$. One inclusion being trivial, we need to show that $I_i \subseteq \downarrow_X(I_i \cap Y)$, for any $i$. Let $x_i \in I_i$. Since the ideals $I_j$ are incomparable for inclusion, there exists $x_i' \in I_i$ such that $x_i \leq x_i'$ and for any $j \neq i$, $x_i' \notin I_j$ ($I_i$ is directed). Besides, $x_i' \in I_i \subseteq \downarrow_X(D \cap Y)$ and thus there is an element $x_i''$ such that $x_i' \leq x_i'' \in D \cap Y$. As the sets $I_j$ are downwards-closed, $x_i''$ cannot belong to any $I_j$ with $j \neq 0$, hence $x_i''$ is in $I_i \cap Y$. Therefore, $x_i \in \downarrow_X(I_i \cap Y)$.

Finally, the ideal decomposition $D \cap Y = \bigcup_j (I_j \cap Y)$ is canonical since the $I_j$'s are incomparable in $X$ (recall the above criterion for inclusion of ideals of $Y$). □

Observe that if $D = \bigcup_i I_i$ then $\downarrow_X(D \cap Y) = \bigcup_i \downarrow_X(I_i \cap Y) = \bigcup_i \mathcal{S}_{\mathrm{I}}(I)$. Thus the canonical representation of $D \cap Y$ is indeed computable from $D \in Down(X)$.

We now present the dual of the previous lemma:

**Lemma 5.16** *Given* $U \in Up(X)$, *a canonical representation of* $U \cap Y$ *(as an upwards-closed set of* $Y$*) can be computed from a canonical representation of* $\uparrow_X(U \cap Y)$ *(as an upwards-closed set of* $X$*).*

*Proof* Let $\bigcup_i \uparrow x_i$ be a canonical filter decomposition (in $X$) of the upwards-closed set $\uparrow_X(U \cap Y)$. We first prove that for every $i$, $x_i$ is equivalent to some element of $Y$. Indeed, since $\uparrow_X x_i \subseteq \uparrow_X(U \cap Y)$, there exists $y \in U \cap Y$ with $y \leq x_i$. But then, $y$ must be in some $\uparrow_X x_j$. Since the decomposition is canonical, the $x_j$'s are incomparable, hence we cannot have $x_j \leq y \leq x_i$ for $j \neq i$. Thus, $x_i \equiv y \in Y$.

Moreover, we can compute a canonical filter decomposition of $\uparrow_X(U \cap Y)$ using only elements in $Y$: for each $x_i$, it is decidable whether $x_i \in Y$ (our first assumption on $Y$). If not, we can enumerate elements of $Y$ until we find some $y_i \equiv x_i$. Such an element exists, and thus the enumeration terminates.

We thus obtain a canonical filter decomposition $\bigcup_i \uparrow y_i$ of $\uparrow_X(U \cap Y)$ with $y_i \in Y$. The rest of the proof is similar to the proof of Lemma 5.15. □

Here also, a canonical representation of $\uparrow_X(U \cap Y)$ is computable from $U$, using the function $\mathcal{S}_{\mathrm{F}}$.

We can now describe procedures for the four remaining operations:

**(CF)**: Given $y \in Y$, the complement of $\uparrow_Y y$ is computed by using $Y \setminus \uparrow_Y y = (X \setminus \uparrow_X y) \cap Y$. Here the downwards-closed set $(X \setminus \uparrow_X y)$ is computable using (CF) for $X$, and its intersection with $Y$ is computable using Lemma 5.15.

**(II)**: Given two ideals $I$ and $I'$ in the adherence of $Y$, the intersection of the ideals they induce is $(I \cap Y) \cap (I' \cap Y) = (I \cap I') \cap Y$, which is computed using (II) for $X$ and Lemma 5.15.

**(IF)**: Computing the intersection of filters is similar to computing the intersection of ideals: given $y_1, y_2 \in Y$, $(\uparrow_Y y_1) \cap (\uparrow_Y y_2) = (\uparrow_X y_1 \cap \uparrow_X y_2) \cap Y$, which is computed using (IF) for $X$ and Lemma 5.16.

**(CI):** Given an ideal $I$ in the adherence of $Y$, $Y \setminus (I \cap Y) = (X \setminus I) \cap Y$, which is computed using (CI) for $X$ and Lemma 5.16.

Finally, and as always, the above presentation can be computed from a presentation of $(X, \leq)$, thanks to the functions $\mathcal{S}_{\mathrm{I}}$ and $\mathcal{S}_{\mathrm{F}}$. Notably, the ideal decomposition of $Y$ can be computed with Lemma 5.15 as the set induced by $X$, seen as a downwards-closed subset, while the filter decomposition of $Y$ can be computed using Lemma 5.16, again as the set induced by $X$ seen this time as an upwards-closed subset.

*Remark 5.17* If $Y$ is a downwards-closed subset of $X$, then $I$ is adherent to $Y$ if and only if $I \subseteq Y$, and therefore $Idl(Y) = Idl(X) \cap \mathcal{P}(Y)$. Moreover, $\mathcal{S}_{\mathrm{I}}$ is computable thanks to (II), and $\mathcal{S}_{\mathrm{F}}(\uparrow x) = \uparrow x$ if $x \in Y$, $\mathcal{S}_{\mathrm{F}}(\uparrow x) = \emptyset$ otherwise. Indeed, if $x \notin Y$, then $\uparrow x \cap Y = \emptyset$.

Similarly, if $Y$ is upwards-closed, $\mathcal{S}_{\mathrm{F}}$ can be computed with (II), and $\mathcal{S}_{\mathrm{I}}(I) = I$ if $Y \cap I \neq \emptyset$, $\mathcal{S}_{\mathrm{I}}(I) = \emptyset$ otherwise. Again, $Y \cap I \neq \emptyset$ if and only if $\exists x \in \min(Y) : x \in I$. Given such an $x$, then $\forall y \in I : \exists z \in I : z \geq x, y$ by directedness. Therefore, $I \subseteq {\downarrow}(I \cap \uparrow x) \subseteq {\downarrow}(I \cap Y)$. □

# 6 Towards a Richer Theory of Ideally Effective WQOs

## 6.1 A Minimal Definition

As we mentioned in the remarks following Definition 3.1, our definition contains redundancies: some of the requirements are implied by the others. Here is the same definition in which we removed redundancies:

**Definition 6.1** (*Simply effective WQOs)* A WQO $(X, \leq)$ further equipped with data structures for $X$ and $Idl(X)$ is *simply effective* if:

**(ID)**: ideal inclusion $\subseteq$ is decidable on $Idl(X)$;
**(PI)**: principal ideals are computable, that is, $x \mapsto \downarrow x$ is computable;
**(CF)**: complementation of filters, denoted $\neg : Fil(X) \to Down(X)$, is computable;
**(II)**: intersection of ideals, denoted $\cap : Idl(X) \times Idl(X) \to Down(X)$, is computable.

A *short presentation* of $(X, \leq)$ is a list of: data structures for $X$ and $Idl(X)$, procedures for the above operations, the ideal decomposition of $X$.

Note that a short presentation of $(X, \leq)$ is obtained from a presentation of $(X, \leq)$ by dropping procedures for (OD), (CI), (IF) and by dropping (XF). Surprisingly, short presentations carry enough information:

**Theorem 6.2** *There exists an algorithm that given a short presentation of* $(X, \leq)$ *outputs a presentation of* $(X, \leq)$.

**Corollary 6.3** *A WQO* $(X, \leq)$ *(with data structures for* $X$ *and* $Idl(X)$*) is ideally effective if and only if it is simply effective.*

Before we proceed to proving Theorem 6.2, why did we bother to display full presentations of WQOs in previous sections? Our proofs of ideal effectiveness would indeed have been shorter.

Our choice is motivated by practical reasons: the algorithms we have given until now are much more efficient than the ones deduced from Theorem 6.2, which is simply impractical. (Indeed, most of these algorithms have been implemented, at the highest level of generality, by the second author.) Theorem 6.2 is more conceptual, and if one only needs computability results, then Theorem 6.2 provides a simpler path to this goal.

As practice goes, we will refine the notion of ideally effective WQOs to "efficient" ideally effective WQOs in Sect. 6.3. Most of the WQOs we have seen earlier are efficient in that sense. By contrast, the presentation of $(X, \leq)$ built from Sect. 6.3 is not *polynomial-time* (see Sect. 6.3 for a definition).

*Proof (of Theorem 6.2)* We explain how to obtain the missing procedures:

**(OD)**: Given $x, y \in X, x \leq y \iff \downarrow x \subseteq \downarrow y$. The latter can be tested using (PI) and (ID).

**(CI)**: We show a stronger statement, denoted (CD), that complementing an arbitrary downwards-closed set is computable. This strengthening is necessary for (IF).

Let $D$ be an arbitrary downwards-closed set. We compute $\neg D$ as follows:

1. Initialize $U := \emptyset$;
2. While $\neg U \not\subseteq D$ do
   (a) pick some $x \in \neg U \cap \neg D$;
   (b) set $U := U \cup \uparrow x$.

Every step of this high-level algorithm is effective. The complement $\neg U$ is computed using the description above: $\neg \bigcup_{i=1}^{n} \uparrow x_i = \bigcap_{i=1}^{n} \neg \uparrow x_i$ which is computed with (CF) and (II) (or with (XI) in case $n = 0$, i.e., for $U = \emptyset$). Then, inclusion $\neg U \subseteq D$ is tested with (ID). If this test fails, then we know $\neg U \cap \neg D$ is not empty, and thus we can enumerate elements $x \in X$ by brute force, and test membership in $U$ and in $D$. Eventually, we will find some $x \in \neg U \cap \neg D$.

To prove partial correctness we use the following loop invariant: $U$ is upwards-closed and $U \subseteq \neg D$. The invariant holds at initialization and is preserved by the loop's body since if $\uparrow x$ is upwards-closed and since $x \notin D$ and $D$ downwards-closed imply $\uparrow X \subseteq \neg D$. Thus when/if the loop terminates, one has both $\neg U \subseteq D$ and the invariant $U \subseteq \neg D$, i.e., $U = \neg D$.

Finally, the algorithm terminates since it builds a strictly increasing sequence of upwards-closed sets, which must be finite by Lemma 2.3.

**(IF)**: This follows from (CF) and (CD), by expressing intersection in terms of complement and union.

Lastly, we need to show that we can retrieve the filter decomposition of $X$. It suffices to use (CD) to compute $X = \neg\emptyset$. □

The algorithm for (CD) computes an upwards-closed set $U$ from an oracle answering queries of the form "Is $U \cap I$ empty?" for ideals $I$. It is an instance of the generalized Valk-Jantzen Lemma [26], an important tool for showing that some upwards-closed sets are computable. This algorithm was originally developed by Valk and Jantzen [59] in the specific case of $(\mathbb{N}^k, \leq_\times)$.

As seen in the above proof, the fact that (ID), (CF), (II) and (PI) entail (CI) is non-trivial. The existence of such a non-trivial redundancy in our definition raises the question of whether there are other hidden redundancies. The following theorem answers the question in the negative.

**Theorem 6.4** *For each operation A among (ID), (CF), (II) and (PI), there exists a WQO $(X_A, \leq_A)$ equipped with data structures for X and Idl$(X)$ for which operation A is not computable, while the other three are.*

This theorem means that short presentations are the shortest possible to capture the information we want. Technically, we should also argue that the ideal decomposition of $X$ cannot be retrieved from procedures for operations (ID), (CF), (II), (PI).

For a full proof of Theorem 6.4, we refer the interested reader to [32, Proposition 8.1.4]. Here we only illustrate the techniques at hand by dealing with one case.

*Example 6.5* For $n \in \mathbb{N}$ we write $T_n$ for the halting time of $M_n$, the $n$-th Turing machine (in some fixed recursive enumeration), letting $T_n = \infty$ if $T_n$ does not halt.

Let now $X_{\text{CF}} = \mathbb{N}^2$ and define an equivalence relation $E$ over $X_{\text{CF}}$ by

$$\langle n, m \rangle E \langle n', m' \rangle \overset{\text{def}}{\Leftrightarrow} n = n' \text{ and } \big(T_n < \min(m, m') \text{ or } T_n \geq \max(m, m')\big)$$

One easily checks that $E$ is compatible with the lexicographic ordering on $\mathbb{N}^2$ in the sense of Sect. 5.2, and we consider the WQO $(X_{\text{CF}}, \leq_{\text{CF}})$ with $\leq_{\text{CF}} \overset{\text{def}}{=} E \circ \leq_{\text{lex}}$. Regarding implementation, we use pairs of natural numbers to represent elements of $X_{\text{CF}}$, as well as the corresponding principal ideals. We also use a special symbol to represent the only non-principal ideal: $X_{\text{CF}}$ itself.

With this representation, $(X_{\text{CF}}, \leq_{\text{CF}})$ is almost ideally effective: deciding whether $\langle n, m \rangle \leq_{\text{CF}} \langle n, m' \rangle$ only requires simulating $M_n$ for $\max(m, m')$ steps (OD); ideal inclusion reduces to comparing elements (ID); creating $\downarrow x$ from $x$ is trivial (PI); as is representing $X_{\text{CF}}$ itself as a sum of ideals (XI).

However, $X_{\text{CF}}$ with the chosen representation does not admit an effective way of computing the complement of filters (CF): indeed the complement of some $\uparrow_X \langle n + 1, 0 \rangle$ must be some $\downarrow \langle n, m \rangle$ with $m > T_n$ if $M_n$ halts (any $m$ is correct if $M_n$ does not halt). Thus a procedure for (CF) could be used to decide the halting problem, which is impossible. □

*Remark 6.6 (On ideally effective extensions)* $(X_{\text{CF}}, \leq_{\text{CF}})$ is obtained as an extension of $(\mathbb{N}^2, \leq_{\text{lex}})$, an ideally effective WQO. This proves that extensions of ideally effective WQOs are not always ideally effective, even in the special case of a quotient by an effective compatible equivalence, and justifies the two extra assumptions we used in Theorem 5.2. More precisely, it justifies that at least one of these assumptions is necessary, and indeed, one can always compute the closure function $\mathcal{C}l_{\text{F}}$ from the closure function $\mathcal{C}l_{\text{I}}$ (but not the converse!), and this in a uniform manner. The latter result relies on an algorithm that is very similar to the generalized Valk and Jantzen Lemma. □

## 6.2 *On Alternative Effectiveness Assumptions*

The set of effectiveness assumptions collected in Definition 3.1 or Definition 4.1 is motivated by the need to perform Boolean operations on (downwards-, upwards-) closed subsets, as illustrated in our motivating examples from Sect. 2.1. Other choices are possible, and we illustrate a possible variant here.

### 6.2.1 A Natural But Not Ideally Effective Constructor

Given two QOs $(X, \leq_X)$ and $(Y, \leq_Y)$, we can define the lexicographic quasi-ordering $\leq_{\text{lex}}$ on $X \times Y$ by:

$$\langle x_1, y_1\rangle \leq_{\text{lex}} \langle x_2, y_2\rangle \overset{\text{def}}{\Leftrightarrow} x_1 <_X x_2 \vee (x_1 \equiv_X x_2 \wedge y_1 \leq_Y y_2),$$

where classically, $\equiv_X$ denotes the equivalence relation $\leq_X \cap \geq_X$ and $<_X$ denotes the strict ordering associated to $X$, defined as $\leq_X \setminus \equiv_X$.

Since $\leq_{\text{lex}}$ is coarser than the product ordering $\leq_\times$ from Sect. 4.3, $(X \times Y, \leq_{\text{lex}})$ is a WQO as soon as $\leq_X$ and $\leq_Y$ are. Besides, when $(X, \leq_X)$ and $(Y, \leq_Y)$ are ordinals, the lexicographic product corresponds to the ordinal multiplication $Y \cdot X$.

This WQO is simple and natural, but it is not always ideally effective in the sense of Definition 3.1 (at least for the natural representation of elements of $X \times Y$). The fact that our definition misses such a simple WQO constructor is disturbing and will be discussed in the next subsection. For now, let us show why lexicographic product is not an ideally effective constructor.

**Proposition 6.7** *Lexicographic product is not an ideally effective constructor. In particular, there exists an ideally effective WQO $X_{PP}$ such that $(X_{PP} \times A_2, \leq_{\text{lex}})$ is not ideally effective for any useful representation.*

*Proof* Recall from Sect. 3.1.1 that $A_2 = \{a, b\}$ is the two-letter alphabet, where $a$ and $b$ are incomparable. We use the following property: Let $(X, \leq)$ be some WQO and $I \in Idl(X)$ be one of its ideals. Then $I$ is principal if, and only if, $I \times A_2$ is not an ideal in the lexicographic product $(X \times A_2, \leq_{\text{lex}})$. Indeed, if $I = \downarrow x$ for some $x \in X_{\text{PP}}$, then $\langle x, a\rangle$ and $\langle x, b\rangle$ do not have a common upper bound in $I \times A_2$ with respect to $\leq_{\text{lex}}$, hence $I \times A_2$ is not directed. Conversely, if $I$ is not principal, then for any two elements $\langle x, c\rangle, \langle y, d\rangle \in I \times A_2$, there is some $z \in I$ such that $z > x$ and $z > y$. The element $\langle z, a\rangle$ is a suitable common upper bound, showing that $I \times A_2$ is directed.

Regarding $X_{\text{PP}}$, we refer to [32, Sect. 8.3] and do not describe it here: it is an ideally effective WQO, similar to $X_{\text{CF}}$ from Example 6.5, and for which it is undecidable whether an ideal $I$ is principal. This is enough to prove that $(X_{\text{PP}} \times A_2, \leq_{\text{lex}})$ is not ideally effective. Assume, by way of contradiction, that it is ideally effective. Then for any $I \in Idl(X)$, one can compute the ideal decomposition of $D = I \times A_2$ and then see whether this downwards-closed set is an ideal. But deciding whether $D$ is an ideal amounts to deciding whether $I$ is not principal, which is impossible in $X_{\text{PP}}$.

Note: the only representation assumption that the proof makes on $X_{\text{PP}} \times A_2$ is that the pairing function $x, c \mapsto \langle x, c\rangle$ is effective. With this assumption $I \times A_2$ can be built in the following manner: (1) compute $X_{\text{PP}} \setminus I = \uparrow x_1 + \cdots + \uparrow x_n$ in $X_{\text{PP}}$; (2) derive $(X_{\text{PP}} \setminus I) \times A_2 = U = \uparrow\langle x_1, a\rangle + \uparrow\langle x_1, b\rangle + \cdots + \uparrow\langle x_n, b\rangle$ using pairings; (3) obtain $I \times A_2$ by complementing $U$ in $X_{\text{PP}} \times A_2$, assumed to be ideally effective. □

### 6.2.2 Deciding Principality

In the previous subsection, we have shown that a very natural constructor, the lexicographic product, is not ideally effective. However, in practice $(X \times Y, \leq_{\text{lex}})$ is usually ideally effective, that is, the lexicographic product of two “actually used” WQOs $(X, \leq_X)$ and $(Y, \leq_Y)$ is ideally effective.

Thus, the problem seems to come from the fact that our definition allows too many exotic WQOs. Indeed, we can show that the lexicographic product of two ideally effective WQOs for which we can decide whether an ideal is principal, is ideally effective [32, Theorem 5.4.2]. All WQOs used in practice trivially meet this extra condition, to the point that we could argue that we should not accept as ideally effective any WQO that would not meet this requirement.

If, in the definition of ideally effective WQOs, one now adds the condition that principality of ideals be decidable, then lexicographic product becomes an ideally effective constructor, most of the constructors described in this chapter remain ideally effective, to the notable exception of extensions and quotients: Theorems 5.2 and 5.7 fail with the new definition (see [32, Sect. 8.3] for details).

### 6.2.3 Directions for Future Work

We would like to mention three directions in which our work can be extended.

The first one was carried out in [19], relying on the topological notion of notion called *Noetherian space* to generalize WQOs, in the following sense. Given a quasi-ordered set $(X, \leq)$, the Alexandroff topology has as open sets exactly the upwards-closed sets for the quasi-ordering $\leq$. It turns out that the Alexandroff topology associated to $\leq$ is Noetherian if and only if $\leq$ is a WQO on $X$. There are also Noetherian topologies that do not arise as Alexandroff topologies, for example the cofinite topology on an infinite set, or the Zariski topology on the spectrum of a Noetherian ring. One advantage of Noetherian spaces is that they are preserved under more constructors than WQOs, e.g., the full powerset of a Noetherian space (with the so-called lower Vietoris topology) is again Noetherian. In [19], the authors define a notion of effectiveness very similar to ours for Noetherian spaces, which however excludes complements and filters, which do not make sense there. Similarly, this notion of effectiveness is preserved under many constructors.

A second extension of this work was carried out in [32, Chap. 9]. The motivation is close to the one above: handling more constructors. As mentioned in Sect. 4.5, the infinite powerset $\mathcal{P}(X)$ of a WQO, ordered with the Hoare ordering is not a WQO in general. However, the class of WQOs for which $(\mathcal{P}(X), \sqsubseteq_H)$ is a WQO is well-known: these WQOs are called $\omega^2$-WQO (e.g., see [34, 48]). The second author [32] proposes a generalization of our notion of ideal effective WQOs which he calls ideal effective $\omega^2$-WQOs (also $Idl^2$-effective WQOs). He then shows that the constructors presented in this chapter also preserve this stronger notion of $Idl^2$-effectiveness, and also prove that, e.g., the powerset of an $Idl^2$-effective WQO, ordered with the Hoare quasi-ordering, is an ideally effective WQO. The notion of $\omega^2$-WQO can be generalized to the notion of $\alpha$-WQO for any indecomposable ordinal $\alpha$, eventually leading to the notion of *better quasi-ordering* ($\alpha$-WQO for every countable $\alpha$). In [32], the author raises the question on how to generalize ideal effectiveness to these classes of quasi-orderings.

Finally, one might challenge our own decision of representing upwards- and downwards-closed sets as their filter/ideal decompositions. Its main advantage is genericity: as proved in Sect. 2, this decomposition is possible in any WQO. It is also very convenient. In the simple cases of $(\mathbb{N}^k, \leq_\times)$ and $(A^*, \leq_*)$, the representations and algorithms we illustrated in Sect. 2.1 have been used for years by researchers who were not aware that they were manipulating ideals. This suggests that the idea is somehow natural.

This does not rule out the existence of better ad-hoc solutions when considering a specific WQO, notably in terms of efficiency. As will be seen in Sect. 6.3, the procedures we have presented in Sect. 4.4 have an exponential-time worst-case complexity. This exponential blow-up essentially occurs when one has to distribute the unions over the products in order to retrieve an actual filter/ideal decomposition. We are not sure this can be averted, but when one only needs to represent certain particular closed subsets of $(X^*, \leq_*)$, better representations do exist: see for instance [23].

## 6.3 On Computational Complexity

In [32], the second author provides a complexity analysis of the algorithms we have described in this chapter. Let us briefly summarize the complexity of the WQO constructors we have considered.

Formally, let us define a *polynomial-time* ideally effective WQO to be an ideally effective WQO for which there exist *polynomial-time* procedures for (OD), (ID), (CF), (IF), (CI), (II), (PI). A presentation of an ideally effective WQO is said to be *polynomial-time* if all the procedures it is composed of run in polynomial time. For instance, $\mathbb{N}$ is a polynomial-time ideally effective WQO, and the presentation we gave for it is polynomial-time. However, a WQO as simple as $(A^*, \leq_*)$, where $A = \{a, b\}$, is not polynomial-time, at least for our choice of data structure for $A^*$ and $Idl(A^*)$. Indeed, observe that the upwards-closed set $U_n = \uparrow a^n \cap \uparrow b^n$ has at least exponentially many (in $n$) minimal elements: any word with $n$ $a$'s and $n$ $b$'s is a minimal element of $U_n$. Therefore, the filter decomposition of $U_n$ is of exponential size in $n$, and thus requires exponential-time to compute.

However, for instance, the Cartesian product $(X \times Y, \leq_\times)$ of polynomial-time ideally effective WQOs is polynomial-time. (That would fail if $X$ or $Y$ were not polynomial-time: for instance, if $(X, \leq_X) = (A^*, \leq_*)$, then the upwards-closed set $\uparrow(a^n, y_1) \cap \uparrow(b^n, y_2)$ has at least exponentially many minimal elements, independently of the filter decomposition of $\uparrow y_1 \cap \uparrow y_2$.) Furthermore, from polynomial-time presentations of $(X, \leq_X)$ and $(Y, \leq_Y)$, the presentation of $(X \times Y, \leq_\times)$ we compute in Sect. 4.3 is polynomial-time as well. This motivates the following definition: an ideally effective constructor $C$ is *polynomial-time* if it is possible to compute a polynomial-time presentation for $C[(X_1, \leq_1), \ldots, (X_n, \leq_n)]$ given polynomial-time presentations of $(X_1, \leq_1), \ldots, (X_n, \leq_n)$. Note that we require that the procedures of the presentation for $C[(X_1, \leq_1), \ldots, (X_n, \leq_n)]$ are polynomial-time, but we do not make any assumption on the complexity of the procedure that builds the new presentation from presentations for each $(X_i, \leq_i)$.

With this definition in mind, here is a summary of the complexity results from [32]:

– Both disjoint sum and lexicographic sum are polynomial-time ideally effective constructors—this is a trivial analysis of the presentation of Sect. 4.2.
– Cartesian product is a polynomial-time ideally effective constructor; that again follows easily from an analysis of Sect. 4.3.
– Higman's sequence extension QO is *not* a polynomial-time ideally effective constructor. As we have seen above, already in the simple case of finite sequences over a finite alphabet, some operations require exponential time. It is not difficult to see that the presentation we gave in Sect. 4.4 consists of exponential time procedures.
– The finite powerset constructor (under the Hoare quasi-ordering) is a polynomial-time ideally effective constructor. This again follows from an easy analysis of Sect. 4.5. This justifies implementing $\mathcal{P}_f(X)$ directly, and not as a quotient of $X^*$.
– The finite multiset constructor, under multiset embedding, is an exponential-time ideally effective constructor, and already $(\mathbb{N}^{2\circledast}, \leq_{\text{emb}})$ is not a polynomial-time ideally effective WQO. However, $(A^{\circledast}, \leq_{\text{emb}})$ and $(\mathbb{N}^{\circledast}, \leq_{\text{emb}})$ are polynomial-time effective WQOs when $A$ is a finite alphabet under equality.

# 7 Concluding Remarks

We have proposed a set of effectiveness assumptions that allow one to compute with upwards-closed and downwards-closed subsets of WQOs, represented as their canonical filter and ideal decompositions respectively. These effectiveness assumptions are fulfilled in the main WQOs that appear in practical computer applications, which are built using constructors that we have shown to be ideally effective. Our algorithms unify and generalize some algorithms that have been used for many years in simple settings, such as $\mathbb{N}^k$ or the set of finite words ordered by embedding.

We have not considered any WQO constructor more complex than sequence extension, and this is an obvious direction for extending this work. How does one compute with closed subsets of finite labeled trees ordered by Kruskal's homeomorphic embedding? Or of some class of finite graphs well-quasi-ordered by some notion of embedding? The case of finite trees has already been partially tackled by the first author, see [19, 27]. The technicalities are daunting, well beyond the ambitions of this chapter, however.

**Acknowledgements** We thank Sylvain Schmitz for his many suggestions and his constant support during the several years that it took us to prepare this chapter. The final text also owes to suggestions made at various times by D. Lugiez, A. Finkel, D. Kuske and, finally, an anonymous reviewer.

## References


1. Abdulla, P. A., Atig, M. F., & Cederberg, J. (2012). Timed lossy channel systems. In *FST&TCS 2012*, LIPIcs 18 (pp. 374–386). Leibniz-Zentrum für Informatik.
2. Abdulla, P. A., Čerāns, K., Jonsson, B., & Tsay, Y.-K. (2000). Algorithmic analysis of programs with well quasi-ordered domains. *Information and Computation*, *160*(1/2), 109–127.
3. Abdulla, P. A., Collomb-Annichini, A., Bouajjani, A., & Jonsson, B. (2004). Using forward reachability analysis for verification of lossy channel systems. *Formal Methods in System Design*, *25*(1), 39–65.
4. Akama, Y. (2011). Set systems: Order types, continuous nondeterministic deformations, and quasi-orders. *Theoretical Computer Science*, *412*(45), 6235–6251.
5. Bachmeier, G., Luttenberger, M., & Schlund, M. (2015). Finite automata for the sub- and superword closure of CFLs: Descriptional and computational complexity. *LATA 2015, LNCS 8977* (pp. 473–485). Cham: Springer.
6. Bertrand, N., & Schnoebelen, P. (2013). Computable fixpoints in well-structured symbolic model checking. *Formal Methods in System Design*, *43*(2), 233–267.
7. Birkhoff, G. (1979). *Lattice Theory*, volume XXV of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, R.I., Corrected reprint of the third edition (1967).
8. Blondin, M., Finkel, A., & McKenzie, P. (2017). Well behaved transition systems. *Logical Methods in Computer Science*, *13*(3), 1–19.
9. Blondin, M., Finkel, A., & McKenzie, P. (2018). Handling infinitely branching well-structured transition systems. *Information and Computation*, *258*, 28–49.
10. R. Bonnet. (1975). On the cardinality of the set of initial intervals of a partially ordered set. In J Bolyai (Ed.) *Infinite and Finite Sets, Colloquia Mathematica Societatis, Keszthely, Hungary, North-Holland, 1973* (pp. 189–198).
11. Chen, K. T., Fox, R. H., & Lyndon, R. C. (1958). Free differential calculus, IV. The quotient groups of the lower central series. *Annals of Mathematics*, *68*(1), 81–95.
12. D'Alessandro, F., & Varricchio, S. (2008). Well quasi-orders in formal language theory. *DLT 2008, LNCS 5257* (pp. 84–95). Berlin, Heidelberg: Springer.
13. Dershowitz, N. (1987). Termination of rewriting. *Journal of Symbolic Computation*, *3*(1–2), 69–115.
14. Dershowitz, N., & Manna, Z. (1979). Proving termination with multiset orderings. *Communications of the ACM*, *22*(8), 465–476.
15. Deutsch, L. P. (1973). An interactive program verifier. Technical Report CSL-73-1, Xerox Palo Alto Research Center.
16. Finkel, A. (2016). The ideal theory for WSTS. *RP 2016, LNCS 9899* (pp. 1–22). Cham: Springer.
17. Finkel, A., & Goubault-Larrecq, J. (2009). Forward analysis for WSTS, part I: Completions. In *STACS 2009*, LIPIcs 3 (pp. 433–444). Leibniz-Zentrum für Informatik.
18. Finkel, A., & Goubault-Larrecq, J. (2009). Forward analysis for WSTS, part II: Complete WSTS. *ICALP 2009, LNCS 5556* (pp. 188–199). Berlin, Heidelberg: Springer.
19. Finkel, A., & Goubault-Larrecq, J. (2019). Forward analysis for WSTS, Part I: Completions. *Mathematical Structures in Computer Science*. Submitted. Journal version of [17].
20. Finkel, A., & Schnoebelen, P. (2001). Well-structured transition systems everywhere!. *Theoretical Computer Science*, *256*(1–2), 63–92.
21. Gallier, J. H. (1991). What's so special about Kruskal's theorem and the ordinal $\Gamma_0$? A survey of some results in proof theory. *Annals of Pure and Applied Logic*, *53*(3), 199–260.
22. Geeraerts, G., Raskin, J.-F., & Van Begin, L. (2006). Expand, enlarge and check: New algorithms for the coverability problem of WSTS. *Journal of Computer and System Sciences*, *72*(1), 180–203.
23. Geffroy, T., Leroux, J., & Sutre, G. (2017). Backward coverability with pruning for lossy channel systems. In *SPIN 2017* (pp. 132–141). ACM Press.
24. Goto, E. (1974). Monocopy and associative algorithms in an extended Lisp. Technical report, University of Tokyo, Japan, May 1974.

25. Goubault, J. (1994). HimML: Standard ML with fast sets and maps. In D. Rémy (Ed.), *5th ACM SIGPLAN Workshop on ML and its Applications, Technical Report 2265*. INRIA: Rocquencourt.
26. Goubault-Larrecq, J. (2009). On a generalization of a result by Valk and Jantzen. Research Report LSV-09-09, Laboratoire Spécification et Vérification, ENS Cachan, France, May 2009.
27. Goubault-Larrecq, J., & Schmitz, S. (2016). Deciding piecewise testable separability for regular tree languages. In *ICALP 2016*, LIPIcs 55 (pp 97:1–97:15). Leibniz-Zentrum für Informatik.
28. Haase, C., Schmitz, S., & Schnoebelen, P. (2014). The power of priority channel systems. *Logical Methods in Computer Science*, *10*(4:4).
29. Habermehl, P., Meyer, R., & Wimmel, H. (2010). The downward-closure of Petri net languages. *ICALP 2010, LNCS 6199* (pp. 466–477). Berlin, Heidelberg: Springer.
30. Haddad, S., Schmitz, S., & Schnoebelen, P. (2012). The ordinal-recursive complexity of timed-arc Petri nets, data nets, and other enriched nets. In *LICS 2012* (pp. 355–364). IEEE Computer Society Press.
31. Haines, L. H. (1969). On free monoids partially ordered by embedding. *Journal of Combinatorial Theory*, *6*(1), 94–98.
32. Halfon, S. (2018). *On Effective Representations of Well Quasi-Orderings*. École Normale Supérieure Paris-Saclay, France: Thèse de doctorat.
33. Higman, G. (1952). Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society*, *2*(7), 326–336.
34. Jančar, P. (1999). A note on well quasi-orderings for powersets. *Information Processing Letters*, *72*(5–6), 155–161.
35. Kabil, M., & Pouzet, M. (1992). Une extension d'un théorème de P. Jullien sur les âges de mots. *RAIRO Theoretical Informatics and Applications*, *26*(5), 449–482.
36. Karp, R. M., & Miller, R. E. (1969). Parallel program schemata. *Journal of Computer and System Sciences*, *3*(2), 147–195.
37. Křížž, I., & Thomas, R. (1990). On well-quasi-ordering finite structures with labels. *Graphs and Combinatorics*, *6*(1), 41–49.
38. Kruskal, J. B. (1960). Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture. *Transactions of the American Mathematical Society*, *95*(2), 210–225.
39. Kruskal, J. B. (1972). The theory of well-quasi-ordering: A frequently discovered concept. *Journal of Combinatorial Theory, Series A*, *13*(3), 297–305.
40. Lazić, R., Newcomb, T., Ouaknine, J., Roscoe, A. W., & Worrell, J. (2008). Nets with tokens which carry data. *Fundamenta Informaticae*, *88*(3), 251–274.
41. Lazić, R., & Schmitz, S. (2015). The ideal view on Rackoff's coverability technique. *RP 2015, LNCS 9328* (pp. 76–88). Cham: Springer.
42. Lazić, R., & Schmitz, S. (2016). The complexity of coverability in $\nu$-Petri nets. In *LICS 2016* (pp. 467–476). ACM Press.
43. Leroux, J., & Schmitz, S. (2015). Demystifying reachability in vector addition systems. In *LICS 2015* (pp. 56–67). IEEE Computer Society Press.
44. Leroux, J., & Schmitz, S. (2016). Ideal decompositions for vector addition systems. In *STACS 2016*, LIPIcs 47 (pp. 1:1–1:13). Leibniz-Zentrum für Informatik.
45. Leuschel, M. (2002). Homeomorphic embedding for online termination of symbolic methods. *The Essence of Computation, Complexity, Analysis, Transformation, LNCS 2566* (pp. 379–403). Berlin, Heidelberg: Springer.
46. Lovász, L. (2006). Graph minor theory. *Bulletin of the American Mathematical Society*, *43*(1), 75–86.
47. Manolios, P., & Vroon, D. (2005). Ordinal arithmetic: Algorithms and mechanization. *Journal of Automated Reasoning*, *34*(4), 387–423.
48. Marcone, A. (2001). Fine analysis of the quasi-orderings on the power set. *Order*, *18*(4), 339–347.
49. Pin, J. É. (1996). Logic, semigroups and automata on words. *Annals of Mathematics and Artificial Intelligence*, *16*(1), 343–384.
50. Rado, R. (1954). Partial well-ordering of sets of vectors. *Mathematika*, *1*(2), 89–95.

51. Robertson, N., & Seymour, P. D. (2004). Graph minors. XX. Wagner's conjecture. *Journal of Combinatorial Theory, Series B*, *92*(2), 325–357.
52. Sacks, G. E. (1990). Constructive ordinals and $\Pi^1_1$ sets. In *Higher Recursion Theory*, *Perspectives in Mathematical Logic* (Chap. 1, Vol. 2 pp. 3–21). Berlin, Heidelberg: Springer.
53. J. Sakarovitch & I. Simon. Subwords. In M. Lothaire (Ed.) *Combinatorics on Words*, *Encyclopedia of Mathematics and Its Applications* (Chap. 6, Vol. 17, pp. 105–142). Cambridge University Press (1983).
54. Schmitz, S. (2014). Complexity bounds for ordinal-based termination. *RP 2014, LNCS 8762* (pp. 1–19). Cham: Springer.
55. Schmitz, S. (2016). Complexity hierarchies beyond Elementary. *ACM Transactions on Computation Theory*, *8*(1),
56. Schmitz, S., & Schnoebelen, P. (2011). Multiply-recursive upper bounds with Higman's lemma. *ICALP 2011, LNCS 6756* (pp. 441–452). Berlin, Heidelberg: Springer.
57. Schmitz, S., & Schnoebelen, P. (2012). Algorithmic Aspects Of WQO Theory. Lecture Notes. https://cel.archives-ouvertes.fr/cel-00727025.
58. Schmitz, S., & Schnoebelen, P. (2013). The power of well-structured systems. *CONCUR 2013, LNCS 8052* (pp. 5–24). Berlin, Heidelberg: Springer.
59. Valk, R., & Jantzen, M. (1985). The residue of vector sets with applications to decidability problems in Petri nets. *Acta Informatica*, *21*(6), 643–674.
60. Zetzsche, G. (2015). Computing downward closures for stacked counter automata. In *STACS 2015*, LIPIcs 30 (pp. 743–756). Leibniz-Zentrum für Informatik.
61. Zetzsche, G. (2018). Separability by piecewise testable languages and downward closures beyond subwords. In *LICS 2018* (pp. 929–938). ACM Press.

# Strong WQO Tree Theorems

**Lev Gordeev**

**Abstract** Ordinal (vertex- and/or edge-) labeled finite trees are well-quasi-ordered by homeomorphic embeddability with sound gap-conditions. Such *strong* generalizations of Harvey Friedman's tree theorem (abbr.: FT) on trees whose vertices are labeled by bounded natural numbers are (a) provable in second-order arithmetic $\Pi^1_1$-$\mathbf{TR}_0$ (also designated $\mathbf{ITR}_0$ below) that extends $\mathbf{ACA}_0$ by transfinite iteration of $\Pi^1_1$-comprehension along arbitrary countable ordinals but (b) not provable in a subsystem thereof that arises by weakening $\Pi^1_1$-transfinite recursion axiom to $\Pi^1_1$-transfinite recursion rule. In particular, I. Křiž's tree theorem (abbr.: KřT) referring to ordinal edge-labeled trees [9] is provable in $\Pi^1_1$-$\mathbf{TR}_0$ (that is weaker than theory $\Pi^1_2$-$\mathbf{CA}$ implicitly used in [9]), which is the main result of the paper. Moreover KřT is proof-theoretically equivalent to the author's analogous theorem (abbr.: GT) referring to ordinal vertex-labeled trees under symmetric gap-condition [6]. Namely, both theorems characterize $\mathbf{ITR}_0$ in the sense of ordinal provability over $\mathbf{ACA}_0$. That is, the supremum of proof-theoretic ordinals provable in $\mathbf{ACA}_0$ extended by GT and/or KřT is the proof-theoretic ordinal of $\mathbf{ITR}_0$ [in symbols: $|\mathbf{ACA}_0 + GT| = \left|\mathbf{ACA}_0 + K\check{r}T\right| = \left|\Pi^1_1\text{-}\mathbf{TR}_0\right| = \psi 0\,(\Phi 10)$ (see [12] for the last equality)]. By contrast, the restricted GT and KřT referring to ordinal labeled intervals (i.e. non-branching trees) both yield analogous characterizations of a weaker (predicative) theory $\mathbf{ATR}_0$, instead of $\mathbf{ITR}_0$ (cf. [5]).

## 1 Introduction

Kruskal's tree theorem (abbr.: KT) states that finite trees are well-quasi-ordered under the homeomorphic embeddability [10]. Beside for its mathematical transparency, this theorem has applications in computer science (eg. in the theory of

Basic version of this paper was presented at 14th British Combinatorial Conference held in University of Keele, 5–9 July, 1993.

L. Gordeev (✉)
Universität Tübingen, Ghent University, Ghent, Belgium
e-mail: lew.gordeew@uni-tuebingen.de

 
P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_4

rewriting,break cf. [1, 3]). However, the most prominent feature of KT is its proof-theoretic strength that exceeds the one of the predicative analysis $\mathbf{ATR}_0$ that is well-known in mathematical logic and foundations of mathematics. In fact $|\mathbf{ACA}_0 + KT| = \vartheta\Omega^{\omega} > \Gamma_0 = |\mathbf{ATR}_0|$ [13]. Thus KT is not provable by predicative means—this fundamental result is due to Harvey Friedman (cf. [14]), who also considered finite trees labeled by bounded natural numbers under the homeomorphic embeddability with certain asymmetric gap-condition, and proved that the corresponding well-quasi-ordering (abbr.: wqo) property is true but not provable in the theory of $\Pi^1_1$-comprehension axiom with arithmetical induction, $\Pi^1_1$-$\mathbf{CA}_0$, that is stronger than $\mathbf{ATR}_0$ (in particular, KT is provable in $\Pi^1_1$-$\mathbf{CA}_0$). To put it more precisely, let $\langle T, \ell \rangle$ denote any $n$-bounded labeled tree in question, where $T$ is a finite tree[1] with labeling function $\ell : \mathrm{V}(T) \to n$, while $\mathrm{V}(T)$ is the set of nodes (vertices) in $T$ and $n = \{0, \cdots, n-1\}$. Now $\langle T_1, \ell_1 \rangle \leq_{\mathrm{A}} \langle T_2, \ell_2 \rangle$ if there is a homeomorphic[2] label-preserving embedding $f : T_1 \to T_2$ that satisfies the following asymmetric gap-condition

($a$) If $\langle x, y \rangle$ is an edge in $T_1$ and $u \in \mathrm{V}(T_2)$ any vertex between $f(x)$ and $f(y)$, in $T_2$, then $\ell_2(u) = \ell_1(y)$.

**Theorem 1** ([14]) *Let $n > 0$ be fixed. The set of n-bounded finite labeled trees is well-quasi-ordered by $\leq_{\mathrm{A}}$. Moreover, this statement, FT, is a theorem of $\Pi^1_1$-$\boldsymbol{CA}_0$. However, the corresponding universal statement* "for every $n > 0$, the set of $n$-bounded finite labeled trees is well-quasi-ordered by $\leq_{\mathrm{A}}$" *is not provable in $\Pi^1_1$-$\boldsymbol{CA}_0$.*

In order to generalize this result, the author considered [5, 6] finite trees labeled with arbitrary (countable) ordinals, instead of natural numbers, under the homeomorphic embeddability $\langle T_1, \ell_1 \rangle \leq_{\mathrm{S}} \langle T_2, \ell_2 \rangle$ determined by the existence of $f : T_1 \to T_2$ that does not decrease the labels, i.e. $\ell_1(x) \leq \ell_2(f(x))$, and satisfies the following symmetric gap-condition

($s$) If $\langle x, y \rangle$ is an edge in $T_1$ and $u \in \mathrm{V}(T_2)$ any vertex between $f(x)$ and $f(y)$, in $T_2$, then $\min\{\ell_1(x), \ell_1(y)\} \leq \ell_2(u)$.

**Theorem 2** ([6]) *The set of finite trees with ordinal-labeled vertices is well-quasi-ordered by $\leq_{\mathrm{S}}$. Moreover, this statement, GT, is provable in impredicative theory $\boldsymbol{ITR}_0$ (that is defined analogously to its predicative counterpart, $\boldsymbol{ATR}_0$, with respect to $\Pi^1_1$-transfinite recursion) but not in a weaker theory that arises by weakening $\Pi^1_1$-transfinite recursion axiom to $\Pi^1_1$-transfinite recursion rule.*

Actually, it follows that GT characterizes $\mathbf{ITR}_0$ in the sense of ordinal provability over $\mathbf{ACA}_0$, i.e. $|\mathbf{ACA}_0 + GT| = |\mathbf{ITR}_0| = \psi 0 (\Phi 10) > \psi 0 \Omega_{\omega} = \left|\Pi^1_1 - \mathbf{CA}_0\right|$ (cf. Abstract and [11, 12]). Hence, proof-theoretically, GT is much stronger than FT. By contrast, [5] shows that GT restricted to intervals analogously characterizes $\mathbf{ATR}_0$ that, in turn, is too weak to prove even (unlabeled) KT (see above).

[1] We consider upward directed structured rooted trees whose roots are the bottom nodes.

[2] Homeomorphisms in question preserve branching order.

Meanwhile, Harvey Friedman conjectured that finite trees with ordinal-labeled edges (instead of vertices) are well-quasi-ordered by the homeomorphic embeddability $\langle T_1, \ell_1 \rangle \leq_{\mathrm{E}} \langle T_2, \ell_2 \rangle$ determined by the existence of $f : T_1 \to T_2$ satisfying the following gap-condition

($e$) Every edge $x$ in $T_1$ is mapped by $f$ onto a path in $T_2$ consisting of edges $u$ such that $\ell_1(x) \leq \ell_2(u)$.

This conjecture was proved by I. Křiž [9] at about the same time as the authors's Theorem 2, although using different ideas that can be formalized in the theory of $\Pi^1_2$-comprehension, $\Pi^1_2$-**CA**, but certainly not in $\mathbf{ITR}_0$.

**Theorem 3** (**KřT** [9]) *The set of finite trees with ordinal-labeled edges is well-quasi-ordered by* $\leq_{\mathrm{E}}$.

In the present paper we show that $\Pi^1_2$-comprehension used in [9] is an "overkill" and demonstrate by the method used in [6] that KřT is actually provable in $\mathbf{ITR}_0$. Hence KřT also characterizes $\mathbf{ITR}_0$ in the sense of ordinal provability over $\mathbf{ACA}_0$. As compared to analogous proof in [6], here we use somewhat different notions of open, closed and nopen trees, in order to make the method work in the case of edge-labeled trees.

By the same token, similar modification of [5] shows that KřT restricted to intervals analogously characterizes $\mathbf{ATR}_0$. Loosely speaking this is because in the case of intervals we can use transfinite iteration of first-order comprehension, instead of $\Pi^1_1$-comprehension being enforced by the branching structure of trees proper (we omit the details).

## 2 Labeling Trees with Ordinals

### *2.1 Basic Notations*

A *finite rooted tree* is a finite partial order $T = \langle S, \trianglelefteq \rangle, \emptyset \neq S \subset_{fin} \mathbb{N}$, with (uniquely determined) minimal element $r_T \in S$ satisfying $(\forall y \in S)\,(r_T \trianglelefteq y)$ and such that for any $x \in S$, $\{y \in S : y \triangleleft x\}$ is linearly ordered by $\triangleleft$. Note that any $x, y \in S$ uniquely determine $\inf\{x, y\} \in S$ such that $\inf\{x, y\} \trianglelefteq x$, $\inf\{x, y\} \trianglelefteq y$ and $(\forall z \in S)\,((z \trianglelefteq x \wedge z \trianglelefteq y) \to z \trianglelefteq \inf\{x, y\})$.

Let $T = \langle S, \trianglelefteq \rangle$ be fixed. Elements of $S$ are called *vertices*, or *nodes*, of $T$. Vertex $r_T$ is called the *root* of $T$. If $(\forall y \in S)\,(x \not\triangleleft y)$ then $x$ is called an *end-node*, or a *leaf*, in $T$. If $x \triangleleft y$, i.e. $x \trianglelefteq y \neq x$, then $x$ is said to occur *below* $y$. If $x \triangleleft y$ and $(\nexists z \in S)\,(x \triangleleft z \triangleleft y)$ then $x$ is called the (uniquely determined) *parent* of $y$ (abbr.: $x = \mathrm{p}(y)$). A $y \in S$ such that $\mathrm{p}(y) = r_T$ is called a *root-neighbor*. Pairs $\mathrm{e}(y) := \langle \mathrm{p}(y), y \rangle$ are called *edges* of $T$. In the sequel we let $\mathrm{V}(T) := S$ and denote by $\mathrm{E}(T)$ and $\mathrm{L}(T)$ the sets of edges and leaves of $T$, respectively.

For any pair of trees $T_1 = \langle S_1, \trianglelefteq_1 \rangle, T_2 = \langle S_2, \trianglelefteq_2 \rangle$ let $\text{HEM} \ni f : T_1 \to T_2$ express that is a *homeomorphic embedding* of $T_1$ into $T_2$, i.e. a monomorphism preserving $\inf\{-, -\}$ and the order of every branching. The latter condition means that if $x = \mathrm{p}(y) = \mathrm{p}(z)$ in $T_1$, $f(x) = \mathrm{p}(u) = \mathrm{p}(v)$ in $T_2$, $u \trianglelefteq_2 f(y)$ and $v \trianglelefteq_2 f(z)$, then $y < z$ implies $u < v$ (in $\mathbb{N}$).

## 2.2 Vertex-Labeled Trees

Let $\mathcal{O} = \langle P, \preceq \rangle$, $P \subseteq \mathbb{N}$, be a fixed (countable) well-order. A *vertex-labeled tree* (abbr.: *vlt*) relative to $\mathcal{O}$ is a structure $V = \langle T, \ell \rangle$, $T = \langle S, \trianglelefteq \rangle$ being a tree and $\ell : \mathrm{V}(T) \to P$ a labeling function. We define the embeddability relations $\leq_{\mathrm{S}}$ and $\leq_{\mathrm{A}}$ on *vlt*'s.

**Definition 4** Let $V_1 = \langle T_1, \ell_1 \rangle = \langle S_1, \trianglelefteq_1, \ell_1 \rangle$ and $V_2 = \langle T_2, \ell_1 \rangle = \langle S_2, \trianglelefteq_2, \ell_2 \rangle$ be any vlt's. Let $f : V_1 \leq_{\mathrm{S}} V_2$ abbreviate the conjunction of the following conditions 1–3, where $x, y$ and $u$ are ranging over $S_1$ and $S_2$, respectively, while $\mathrm{p}_1(-)$ is $\mathrm{p}(-)$ of $T_1$ (if defined). Let $V_1 \leq_{\mathrm{S}} V_2$ abbreviate $\exists f : V_1 \leq_{\mathrm{S}} V_2$.

1. $\text{HEM} \ni f : T_1 \to T_2$ and $f\left(\mathrm{L}(T_1)\right) \subseteq \mathrm{L}(T_2)$.
2. $\ell_1(x) \preceq \ell_2(f(x))$.
3. If $f(\mathrm{p}_1(y)) \lhd_2 u \lhd_2 f(y)$ then $\min\{\ell_1(\mathrm{p}_1(y)), \ell_1(y)\} \preceq \ell_2(u)$.
   The embeddability relation $\leq_{\mathrm{A}}$ is defined analogously using conditions 1, 2 and 4, where
4. If $f(\mathrm{p}_1(y)) \lhd_2 u \lhd_2 f(y)$, then $\ell_1(y) \preceq \ell_2(u)$.

Conditions 3 and 4 are referred to as *symmetric* and *asymmetric gap-conditions*, respectively.[3] Note that $\leq_{\mathrm{A}}$ implies $\leq_{\mathrm{S}}$.

## 2.3 Edge-Labeled Trees

Let $\mathcal{O} = \langle P, \preceq \rangle$ be as above. An *edge-labeled tree* (abbr.: *elt*) relative to $\mathcal{O}$ is a structure $E = \langle T, \ell \rangle, T = \langle S, \trianglelefteq \rangle$ being a tree and $\ell : \mathrm{E}(T) \to P$ a labeling function. The embeddability relation $\leq_{\mathrm{E}}$ on *elt*'s is defined as follows (cf. [9]).

**Definition 5** Let $E_1 = \langle T_1, \ell_1 \rangle = \langle S_1, \trianglelefteq_1, \ell_1 \rangle$ and $E_2 = \langle T_2, \ell_1 \rangle = \langle S_2, \trianglelefteq_2, \ell_2 \rangle$ be any *elt*'s. Let $f : E_1 \leq_{\mathrm{E}} E_2$ abbreviate the conjunction of the following two conditions, where $y$ and $v$ are ranging over $S_1$ and $S_2$, respectively, while $\mathrm{p}_i(-)$ and $\mathrm{e}_i(-)$ are $\mathrm{p}(-)$ and $\mathrm{e}(-)$ of $T_i$, respectively. Let $E_1 \leq_{\mathrm{E}} E_2$ abbreviate $\exists f : E_1 \leq_{\mathrm{E}} E_2$.

[3] This asymmetric gap-condition, due to M. Okada, upgrades Friedman's asymmetric gap-condition (a) (see Introduction) that fails in the infinite ordinal domain (cf. [5, 7]).

1. $\textsc{Hem} \ni f : T_1 \to T_2$ and $f\left(\mathrm{L}(T_1)\right) \subseteq \mathrm{L}(T_2)$.
2. If $f(\mathrm{p}_1(y)) \lhd_2 v \unlhd_2 f(y)$ then $\ell_1(\mathrm{e}_1(y)) \preceq \ell_2(\mathrm{e}_2(v))$.

Condition 2 is referred to as the *edge-gap-condition*.

## 2.4 Complex Edge-Labeled Trees

Let $\mathcal{O} = \langle P, \preceq \rangle$ be as above. A *complex edge-labeled tree* (abbr.: *celt*) relative to $\mathcal{O}$ is a structure $C = \langle T, M, \ell \rangle$, $T = \langle S, \unlhd \rangle$ being a tree, $\ell : \mathrm{E}(T) \to P$ a labeling function and $M \subseteq S \setminus \{r_T\}$ a set of distinguished vertices called *marks*, provided that for any $x, y \in S$ the following two conditions hold. (Note that *elt*'s are *celt*'s with $M = \emptyset$.)

1. If $y \in M$ and $r_T \neq x \lhd y$, then $\ell(\mathrm{e}(y)) \prec \ell(\mathrm{e}(x))$.
2. If $y \in M$, $y \lhd x$ and $(\nexists z \in M)(y \lhd z \unlhd x)$, then $\ell(\mathrm{e}(y)) \preceq \ell(\mathrm{e}(x))$.

Below we'll often rename $r_T$ as $r_C$. The corresponding reflexive (but not necessarily transitive!) embeddability relation $\leq_{\mathrm{C}}$ is defined as follows.

**Definition 6** Let $C_1 = \langle T_1, M_1, \ell_1 \rangle = \langle S_1, \unlhd_1, M_1, \ell_1 \rangle$ and $C_2 = \langle T_2, M_2, \ell_1 \rangle = \langle S_2, \unlhd_2, M_2, \ell_2 \rangle$ be any *celt*'s. Let $f : C_1 \leq_{\mathrm{C}} C_2$ abbreviate the conjunction of the following three conditions, where $y$ and $v$ are ranging over $S_1$ and $S_2$, respectively, while $\mathrm{p}_i(-)$ and $\mathrm{e}_i(-)$ are $\mathrm{p}(-)$ and $\mathrm{e}(-)$ of $T_i$, respectively. Let $C_1 \leq_{\mathrm{C}} C_2$ abbreviate $\exists f : C_1 \leq_{\mathrm{C}} C_2$.

1. $\textsc{Hem} \ni f : T_1 \to T_2$ and $f\left(\mathrm{L}(T_1)\right) \subseteq \mathrm{L}(T_2)$.
2. If $f(\mathrm{p}_1(y)) \lhd_2 v \unlhd_2 f(y)$ then $\ell_1(\mathrm{e}_1(y)) \preceq \ell_2(\mathrm{e}_2(v))$.
3. If $\mathrm{p}_1(y) = r_{T_1}$ and $(\exists w \in M_2)(w \unlhd_2 v \unlhd_2 f(y))$, then $\ell_1(\mathrm{e}_1(y)) \preceq \ell_2(\mathrm{e}_2(v))$.

## 2.5 Main Propositions

**Definition 7** Let $\mathcal{O} = \langle P, \preceq \rangle$ be as above. Let $\mathcal{S} = \{\mathcal{S}(0), \cdots, \mathcal{S}(n), \cdots\}$ be an infinite sequence of *vlt*'s, relative to $\mathcal{O}$. $\mathcal{S}$ is called $\leq_{\mathrm{S}}$-*bad* if $\mathcal{S}(i) \nleq_{\mathrm{S}} \mathcal{S}(j)$ fails for all $i < j \in \mathbb{N} = \{0, 1, \cdots\}$. The $\leq_{\mathrm{A}}$-*bad* sequences of *vlt*'s, $\leq_{\mathrm{E}}$-*bad* sequences of *elt*'s and $\leq_{\mathrm{C}}$-*bad* sequences of *celt*'s, relative to $\mathcal{O}$, are defined analogously.

Let *Propositions* $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$ express that for any well-order $\mathcal{O}$, there is no $\leq_{\mathrm{S}}$ -bad sequence of *vlt*'s, no $\leq_{\mathrm{A}}$ -bad sequence of *vlt*'s, no $\leq_{\mathrm{E}}$ -bad sequence of *elt*'s and no $\leq_{\mathrm{C}}$ -bad sequence of *celt*'s, respectively, relative to $\mathcal{O}$.

Propositions **A**, **B** and **C** respectively say that the corresponding orders on *vlt*'s and *elt*'s are wqo's. Recall that **A** and **C** are true propositions GT and KřT according to [6, 9], respectively (cf. Theorems 2, 3), while **B** is stronger than **A** as $\leq_A$ implies $\leq_S$. Moreover, by Theorem 2, Proposition **A** is provable in $\mathbf{ITR}_0$ that extends $\mathbf{ACA}_0$ (being a second-order conservative extension of Peano Arithmetic) by the axiom expressing in the second-order language that for any well-order $\mathcal{O}$, there exists the hyperjump-hierarchy of sets along $\mathcal{O}$ (cf. e.g. [5, 6]; see also [2, 11, 12, 14] for more information about subsystems of analysis in question).

**Main Theorem**. *Propositions **A**, **B** and **C** are all provable in* $\mathbf{ITR}_0$

In what follows we prove that Proposition **C** is a theorem of $\mathbf{ITR}_0$ (which is the paper's main result), while in the final chapter we'll show that so is **D**, too. The remainder follows from

**Lemma 8** *Proposition **C** infers Proposition **B**, and hence **A**, provably in* $\mathbf{ACA}_0$.

*Proof* Assuming **C** suppose that $\mathcal{S} = \{\mathcal{S}(0), \cdots, \mathcal{S}(n), \cdots\}$ is an infinite $\leq_A$-bad sequence of *vlt*'s, relative to a given well-order $\mathcal{O} = \langle P, \preceq \rangle$. For any $n \geq 0$, let $\rho(n) \in P$ be the root-label of $\mathcal{S}(n)$. Arguing in $\mathbf{ACA}_0$ there is a strictly increasing function $\varphi : \mathbb{N} \to \mathbb{N}$ such that $(\forall i < j)\, \rho(\mathcal{S}(\varphi(i))) \preceq \rho(\mathcal{S}(\varphi(j)))$. Let $\mathcal{S}_\varphi := \mathcal{S} \circ \varphi = \{\mathcal{S}(\varphi(0)) = \langle T_0, \ell_0 \rangle, \cdots, \mathcal{S}(\varphi(n)) = \langle T_n, \ell_n \rangle, \cdots\}$. Clearly $\mathcal{S}_\varphi$ is an infinite $\leq_A$-bad subsequence of $\mathcal{S}$. Let $\mathcal{O}'$ be a minimal well-ordered extension of $\mathcal{O}$ and $\mathcal{O}_1 = \langle P_1, \preceq_1 \rangle := \mathcal{O}' \oplus \mathcal{O}$ be the corresponding (well-ordered) disjoint sum. The ordinal of $\mathcal{O}_1$ is $(\sigma + 1) + \sigma$, where $\sigma = \sup \mathcal{O}$. Let $\mathcal{S}_1 = \left\{\mathcal{S}_1(0) = \left\langle T_0, \ell_0^1 \right\rangle, \cdots, \mathcal{S}_1(n) = \left\langle T_n, \ell_n^1 \right\rangle, \cdots\right\}$ be an infinite sequence of *vlt*'s relative to $\mathcal{O}_1$ that arises by replacing, in every $T_n$, the labels $\ell_n(x) \in P$ of the root-neighbors $x$ by the labels $\ell_n^1(x)$ corresponding to $\sigma + \ell_n(x)$ in $P_1$. We claim that $\mathcal{S}_1$ is $\leq_A$-bad relative to $\mathcal{O}_1$. For suppose $h : \left\langle T_i, \ell_i^1 \right\rangle \leq_A \left\langle T_j, \ell_j^1 \right\rangle$ holds for some $i < j$. Then, by conditions 2, 4 of Definition 4, we conclude that $h\left(r_{T_i}\right) = r_{T_j}$, while for any root-neighbor $x$ of $r_{T_i}$, $h(x)$ is a root neighbor of $r_{T_j}$, as $\max\left\{Rng(\ell_i), Rng\left(\ell_j\right)\right\} \prec_1 \sigma$. But this easily implies $h : \langle T_i, \ell_i \rangle \leq_A \left\langle T_j, \ell_j \right\rangle$—a contradiction to the $\leq_A$-badness of $\mathcal{S}_\varphi$. Now let $\mathcal{S}_2 = \left\{\mathcal{S}_2(0) = \left\langle T_0, \ell_0^2 \right\rangle, \cdots, \mathcal{S}_2(n) = \left\langle T_n, \ell_n^2 \right\rangle, \cdots\right\}$ be an infinite sequence of *elt*'s relative to $\mathcal{O}_1$ that arises by setting $\ell_n^2(\mathrm{e}(y)) := \ell_n^1(y)$ for all $n$ and $y \in \mathrm{V}(T_n) \setminus \left\{r_{T_n}\right\}$. The $\leq_A$-badness of $\mathcal{S}_1$ easily implies that $\mathcal{S}_2$ is $\leq_E$-bad relative to $\mathcal{O}_1$—a contradiction to Proposition **C**. ■

*Remark 9* Theorem 2, Main Theorem and Lemma 8 reveal that $\mathbf{ACA}_0$ extended by Proposition **C** (and/or Proposition **B**) has the same proof-theoretic strength as $\mathbf{ITR}_0$, since this is the case of Proposition **A** ([4, 6, 7], see also [9]). Thus Propositions **A**, **B** and **C** are equivalent modulo ordinal provability over $\mathbf{ACA}_0$. In particular, none of Propositions **A**, **B**, **C** is provable in the weakening of $\mathbf{ITR}_0$ that arises by replacing $\Pi_1^1$-transfinite recursion axiom by $\Pi_1^1$-transfinite recursion rule.

In the sequel we use Higman theorem in the following "reflexive" form (it has the same proof-theoretic strength as classical wqo result from [8]).

**Theorem 10** (Higman) *Let $Q = \langle S, R\rangle$, $R \subset S \times S$, be any reflexive order that has no infinite R-bad sequence, i.e. no $\{x_n \in S\}_{n\in\omega}$ such that $(\nexists i < j)\, x_i R x_j$. Let $Q^* := \langle S^*, R^*\rangle$, where $S^*$ consists of all finite tuples $\langle x_0, \cdots, x_n\rangle$, $x_i \in S$, and $R^* \subset S^* \times S^*$ is given by*

$$\langle x_0, \cdots, x_n\rangle R^* \left\langle x'_0, \cdots, x'_m\right\rangle :\Leftrightarrow (\exists 0 \le \zeta(0) < \cdots < \zeta(n) \le m)\, (\forall i \le m)\, x_i R x'_{\zeta(i)}.$$

*Then $Q^*$ has no infinite $R^*$-bad sequence.*

## 3 Proof of Proposition C. Part 1

### *3.1 Basic Definitions and Notations*

To begin with we specify complex edge-labeled trees to be used in the proof.

**Definition 11** Let $\mathcal{O} = \langle P, \preceq\rangle$ be a fixed well-order. Arbitrary elements of $P$ ($\delta$, $\sigma$, $\tau$, $\nu$, etc.) are also called ordinals. A *regular celt* (abbr.: *relt*) relative to $\sigma \preceq \nu \in P$ is a structure $C = \langle T, M, \ell\rangle$, $T = \langle S, \trianglelefteq\rangle$ being a tree and $\ell : \mathrm{E}(T) \to P$ a labeling function such that for any $r_T \neq x, y \in S$ the following five conditions hold.

1. $\ell(\mathrm{e}(x)) \prec \nu$.
2. If $y \in M$ and $x \vartriangleleft y$, then $\ell(\mathrm{e}(y)) \prec \ell(\mathrm{e}(x))$.
3. If $y \in M$, $y \vartriangleleft x$ and $(\nexists z \in M)\, (y \vartriangleleft z \trianglelefteq x)$, then $\ell(\mathrm{e}(y)) \preceq \ell(\mathrm{e}(x))$.
4. If $y \in M$ then $\ell(\mathrm{e}(y)) \prec \sigma$.
5. If $(\nexists z \in M)\, (z \trianglelefteq y)$ then $\sigma \preceq \ell(\mathrm{e}(y))$.

**Notations 12** Let $\mathcal{R}^{\nu}_{\sigma}[\mathcal{O}]$ denote the set of *relt*'s relative to $\sigma \preceq \nu$. For any *celt*'s (*relt*'s) $C_i = \langle T_i, M_i, \ell_i\rangle = \langle S_i, \trianglelefteq_i, M_i, \ell_i\rangle$, $i \in \{\emptyset, 1, 2\}$, by $C_1 \subseteq C_2$ we abbreviate that $C_1$ is a *subcelt* (*subrelt*) of $C_2$, i.e. $r_{C_1} \in S_2$, $S_1 = \left\{x \in S_2 : r_{C_1} \trianglelefteq_2 x\right\}$, $\trianglelefteq_1 = \trianglelefteq_2\restriction_{S_1}$, $\ell_1 = \ell_2 \restriction_{\mathrm{E}(T_1)}$ and $M_1 = M_2 \cap S_1$. In other words, $C_1 \subseteq C_2$ iff $r_{C_1} \in S_2$ and $C_1 = (C_2)_{r_{C_1}}$, where $(C)_x$ denotes the *subcelt* of $C$ rooted in $x$. The *subcelt*'s (*subrelt*'s) are supposed to be ordered (say, lexicographically) by a linear size-preserving relation $<_{\mathrm{LEX}}$ (thus $C_1 \subsetneq C_2$ implies $C_1 <_{\mathrm{LEX}} C_2$, where as usual $\subsetneq_{(-)}$ stands for $\subseteq_{(-)} \cap \neq$). Furthermore:

- $C$ is called *open* if it has a root-neighbor $x \notin M$; otherwise $C$ is called *nopen*. A nopen $C$ is *closed* if it has just one root-neighbor $x$ (hence $x \in M$); in this case, we denote by $e_C$ and $\rho_C$ the (uniquely determined) *root-edge* $\mathrm{e}(x) = \langle r_C, x\rangle$ and *root-ordinal* $\ell(e_C)$, respectively. A closed $C$ is *closed below* (resp. *above*) $\delta \in P$ if $\ell(e_C) \prec \delta$ (resp. $\ell(e_C) \succeq \delta$). Thus any closed $C \in \mathcal{R}^{\nu}_{\sigma}[\mathcal{O}]$ is closed below $\nu$.
- $y \in S$ occurs in the *root-piece* of $C$ (abbr.: $y \in_{\mathrm{RP}} C$) if $(\nexists x \in M)\, x \trianglelefteq y$. $C_1$ occurs in the *root-piece* of $C_2$ (abbr.: $C_1 \subseteq_{\mathrm{RP}} C_2$) if $r_{C_1} \in_{\mathrm{RP}} C_2$ and $C_1 \subseteq C_2$.

- $C_1$ is *closed in* $C_2$ (abbr.: $C_1 \subseteq_{\mathrm{CL}} C_2$) if $C_1$ is closed, $r_{C_1} \in_{\mathrm{RP}} C_2$, $e_{C_1} = \left\langle r_{C_1}, x\right\rangle \in \mathrm{E}(T_2)$, $\rho_{C_1} = \ell_2\left(e_{C_1}\right)$ and $(C_1)_x \subseteq C_2$. Relations $\subseteq_{\mathrm{CL}}^{\prec\delta}$ and $\subseteq_{\mathrm{CL}}^{\succeq\delta}$ specify $\subseteq_{\mathrm{CL}}$ by adding conditions $\rho_{C_1} \prec \delta$ and $\rho_{C_1} \succeq \delta$, respectively.
- $T$ and $\langle T, M\rangle$ are called the *skeletons* of $C$. Relations $\subseteq$, $\subseteq_{\mathrm{RP}}$, $\subseteq_{\mathrm{CL}}$ are naturally adapted to skeletons; we also let $r_T := r_C$.

**Definition 13** A set $\mathcal{S} \subseteq \mathcal{R}_\sigma^\nu[\mathcal{O}]$ is $\leq_{\mathrm{C}}$*-wqo* if there is no $\leq_{\mathrm{C}}$-bad sequence $g : \mathbb{N} \to \mathcal{S}$. A sequence $f : \mathbb{N} \to \mathcal{R}_\sigma^\nu[\mathcal{O}]$ is $\leq_{\mathrm{C}}$*-odd* (resp. $\leq_{\mathrm{C}}$*-flawful)* if it is $\leq_{\mathrm{C}}$-bad, although $\mathcal{S}_0$ (resp. $\mathcal{S}_1$) is $\leq_{\mathrm{C}}$-wqo, where $\mathcal{S}_0 = \{C : (\exists n \in \mathbb{N})(C \subsetneq_{\mathrm{CL}} f(n))\}$ and $\mathcal{S}_1 = \{C : (\exists n \in \mathbb{N})(C \subseteq_{\mathrm{CL}} f(n))\}$.

Let $\mathcal{R}_\sigma^\nu[\mathcal{O}]^+ := \mathcal{R}_\sigma^\nu[\mathcal{O}] \cup \{\emptyset\}$. A sequence $f : \mathbb{N} \to \mathcal{R}_\sigma^\nu[\mathcal{O}]^+$ is *normal* if $\{n \in \mathbb{N} : f(n) \neq \emptyset\}$ is infinite. For any normal $f : \mathbb{N} \to \mathcal{R}_\sigma^\nu[\mathcal{O}]^+$ let $\widehat{f} : \mathbb{N} \to \mathcal{R}_\sigma^\nu[\mathcal{O}]$ be an infinite subsequence enumerating the nonempty components; it is defined by $\widehat{f}(n) := f(\varphi(n))$ where $\varphi$ is given by $\varphi(0) := \min\{j \geq 0 : f(j) \neq \emptyset\}$ and $\varphi(i+1) := \min\{j > \varphi(i) : f(j) \neq \emptyset\}$. A normal $f : \mathbb{N} \to \mathcal{R}_\sigma^\nu[\mathcal{O}]^+$ is $\leq_{\mathrm{C}}$*-bad* ($\leq_{\mathrm{C}}$*-odd*, $\leq_{\mathrm{C}}$*-flawful*) if so is the corresponding $\widehat{f} : \mathbb{N} \to \mathcal{R}_\sigma^\nu[\mathcal{O}]$. Note that "$f$ *is* $\leq_{\mathrm{C}}$*-flawful*" implies "$f$ *is* $\leq_{\mathrm{C}}$*-odd*" implies "$f$ *is* $\leq_{\mathrm{C}}$*-bad*".

**Definition 14** Let $f : \mathbb{N} \to \mathcal{R}_\sigma^\nu[\mathcal{O}]^+$ be normal and $\leq_{\mathrm{C}}$-bad. *Minimal-cofinal* sequence $f_{\mathrm{MC}} : \mathbb{N} \to \mathcal{R}_\sigma^\nu[\mathcal{O}]^+$ and correlated strictly increasing function $\varphi : \mathbb{N} \to \mathbb{N}$ are defined by simultaneous recursion, as follows.

1. Let $\varphi(0) := \min\{i : f(i) \neq \emptyset\}$ and for every $j < \varphi(0)$ let $f_{\mathrm{MC}}(j) := \emptyset$.
2. Suppose $\varphi(i)$ and $f_{\mathrm{MC}}(j)$ are defined for all $i \leq k$ and $j < \varphi(k)$, where $f(\varphi(k)) \neq \emptyset$. Then:
   (a) Let $f_{\mathrm{MC}}(\varphi(k))$ be the minimal (lexicographically) $C$ for which there exists a strictly increasing $\psi : \mathbb{N} \to \mathbb{N}$ together with a $\leq_{\mathrm{C}}$-bad sequence $g : \mathbb{N} \to \mathcal{R}_\sigma^\nu[\mathcal{O}]$ such that $\psi(0) = \varphi(k)$, $g(0) = C$ and for all $n \geq 0$ and $j < \psi(0)$ the following two conditions hold.
      i. If $f_{\mathrm{MC}}(j) \neq \emptyset$ then $f_{\mathrm{MC}}(j) \nleq_{\mathrm{C}} g(n)$.
      ii. $g(n) \subseteq_{\mathrm{RP}} f(\psi(n))$.
   (b) Let $\varphi(k+1)$ be the minimum $i > \varphi(k)$ such that $f(i) \neq \emptyset$ and there exist $\psi$ and $g$ as in (a) such that $\psi(0) = i$ and for all $n \geq 0$ and $j \leq \varphi(k)$ the conditions (a) i. and (a) ii. hold.
   (c) For every $m : \varphi(k) < m < \varphi(k+1)$ let $f_{\mathrm{MC}}(m) := \emptyset$.

Now let $\sigma \succeq \delta \in P$. *Minimal-limit-above-*$\delta$ sequence $f_{\mathrm{ML}}^\delta : \mathbb{N} \to \mathcal{R}_\sigma^\nu[\mathcal{O}]^+$ and correlated strictly increasing function $\varphi^\delta : \mathbb{N} \to \mathbb{N}$ are defined analogously, while replacing everywhere $f_{\mathrm{MC}}$ and $\varphi$ by $f_{\mathrm{ML}}^\delta$ and $\varphi^\delta$, respectively, and 2 (a) ii. by

*2 (a) iii.* $g(n) = f(\psi(n))$ or $g(n) \subseteq_{\mathrm{CL}}^{\succeq\delta} f(\psi(n))$.

## 3.2 Basic Observations

**Lemma 15** *For any* $f : \mathbb{N} \to \mathcal{R}_{\sigma}^{\nu}[\mathcal{O}]$, *if* $\{C : (\exists n \in \mathbb{N}) C \subsetneq_{\mathrm{CL}} f(n)\}$ *is* $\leq_{\mathrm{C}}$-wqo *then so is* $\{f(n) : f(n)$ *is neither open nor closed*$\}$, *and if* $\{C : (\exists n \in \mathbb{N}) C \subseteq_{\mathrm{CL}} f(n)\}$ *is* $\leq_{\mathrm{C}}$-wqo *then so is* $\{f(n) : f(n)$ *is nopen*$\}$.

*Proof* This is an easy consequence of the Higman theorem, if we regard nopen *celt*'s as tuples of their closed components. ■

**Lemma 16** *Let* $\delta \preceq \sigma \preceq \nu$ *and* $f, f_{\mathrm{MC}}, f_{\mathrm{ML}}^{\delta} : \mathbb{N} \to \mathcal{R}_{\sigma}^{\nu}[\mathcal{O}]^{+}$ *be as in Definition 14. Then* $f_{\mathrm{MC}}$ *and* $f_{\mathrm{ML}}^{\delta} : \mathbb{N} \to \mathcal{R}_{\sigma}^{\nu}[\mathcal{O}]^{+}$*are both normal and* $\leq_{\mathrm{C}}$*-bad. Moreover, if* $f$ *is* $\leq_{\mathrm{C}}$*-odd then so are* $f_{\mathrm{MC}}$ *and* $f_{\mathrm{ML}}^{\delta}$*, and if* $f$ *is* $\leq_{\mathrm{C}}$*-flawful then so is* $f_{\mathrm{MC}}$*. Furthermore, the following three conditions hold.*

*1. There is a strictly increasing* $\varphi : \mathbb{N} \to \mathbb{N}$ *such that for every* $n \in \mathbb{N}$, $\emptyset \neq f_{\mathrm{MC}}(\varphi(n)) \subseteq_{\mathrm{RP}} f(\varphi(n))$*. Moreover there is no* $\leq_{\mathrm{C}}$*-bad* $g : \mathbb{N} \to \mathcal{R}_{\sigma}^{\nu}[\mathcal{O}]$ *with a strictly increasing* $\psi : \mathbb{N} \to \mathbb{N}$ *such that for every* $n \in \mathbb{N}$, $g(n) \subsetneq_{\mathrm{RP}} f_{\mathrm{MC}}(\psi(n))$*; equivalently,* $\{C \neq \emptyset : (\exists n \in \mathbb{N}) C \subsetneq_{\mathrm{RP}} f_{\mathrm{MC}}(n)\}$ *is* $\leq_{\mathrm{C}}$*-wqo.*

*2. There is a strictly increasing* $\varphi : \mathbb{N} \to \mathbb{N}$ *such that for every* $n \in \mathbb{N}$*, either* $f(\varphi(n)) = f_{\mathrm{ML}}^{\delta}(\varphi(n))$ *or* $\emptyset \neq f_{\mathrm{ML}}^{\delta}(\varphi(n)) \subseteq_{\mathrm{CL}}^{\succeq\delta} f(\varphi(n))$*. Moreover there is no* $\leq_{\mathrm{C}}$*-bad* $g : \mathbb{N} \to \mathcal{R}_{\sigma}^{\nu}[\mathcal{O}]$ *with a strictly increasing* $\psi : \mathbb{N} \to \mathbb{N}$ *such that for every* $n \in \mathbb{N}$, $g(n) \subsetneq_{\mathrm{CL}}^{\succeq\delta} f_{\mathrm{ML}}^{\delta}(\psi(n))$*; equivalently,* $\left\{C \neq \emptyset : (\exists n \in \mathbb{N}) C \subsetneq_{\mathrm{CL}}^{\succeq\delta} f_{\mathrm{ML}}^{\delta}(n)\right\}$ *is* $\leq_{\mathrm{C}}$*-wqo.*

*3. If* $\mathcal{S}_{\prec\delta} := \left\{C \neq \emptyset : (\exists n \in \mathbb{N}) C \subseteq_{\mathrm{CL}}^{\prec\delta} f(n)\right\}$ *is* $\leq_{\mathrm{C}}$*-wqo, then* $Rng\left(f_{\mathrm{ML}}^{\delta}\right)$ *contains only finite collection of* celt*'s that are neither open nor closed above* $\delta$*.*

*Proof* That $f_{\mathrm{MC}}$ and $f_{\mathrm{ML}}^{\delta}$ are normal and $\leq_{\mathrm{C}}$-bad follows by induction from Definition 14. The $\leq_{\mathrm{C}}$-odd ($\leq_{\mathrm{C}}$-flawful) strengthenings in question are obvious. Note that if $f_{\mathrm{ML}}^{\delta}$ is $\leq_{\mathrm{C}}$-odd but not $\leq_{\mathrm{C}}$-flawful, then $f_{\mathrm{ML}}^{\delta}(n) = f(n)$ holds for almost all (nonempty) components of $Rng\left(f_{\mathrm{ML}}^{\delta}\right)$.

1. The former assertion follows by induction from Definition 14. To prove the latter suppose there exists a $\leq_{\mathrm{C}}$-bad $g : \mathbb{N} \to \mathcal{R}_{\sigma}^{\nu}[\mathcal{O}]$ with strictly increasing $\psi : \mathbb{N} \to \mathbb{N}$ such that for every $n$, $g(n)$ is a proper *subrelt* of $f_{\mathrm{MC}}(\psi(n))$ whose root occurs in the root-piece of $f(n)$. But then $g$ should arise at some step $k$ of the construction of $f_{\mathrm{MC}}$ as a possible infinite $\leq_{\mathrm{C}}$-bad extension of the previously defined initial segment $f_{\mathrm{MC}}(0), \cdots, f_{\mathrm{MC}}(\varphi(k) - 1)$. For to verify the only nontrivial condition 2 (a) i. it will suffice to observe that the assumption $\emptyset \neq f_{\mathrm{MC}}(j) \leq_{\mathrm{C}} g(n) \subseteq_{\mathrm{RP}} f_{\mathrm{MC}}(\psi(n))$ would infer $f_{\mathrm{MC}}(j) \leq_{\mathrm{C}} f_{\mathrm{MC}}(\psi(n))$ (in contradiction to proven $\leq_{\mathrm{C}}$-badness of $f_{\mathrm{MC}}$), as $g(n)$ and $f_{\mathrm{MC}}(\psi(n))$ both have the same marks. Thus $g$ would be a legitimate candidate for an infinite $\leq_{\mathrm{C}}$-bad extension in question. But $g(0)$ is smaller than $f_{\mathrm{MC}}(\psi(0))$—a contradiction.

2. The former assertion follows by induction from Definition 14. To prove the latter suppose there exists a $\leq_{\mathrm{C}}$-bad $g : \mathbb{N} \to \mathcal{R}_{\sigma}^{\nu}[\mathcal{O}]$ with strictly increasing $\psi : \mathbb{N} \to \mathbb{N}$ such that for every $n$, $g(n)$ is a proper *subrelt* of $f_{\mathrm{ML}}^{\delta}(\psi(n))$ that is closed above $\delta$. But then arguing as above we show that $g$ should arise at some step $k$ of the construction of $f_{\mathrm{ML}}^{\delta}$ as a possible infinite $\leq_{\mathrm{C}}$-bad extension of the previously

defined initial segment $f^{\delta}_{\text{ML}}(0), \cdots, f^{\delta}_{\text{ML}}(\varphi(k)-1)$, although $g(0)$ is smaller than $f^{\delta}_{\text{ML}}(\psi(0))$—a contradiction.

3. Arguing by contraposition suppose that there exists a strictly increasing $\varphi : \mathbb{N} \to \mathbb{N}$ such that for every $n$, $h(n) := f^{\delta}_{\text{ML}}(\varphi(n)) \in \mathcal{R}^{\nu}_{\sigma}[\mathcal{O}]$ is neither open nor closed above $\delta$. Thus for any $h(n) = \langle T_n, M_n, \ell_n \rangle = \langle S_n, \trianglelefteq_n, M_n, \ell_n \rangle$ and any root-neighbor $y \in S_n$ we have $y \in M_n$. Moreover, since $\mathcal{S}_{\prec\delta}$ is $\leq_{\text{C}}$-wqo, $h : \mathbb{N} \to \mathcal{R}^{\nu}_{\sigma}[\mathcal{O}]$ has no infinite subsequence consisting of closed *relt*'s (otherwise there would be a one consisting of those closed below $\delta$). So we may just as well assume that every $T_n$ has at least two distinct root-neighbors. Let $\mathcal{S}' \subset \mathcal{R}^{\nu}_{\sigma}[\mathcal{O}]$ be the collection of all $C_{n,y} = \langle T_{n,y}, M_{n,y}, \ell_{n,y} \rangle = \langle S_{n,y}, \trianglelefteq_{n,y}, M_{n,y}, \ell_{n,y} \rangle \subsetneq h(n)$, $n \in \mathbb{N}$, where $y \in S_n$ are the root-neighbors in $T_n$ and $S_{n,y} = \{r_{T_n}\} \cup \{x \in S_n : y \trianglelefteq_n x\}$, $\trianglelefteq_{n,y} = \trianglelefteq_n \restriction_{S_{n,y}}$, $M_{n,y} = M_n \cap S_{n,y}$ and $\ell_{n,y} = \ell_n \restriction_{S_{n,y}}$. Since $h$ is $\leq_{\text{C}}$-bad, by an obvious specialization of the Higman's theorem (cf. Lemma 15) we conclude that there exists an infinite $\leq_{\text{C}}$-bad sequence $h' : \mathbb{N} \to \mathcal{S}'$ such that every $h'(n)$ is closed, and hence also closed above $\delta$, since $\mathcal{S}_{\prec\delta}$ is $\leq_{\text{C}}$-wqo. Now that all $h'(n)$ are proper *subrelt*'s of the corresponding components of $h$, this contradicts the minimality of $f^{\delta}_{\text{ML}}$ as in the proof of 1. ■

**Lemma 17** *For any well-order $\mathcal{O} = \langle P, \preceq \rangle$ and $\nu \in P$, there is no normal $\leq_{\text{C}}$-flawful sequence $f : \mathbb{N} \to \mathcal{R}^{\nu}_{\nu}[\mathcal{O}]^{+}$.*

*Proof* Suppose that a normal $f : \mathbb{N} \to \mathcal{R}^{\nu}_{\nu}[\mathcal{O}]^{+}$ is $\leq_{\text{C}}$-flawful. By Definition 11 (1) all labels in $f(n) = \langle T_n, M_n, \ell_n \rangle \neq \emptyset$, $n \in \mathbb{N}$, are $\prec \nu$. By Definition 11 (5) the root-piece of $T_n$ is empty, and hence $\ell_n(y) \in M_n$ holds for every root-neighbor $y$, i.e. every $f(n)$ is nopen. Since $f$ is $\leq_{\text{C}}$-flawful, the set of closed *subcelt*'s occurring in $Rng(f)$ is $\leq_{\text{C}}$-wqo. Hence, by Lemma 15, so is $Rng(f)$ as well—a contradiction to $\leq_{\text{C}}$-badness of $f$. ■

## 4 Proof of Proposition C. Part 2

Denote by $\mathcal{E}^{\nu}[\mathcal{O}]$ ($\mathcal{C}^{\nu}[\mathcal{O}]$) the set of *elt*'s (*celt*'s) relative to $\mathcal{O}_{\prec\nu} := \{\delta \in P : \delta \prec \nu\}$ and let $\mathcal{E}^{\nu}[\mathcal{O}]^{+} := \mathcal{E}^{\nu}[\mathcal{O}] \cup \{\emptyset\}$ and $\mathcal{C}^{\nu}[\mathcal{O}]^{+} := \mathcal{C}^{\nu}[\mathcal{O}] \cup \{\emptyset\}$. Note that $\mathcal{R}^{\nu}_{\sigma}[\mathcal{O}] \subset \mathcal{C}^{\nu}[\mathcal{O}]$ and $\mathcal{E}^{\nu}[\mathcal{O}] \subset \mathcal{R}^{\nu}_{0}[\mathcal{O}]$, and hence every infinite $\leq_{\text{E}}$-bad sequence $g : \mathbb{N} \to \mathcal{E}^{\nu}[\mathcal{O}]$ is normal and $\leq_{\text{C}}$-flawful in $\mathcal{R}^{\nu}_{0}[\mathcal{O}]^{+}$. In the sequel we consider normal sequences $f, g, h$ of the types $\mathbb{N} \to \mathcal{C}^{\nu}[\mathcal{O}]^{+}$ and in particular $\mathbb{N} \to \mathcal{R}^{\nu}_{\sigma}[\mathcal{O}]^{+}$, $\sigma \preceq \nu$, while keeping in mind that $\leq_{\text{C}}$, $\subseteq_{(-)}$ are relations on $\mathcal{C}^{\nu}[\mathcal{O}]$ and $\mathcal{R}^{\nu}_{\sigma}[\mathcal{O}]$; thus in particular $C_1 \not\leq_{\text{C}} C_2$ if $\emptyset \in \{C_1, C_2\}$. Proposition **C** now follows from the alleged existence of a $\leq_{\text{E}}$-bad sequence $g : \mathbb{N} \to \mathcal{E}^{\nu}[\mathcal{O}]$ by Lemma 17 together with following theorem (for $\sigma := 0$ and $\tau := \nu$).

**Theorem 18** *For any well-order $\mathcal{O} = \langle P, \preceq \rangle$ and $\sigma \preceq \tau \preceq \nu \in P$ the following holds. Suppose there exists a normal $\leq_{\text{C}}$-flawful sequence $g : \mathbb{N} \to \mathcal{R}^{\nu}_{\sigma}[\mathcal{O}]^{+}$. Then there exists a normal $\leq_{\text{C}}$-flawful sequence $h : \mathbb{N} \to \mathcal{R}^{\nu}_{\tau}[\mathcal{O}]^{+}$.*

## 4.1 Proof of Theorem 18. Part 1 (Construction)

We define an operator $\mathfrak{R}$ that for any $\sigma \preceq \tau \preceq \nu \in P$ and any normal $\leq_{\mathrm{C}}$-flawful sequence $g : \mathbb{N} \to \mathcal{R}_{\sigma}^{\nu}[\mathcal{O}]^{+}$ will produce a required normal $\leq_{\mathrm{C}}$-flawful sequence $\mathfrak{R}(\sigma, \tau, g) : \mathbb{N} \to \mathcal{R}_{\tau}^{\nu}[\mathcal{O}]^{+}$. Now $\mathfrak{R}(\sigma, \tau, g)$ is defined as follows by transfinite recursion on $\tau$, where we show only nonempty components of $g$ and leave empty ones unchanged, i.e. $\mathfrak{R}(\sigma, \tau, g)(n) := \emptyset$ if $g(n) = \emptyset$. Note that all ordinal labels occurring in $\mathfrak{R}(\sigma, \tau, g)(n)$ already occur in $g$. For the sake of brevity we'll adopt operational formalism to the previously defined minimal cofinal and limit after $\nu$ infinite sequences by setting $\mathfrak{M}_{\mathrm{MC}}(f) := f_{\mathrm{MC}}$ and $\mathfrak{M}_{\mathrm{ML}}^{\nu}(f) := f_{\mathrm{ML}}^{\nu}$.

**Transfinite recursion**.

**Basis**. Let $\mathfrak{R}(\sigma, \sigma, g) := g$.

**Successor clause**. Suppose $\tau = \delta + 1$ in $\mathcal{O}$. Let

$$\boxed{\mathfrak{R}(\sigma, \tau, g) := g_3}$$

where $g_1 - g_3 : \mathbb{N} \to \mathcal{C}^{\nu}[\mathcal{O}]^{+}$ are as follows.

(1) $g_1 := \mathfrak{R}(\sigma, \delta, g)$.

(2) $g_2 := \mathfrak{M}_{\mathrm{MC}}(g_1)$.

(3) For any $n \in \mathbb{N}$ with $g_2(n) = \emptyset$ let $g_3(n) := \emptyset$. For any $n \in \mathbb{N}$ with $\emptyset \neq g_2(n) = \langle T, M, \ell \rangle$, let $g_3(n) := \langle T, M^{+}, \ell \rangle$, where $M^{+} \supseteq M$ arises by adding all lowermost nodes $y \in_{\mathrm{RP}} \langle T, M \rangle$ such that $\ell_n(\mathrm{e}(y)) = \delta$ (if any exists, otherwise $M^{+} := M$).

**Limit clause**. Suppose $\tau = \lim_{i \in \mathbb{N}} \{\tau\}(i)$ in $\mathcal{O}$, $\{\tau\} : \mathbb{N} \to P$ a $\tau$-fundamental sequence with $(\forall i < j)(\{\tau\}(i) \prec \{\tau\}(j) \prec \tau)$ and $(\forall \alpha \prec \tau)\, \exists i\, (\alpha \prec \{\tau\}(i))$. Let $\{\tau\}_{\sigma}$ be a restriction of $\{\tau\}$ that is defined by $\{\tau\}_{\sigma}(0) := \sigma$ and $\{\tau\}_{\sigma}(i+1) := \{\tau\}(i + \min\{j : \sigma \prec \{\tau\}(j)\})$. Now let

$$\boxed{\mathfrak{R}(\sigma, \tau, g) := g_5}$$

where $g_0 : \mathbb{N} \to \left(\mathbb{N} \to \mathcal{C}^{\nu}[\mathcal{O}]^{+}\right)$ and $g_1 - g_5 : \mathbb{N} \to \mathcal{C}^{\nu}[\mathcal{O}]^{+}$ are as follows.

(0) $g_0$ is defined recursively by

$g_0(0) := g$ and $g_0(i+1) := \mathfrak{R}(\{\tau\}_{\sigma}(i), \{\tau\}_{\sigma}(i+1), g_0(i))$.

(1) $g_1$ (being a diagonalization of $g_0$) is defined for any $n \in \mathbb{N}$ by

$g_1(n) := g_0(s(n))(n)$, where $s(n) := \min\{i : (\forall j > i)\, g_0(i)(n) = g_0(j)(n)\}$.[4]

Now consider two cases:

**Case 1**. Suppose that the root-labels of closed *relt*'s $g_1(n), n \in \mathbb{N}$, are all bounded by a fixed $\zeta \prec \tau$. Then for every $i = 2, \cdots, 5$ we let $g_i := g_1$.

**Case 2**. Otherwise, $g_2 - g_5$ are defined as follows.

(2) $g_2 := \mathfrak{M}_{\mathrm{ML}}^{\sigma}(g_1)$.

(3) $g_3$ is defined by recursion. Let $g_3(0) := g_2(0)$ and for any given $k \geq 0$ suppose that $g_3(0), \cdots, g_3(k)$ are already defined, where for any $i \leq k$ with $g_3(i) \neq \emptyset$ we

[4] In the next section we'll show that $s(n) \in \mathbb{N}$ holds for every $n \in \mathbb{N}$.

have $g_3(i) = \langle T_i, M_i, \ell_i \rangle$. Then $g_3(k+1) := g_2(k+1)$, except that $g_2(k+1)$ is closed above $\sigma$ and the following condition $(*)$ :

$$(\forall i \leq k : g_3(i) \neq \emptyset)\,(\forall e \in \mathrm{E}(T_i) : \ell_i(e) \prec \tau) \max\{\ell_i(e), \{\tau\}_\sigma(s(i))\} \preceq \rho_{g_2(k+1)}$$

holds for $s : \mathbb{N} \to \mathbb{N}$ as in (1). In the remaining case of $g_2(k+1) = \langle T, M, \ell \rangle$ that is closed above $\sigma$ and satisfies $(*)$, let $g_3(k+1) := \left\langle T, M^-, \ell' \right\rangle$ for $M^- := M \setminus \{y\}$, where $y$ is the root-neighbor in $T$, while for any $e = \langle u, v \rangle \in \mathrm{E}(T)$ we let

$$\ell'(e) := \begin{cases} \max\left\{\ell\left(\mathrm{e}\left(v'\right)\right) \prec \tau : v' \in_{\mathrm{RP}} \left\langle T, M^- \right\rangle\right\}, & \text{if } \quad v \in_{\mathrm{RP}} \left\langle T, M^- \right\rangle \text{ and } \ell(e) \prec \tau, \\ \ell(e), & \text{else.} \end{cases}$$

(4) $g_4 := \mathfrak{M}_{\mathrm{MC}}(g_3)$.

(5) For any $n \in \mathbb{N}$ with $g_4(n) = \emptyset$ let $g_5(n) := \emptyset$. For any $n \in \mathbb{N}$ with $\emptyset \neq g_4(n) = \langle T, M, \ell \rangle$ let $g_5(n) := \left\langle T, M^+, \ell \right\rangle$, where $M^+ \supseteq M$ arises by adding all lowermost nodes $y \in_{\mathrm{RP}} \langle T, M \rangle$ such that $\ell(\mathrm{e}(y)) \prec \tau$ (if any exists, otherwise $M^+ := M$).

This yields $\mathfrak{R}(\sigma, \tau, g) = g_5$ for limit $\tau$ and thereby completes the definition of $\mathfrak{R}$ by transfinite recursion on $\tau \preceq \nu$.

## 4.2 *Proof of Theorem 18. Part 2 (Soundness)*

To complete the proof we must establish the soundness of the construction, i.e. that at each stage we indeed obtain a normal $\leq_{\mathrm{C}}$-flawful sequence $\mathfrak{R}(\sigma, \tau, g)$ in $\mathcal{R}^\nu_\tau[\mathcal{O}]^+$. Case $\sigma = \tau$ is trivial. Now assuming $\sigma \prec \tau$ the soundness is proved below together with the additional assertion of coherency by simultaneous transfinite induction on $\tau \preceq \nu$.

**Definition 19** (*coherency*) Let $f^\sigma : \mathbb{N} \to \mathcal{R}^\nu_\sigma[\mathcal{O}]^+$ and $f^\tau : \mathbb{N} \to \mathcal{R}^\nu_\tau[\mathcal{O}]^+$ be normal. We say that $f^\sigma$ and $f^\tau$ are *coherent* if for any $m, n \in \mathbb{N}$ the following conditions 1–5 hold, where $f^\xi(k) = \left\langle T^\xi_k, M^\xi_k, \ell^\xi_k \right\rangle = \left\langle S^\xi_k, \trianglelefteq^\xi_k, M^\xi_k, \ell^\xi_k \right\rangle$ are the nonempty components of $f^\xi$, $\xi \in \{\sigma, \tau\}$.

1. If $f^\sigma(n) = \emptyset$ then $f^\tau(n) = \emptyset$, else $T^\tau_n \subseteq_\circ T^\sigma_n$, where $\subseteq_\circ$ is a chain of $\subseteq$ and $\subseteq_{\mathrm{CL}}$. If $f^\sigma(n) \neq \emptyset \neq f^\tau(n)$ then $(\forall e \in \mathrm{E}(f^\tau(n)))\left(\exists e' \in \mathrm{E}(f^\sigma(n))\right) \ell^\tau_n(e) = \ell^\sigma_n\left(e'\right)$.
2. For any $x \in M^\tau_n$ there is $y \trianglelefteq^\sigma_n x$ with $y \in M^\sigma_n$ and $\ell^\sigma_n(\mathrm{e}(y)) \preceq \ell^\tau_n(\mathrm{e}(x))$.
3. Suppose $M^\tau_n \ni x \trianglelefteq^\tau_n y$ and $\ell^\tau_n(\mathrm{e}(x)) \prec \sigma$. Then $\ell^\tau_n(\mathrm{e}(y)) = \ell^\sigma_n(\mathrm{e}(y))$. Moreover $y \in M^\tau_n$ iff $y \in M^\sigma_n$.
4. Suppose $r_{f^\tau(n)} \neq y \in_{\mathrm{RP}} f^\tau(n)$ and $\ell^\tau_n(\mathrm{e}(y)) \succeq \sigma$. Then $\ell^\tau_n(\mathrm{e}(y)) \succeq \tau$ implies $\ell^\tau_n(\mathrm{e}(y)) = \ell^\sigma_n(\mathrm{e}(y))$, whereas $\ell^\tau_n(\mathrm{e}(y)) \prec \tau$ implies $\ell^\sigma_n(\mathrm{e}(y)) \preceq \ell^\tau_n(\mathrm{e}(y))$ and $\left(\exists y' \in_{\mathrm{RP}} f^\sigma(n)\right) \ell^\tau_n(\mathrm{e}(y)) = \ell^\sigma_n\left(\mathrm{e}\left(y'\right)\right)$.
5. If $f^\sigma(n) \neq \emptyset$ and $(\forall m < n)\, f^\tau(m) = f^\sigma(m)$, then $f^\tau(n) \neq \emptyset$.

**Theorem 20** *Let $\sigma \prec \nu$ and a normal $\leq_{\mathrm{C}}$-flawful sequence $f^{\sigma}: \mathbb{N} \to \mathcal{R}_{\sigma}^{\nu}[\mathcal{O}]^{+}$ be fixed. For any $\tau \in [\sigma, \nu]$ let $f^{\tau} := \mathfrak{R}(\sigma, \tau, f^{\sigma})$. Then $f^{\tau}$ is a normal $\leq_{\mathrm{C}}$-flawful sequence of the type $\mathbb{N} \to \mathcal{R}_{\tau}^{\nu}[\mathcal{O}]^{+}$. Moreover $f^{\sigma}$ and $f^{\tau}$ are coherent.*

*Proof* We argue by simultaneous transfinite induction on $\tau$. $Rng(f^{\tau}) \subseteq \mathcal{R}_{\tau}^{\nu}[\mathcal{O}]^{+}$ and coherency conditions 1–4 are readily verified by the definition of $\mathfrak{R}(\sigma, \tau, f^{\sigma})$. Consider the rest.

**Basis**. Case $\tau := \sigma$ is trivial, as $\mathfrak{R}(\sigma, \sigma, f^{\sigma}) = f^{\sigma}$.

**Successor clause**. Suppose $\tau = \delta + 1$ and let $g_1 = f^{\delta}$, $g_2 = \mathfrak{M}_{\mathrm{MC}}(f^{\delta})$ and $g_3 = f^{\delta+1}$ be as in the definition of $\mathfrak{R}(\sigma, \delta+1, f^{\sigma})$. That $g_1$ and $g_2$ are normal $\leq_{\mathrm{C}}$-flawful sequences in $\mathcal{R}_{\delta}^{\nu}[\mathcal{O}]^{+}$ is readily seen by the induction hypothesis and definition of $\mathfrak{M}_{\mathrm{MC}}$ along with Lemma 16 (1), respectively. $g_3: \mathbb{N} \to \mathcal{R}_{\tau}^{\nu}[\mathcal{O}]^{+}$ is normal and $\leq_{\mathrm{C}}$-bad, as adding new marks can only destroy the embeddability $\leq_{\mathrm{C}}$. Consider the $\leq_{\mathrm{C}}$-flawfulness of $g_3$. Since closed *subrelt*'s occurring in $Rng(g_2)$ are the ones occurring in $Rng(g_1) = Rng(f^{\delta})$, by the induction hypothesis the set of these *subrelt*'s is a $\leq_{\mathrm{C}}$-wqo. Hence $g_3$ is $\leq_{\mathrm{C}}$-odd (at least). It remains to show that the set of new closed *subrelt*'s in $Rng(g_3)$ that arise by adding marks whose edges are labeled with $\delta$ is a $\leq_{\mathrm{C}}$-wqo too. Suppose not and consider an infinite $\leq_{\mathrm{C}}$-bad sequence $h: \mathbb{N} \to \mathcal{R}_{\tau}^{\nu}[\mathcal{O}]$ and a strictly increasing $\psi: \mathbb{N} \to \mathbb{N}$ such that for any $n \in \mathbb{N}$, $h(n) \subseteq_{\mathrm{CL}} g_3(\psi(n)) = \langle T_n, M_n, \ell_n \rangle$ with $\rho_{h(n)} = \delta$. Let $h^{-}(n)$ be a *subrelt* of $g_2(\psi(n))$ obtained by deleting the root mark $x$ and $h^{-}: \mathbb{N} \to \mathcal{R}_{\delta}^{\nu}[\mathcal{O}]$ be the corresponding sequence of *subrelt*'s $h^{-}(n) \subsetneq_{\mathrm{RP}} g_2(\psi(n))$. Note that for any $m < n$, $h^{-}(m) \leq_{\mathrm{C}} h^{-}(n)$ implies $h(m) \leq_{\mathrm{C}} h(n)$, as $\delta$ is the minimum ordinal occurring in the root-piece of $g_2(\psi(m))$. Since $Rng(h^{-})$ is $\leq_{\mathrm{C}}$-wqo (cf. Lemma 16 (1)), we arrive at a contradiction to $\leq_{\mathrm{C}}$-badness of $h$. Thus $g_3 = f^{\tau}$ is $\leq_{\mathrm{C}}$-flawful in $\mathcal{R}_{\tau}^{\nu}[\mathcal{O}]^{+}$, as required.

Consider coherency condition 5. Suppose $f^{\sigma}(n) \neq \emptyset$ and for every $m < n$, $f^{\sigma}(m) = f^{\delta+1}(m) = g_3(m) \neq \emptyset$. By the induction hypothesis with respect to conditions 1–4 together with obvious skeleton-monotonicity of $g_1 - g_3$ we obtain $f^{\sigma}(m) = f^{\delta}(m) = g_1(m) = g_2(m)$, which by the induction hypothesis with respect to condition 5 yields $f^{\delta}(n) = g_1(n) \neq \emptyset$. Furthermore, $g_2(n) \neq \emptyset$ follows from $g_1(n) \neq \emptyset$ and $(\forall m < n)\, g_1(m) = g_2(m)$ by the definition of $g_2 = \mathfrak{M}_{\mathrm{MC}}(g_1)$, while $f^{\delta+1}(n) = g_3(n) \neq \emptyset$ easily follows from $g_2(n) \neq \emptyset$ by the definition of $g_2$.

**Limit clause**. Suppose $\tau = \lim_{i \in \mathbb{N}} \{\tau\}(i)$ and consider sequences $g_0 - g_5$ as in the definition of $\mathfrak{R}(\sigma, \tau, f^{\sigma})$. First we prove the soundness of $g_1 - g_5$.

$(g_1)$ $s(n) = \min\{i : (\forall j > i)\, g_0(i)(n) = g_0(j)(n)\}$ is well-defined for every $n \in \mathbb{N}$, where $g_0: \mathbb{N} \to (\mathbb{N} \to \mathcal{C}^{\nu}[\mathcal{O}]^{+})$ arises by recursive clauses $g_0(0) := f^{\sigma}$ and $g_0(i+1) := \mathfrak{R}(\{\tau\}_{\sigma}(i), \{\tau\}_{\sigma}(i+1), g_0(i))$. Indeed, by the induction hypothesis, every $g_0(i)$ is normal in $\mathcal{R}_{\{\tau\}_{\sigma}(i)}^{\nu}[\mathcal{O}]^{+}$. Consider an infinite sequence $\{g_0(i)(n)\}_{i \in \mathbb{N}}$. By the induction hypothesis with respect to conditions 1–4, the skeletons of $g_0(i)(n)$ weakly decrease while their labels and marks weakly increase with regard to $i$ such that every new ordinal already occurs in $f^{\sigma}(n)$. Hence $g_0(0)(n), \cdots, g_0(i)(n), \cdots$ must stabilize after a certain $i =: s(n)$, i.e. $(\forall j \geq s(n))\, g_0(j)(n) = g_0(s(n))(n)$. Hence $s(n)$ and $g_1(n) := g_0(s(n))(n) \in \mathcal{R}_{\tau}^{\nu}[\mathcal{O}]^{+}$ are well-defined. We claim that $g_1$ is normal, i.e. it contains an infinite set of nonempty components $g_1(m) =$

$\langle T_m, M_m, \ell_m \rangle$, $m \in \mathbb{N}$. Suppose not and let

$$n_0 := \min \{n : (\forall m > n)\, g_1(m) = \emptyset\},\ i_0 := \max \{s(0), \cdots, s(n_0)\},$$

$$n_1 := \min \{m > n_0 : (\exists i > i_0)\, g_0(i)(m) \neq \emptyset\},\ i_1 := \min \{i > i_0 : g_0(i)(n_1) \neq \emptyset\}.$$

(The existence of $n_1$ and $i_1$ follows from the normality of $g_0(i)$'s.) Thus for any $j \geq i_1, m \leq n_0$ implies $g_0(j)(m) = g_0(i_0)(m) = g_1(m)$, whereas $n_0 < m < n_1$ implies $g_1(m) = \emptyset$ together with $g_0(j)(m) = \emptyset$. Hence for all $j, j' \geq i_1$ and $m < n_1$ we have $g_0(j)(m) = g_0(j')(m) = g_1(m)$ and in particular $g_0(j)(m) = g_0(i_1)(m)$. Having this, by the induction hypothesis with respect to coherency condition 5, we can pass from $g_0(i_1)(n_1) \neq \emptyset$ to $g_0(j)(n_1) \neq \emptyset$, for all $j$ from $i_1 + 1$ to $s(n_1)$, and eventually arrive at $g_0(s(n_1))(n_1) = g_1(n_1) \neq \emptyset$–a contradiction, as $n_1 > n_0$. So $g_1 : \mathbb{N} \to \mathcal{R}^{\nu}_{\tau}[\mathcal{O}]^+$ is normal and $\leq_{\mathrm{C}}$-bad, as for any $m < n$, $g_1(m) = g_0(s(m))(m) \leq_{\mathrm{C}} g_1(n) = g_0(s(n))(n)$ would infer $g_0(i)(m) \leq_{\mathrm{C}} g_0(i)(n)$ for $i := \max\{s(m), s(n)\}$, in contradiction to $\leq_{\mathrm{C}}$-badness of $g_0(i)$.

We can't guarantee that $g_1$ is $\leq_{\mathrm{C}}$-flawful. However, we can prove that for every fixed $\zeta \prec \tau$, the set $\mathcal{S}^{\zeta}_1 = \left\{C : (\exists n \in \mathbb{N})\left(C \subseteq^{\prec\zeta}_{\mathrm{CL}} g_1(n)\right)\right\}$ is a $\leq_{\mathrm{C}}$-wqo. Since by the induction hypothesis every single $g_0(i)$ is $\leq_{\mathrm{C}}$-flawful, it will suffice to show that $C \subseteq^{\prec\zeta}_{\mathrm{CL}} g_1(n)$ implies $C \subseteq_{\mathrm{CL}} g_0(i_\zeta)(n)$, and hence $C \subseteq^{\prec\zeta}_{\mathrm{CL}} g_0(i_\zeta)(n)$, where $i_\zeta = \min\{i \leq s(n) : \zeta \prec \{\tau\}_\sigma(i)\}$. Now suppose $C = \langle T, M, \ell \rangle \subseteq_{\mathrm{CL}} g_1(n) = g_0(s(n))(n)$, while $\sigma \preceq \rho_C = \ell(e_C) \prec \zeta$ and $e_C = \langle r_C, x \rangle$, $x \in M$. For any $j \leq s(n)$ let $g_0(j)(n) = \langle T_j, M_j, \ell_j \rangle \in \mathcal{R}^{\nu}_{\{\tau\}_\sigma(j)}[\mathcal{O}]$. By the coherency conditions 1–4, there exists $j_0 < s(n)$ such that $x \in M_{j_0+1} \setminus M_{j_0}$ is a lowermost mark in $g_0(j_0+1)(n)$, and hence $e_C \in_{\mathrm{RP}} g_0(j_0)(n)$. This yields $\{\tau\}_\sigma(j_0) \preceq \rho_C \prec \zeta$ (cf. Definition 11 (5)). Hence $j_0 + 1 \leq i_\zeta \leq s(n)$, as $\zeta \prec \{\tau\}_\sigma(i_\zeta)$. Moreover $C \subseteq_{\mathrm{CL}} g_0(j_0+1)(n)$, while $x \in M_j$ if $j_0 < j \leq s(n)$, which by (iterated) coherency condition 2 yields $C \subseteq_{\mathrm{CL}} g_0(i_\zeta)(n)$, as desired. Thus $\mathcal{S}^{\zeta}_1$ is a $\leq_{\mathrm{C}}$-wqo for every fixed $\zeta \prec \tau$, and hence $g_5 := g_1$ is $\leq_{\mathrm{C}}$-flawful in the Case 1 of the definition of $\mathfrak{R}(\sigma, \tau, f^\sigma)$. Otherwise, in the corresponding Case 2, consider a more sophisticated sequences $g_2 - g_5$.

($g_2$) That $g_2 : \mathbb{N} \to \mathcal{R}^{\nu}_{\tau}[\mathcal{O}]^+$ is normal and $\leq_{\mathrm{C}}$-bad follows by Lemma 16 (2) from already proven normality and $\leq_{\mathrm{C}}$-badness of $g_1$. Actually it follows that $g_2$ is $\leq_{\mathrm{C}}$-odd (although not necessarily $\leq_{\mathrm{C}}$-flawful, see above).

($g_3$) Hence $g_3$ is also normal and $\{C : (\exists n \in \mathbb{N})(C \subsetneq_{\mathrm{CL}} g_3(n))\}$ is a $\leq_{\mathrm{C}}$-wqo. Moreover, the $\leq_{\mathrm{C}}$-badness of $g_2$ implies that of $g_3$. Indeed, in the only nontrivial $g_3$-recursion step case, $m < n$ with $g_3(m) \leq_{\mathrm{C}} g_3(n)$ would imply $g_2(m) \leq_{\mathrm{C}} g_2(n)$, since $\rho_{g_2(n)}$ is not smaller than any ordinal label $\prec \tau$ occurring in $g_3(m)$, due to $g_2(n)$'s condition $(*)$. [However, we can't guarantee that ordinal labels occurring in the root-pieces of $g_3(n)$ are $\succeq \tau$; thus $Rng(g_3) \subseteq \mathcal{R}^{\nu}_{\tau}[\mathcal{O}]^+$ can fail (cf. Definition 11 (5)).]

($g_4$) That $g_4$ is normal and $\leq_{\mathrm{C}}$-odd now easily follows by Lemma 16 (1).

($g_5$) Since adding new marks in $g_4(n)$, $n \in \mathbb{N}$, can only destroy $\leq_{\mathrm{C}}$, $g_5$ is still normal and $\leq_{\mathrm{C}}$-bad. Moreover $Rng(g_5) \subseteq \mathcal{R}^{\nu}_{\tau}[\mathcal{O}]^+$. To prove that $g_5$ is $\leq_{\mathrm{C}}$-flawful it will suffice to show that the set of corresponding new lowermost closed (sub)*relt*'s is a $\leq_{C}$-wqo. Suppose not and consider a $\leq_{C}$-bad sequence $h : \mathbb{N} \to \mathcal{R}_{\tau}^{\nu}[\mathcal{O}]$ together with correlated strictly increasing sequence $\psi : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$, $h(n) \subseteq_{CL} g_5(\psi(n))$ and $\rho_{h(n)} \prec \tau$. Moreover, since $\mathcal{O}$ is well-ordered, we can just as well assume that $\rho_{h(m)} \preceq \rho_{h(n)}$ holds for all $m < n$. Now let $h^{-} : \mathbb{N} \to \mathcal{R}_{\tau}^{\nu}[\mathcal{O}]$ be a sequence of *subrelt*'s $h^{-}(n) \subsetneq_{RP} g_4(\psi(n))$ obtained by deleting the root marks of $h(n)$. It follows that for any $m < n$, $h^{-}(m) \leq_{C} h^{-}(n)$ implies $h(m) \leq_{C} h(n)$. Since $Rng\left(h^{-}\right)$ is a $\leq_{C}$-wqo (cf. Lemma 16 (1)), we arrive at a contradiction to $\leq_{C}$-badness of $h$. Hence $Rng(h)$ is $\leq_{C}$-excellent and $g_5 = f_{\sigma}^{\tau}$ is $\leq_{C}$-flawful in $\mathcal{R}_{\tau}^{\nu}[\mathcal{O}]^{+}$, as desired.

Consider coherency condition 5. Suppose $f^{\sigma}(n) \neq \emptyset$ and for all $m < n$, $f^{\sigma}(m) = f^{\tau}(m) \neq \emptyset$. By the induction hypothesis with respect to coherency conditions 1–4 together with obvious skeleton-monotonicity of $g_1 - g_5$ we get $f^{\sigma}(m) = f^{\tau}(m) = g_0(s(m))(m) = g_1(m)$ and $f^{\sigma}(m) = f^{\tau}(m) = g_0(s(m))(m) = g_1(m) = \cdots = g_5(m)$ in the corresponding Case 1 and Case 2, respectively. In both cases, the induction hypothesis with respect to condition 5 yields $g_1(n) \neq \emptyset$ via $g_1(i) = g_0(s(i))(i) = g_0(i_n)(i)$ for all $i \leq n$, where $i_n := \max\{s(i) : i \leq n\}$. In the Case 1 this already yields $f^{\tau}(n) = g_5(n) \neq \emptyset$, as $g_1 = g_5$. Furthermore, in the Case 2, $g_2(n) \neq \emptyset$ follows from $g_1(n) \neq \emptyset$ and $(\forall m < n)\, g_1(m) = g_2(m)$ by the definition of $g_2 = \mathfrak{M}_{\mathrm{ML}}^{\sigma}(g_1) = (g_1)_{\mathrm{ML}}^{\sigma}$, while $g_3(n) \neq \emptyset$ easily follows from $g_2(n) \neq \emptyset$ by the definition of $g_3$. By the same token, we pass from $g_3(n) \neq \emptyset$ and $(\forall m < n)\, g_3(m) = g_4(m)$ to $g_4(n) \neq \emptyset$ by the definition of $g_4 = \mathfrak{M}_{\mathrm{MC}}(g_3) = (g_3)_{\mathrm{MC}}$, and finally from $g_4(n) \neq \emptyset$ to $f^{\tau}(n) = g_5(n) \neq \emptyset$ by the definition of $g_5$. This completes the condition 5 and thereby the inductive proof of Theorem 18 along with the whole proof of Proposition **C** (see above). ■

In the next section we show that this proof can be formalized in $\mathbf{ITR}_0$.

## 4.3 Formalization

Recall that the Higman's theorem is provable in $\mathbf{ACA}_0$ (cf. e.g. [5, 15]). Furthermore, arguing in $\Pi_1^1$-$\mathbf{CA}_0$ (=$\mathbf{ACA}_0$ plus $\Pi_1^1$-comprehension axiom) we'll assume that any well-order $\mathcal{O} = \langle P, \preceq \rangle$ under consideration permits standard operations of ordinal sum $\alpha + \beta$ and ordinal exponentiation $\omega^{\alpha}$, for all $\alpha \in P$. Consider operators $\mathfrak{M}_{\mathrm{MC}}$ and $\mathfrak{M}_{\mathrm{ML}}^{\sigma}$, $\sigma \in P$. It is readily seen that sequences $f_{\mathrm{MC}} = \mathfrak{M}_{\mathrm{MC}}(f)$ and $f_{\mathrm{ML}}^{\sigma} = \mathfrak{M}_{\mathrm{ML}}^{\sigma}(f)$, regarded as second-order objects, are definable by $\Sigma_1^1$-formulas with second-order parameters $f$ and $\mathcal{O}$. Consequently, $\mathfrak{M}_{\mathrm{MC}}(f)$ and $\mathfrak{M}_{\mathrm{ML}}^{\sigma}(f)$ are both recursive in $U_1(f, \mathcal{O}) := \{x : U_1(x, f, \mathcal{O})\}$ for a fixed universal $\Pi_1^1$-predicate $U_1(x, f, \mathcal{O})$. Now consider the first part of the proof of Theorem 18, i.e. the construction of $\mathfrak{R}(\sigma, \tau, g)$ by transfinite recursion on $\tau$.

**Lemma 21** *For every $\sigma \preceq \tau \preceq \nu$, $\mathfrak{R}(\sigma, \tau, g)$ is recursive in* $\left(\omega^{\omega^{\tau}} + 3\right)$th *iteration of $\lambda f.U_1(f, \mathcal{O})$ starting with $f := g$ (abbr.: $\lambda f_{[g]}.U_1(f, \mathcal{O})$).*

*Proof* The initial case $\tau = \sigma$ is trivial as $\mathfrak{R}(\sigma, \sigma, g) = g$.

Let $\tau = \delta + 1$. By the definition, $\mathfrak{R}(\sigma, \tau, g)$ is recursive in $\mathfrak{M}_{\mathrm{MC}}(\mathfrak{R}(\sigma, \delta, g))$, while by the induction hypothesis $\mathfrak{R}(\sigma, \delta, g)$ is recursive in $\left(\omega^{\omega^{\delta}} + 3\right)th$ iteration of $\lambda f_{[g]}.U_1(f, \mathcal{O})$. Hence by the above estimate of $\mathfrak{M}_{\mathrm{MC}}(f)$, $\mathfrak{R}(\sigma, \tau, g)$ is recursive in $\left(\omega^{\omega^{\delta}+\omega} + 3 \prec \omega^{\omega^{\delta+1}} \prec \omega^{\omega^{\tau}} + 3\right)th$ iteration of $\lambda f_{[g]}.U_1(f, \mathcal{O})$, as required.

Let $\tau = \lim_{i \in \mathbb{N}} \{\tau\}(i)$. We note that $g_1$ is first-order definable over the collection $\{g_0(i)\}_{i \in \mathbb{N}}$, where $g_0(0) = g$ and $g_0(i+1) = \mathfrak{R}(\{\tau\}_\sigma(i), \{\tau\}_\sigma(i+1), g_0(i))$. By the induction hypothesis, every $g_0(i+1)$ is recursive in $\left(\omega^{\omega^{\{\tau\}_\sigma(i+1)}} + 3\right)th$ iteration of $\lambda f_{[g_0(i)]}.U_1(f, \mathcal{O})$. Having this we conclude that $g_0(i+1)$ is recursive in $\left(\prod_{j=0}^{i+1}\left(\omega^{\omega^{\{\tau\}_\sigma(j+1)}} + 3\right) \prec \prod_{j=0}^{i+1}\left(\omega^{\omega^{\{\tau\}_\sigma(j+1)}+1}\right) \prec \omega^{\omega^{\{\tau\}_\sigma(i+1)+1}} \prec \omega^{\omega^{\tau}}\right)th$ iteration of $\lambda f_{[g]}.U_1(f, \mathcal{O})$. Consequently, $g_1$ is recursive in $\left(\omega^{\omega^{\tau}} + 1 \prec \omega^{\omega^{\tau}} + 3\right)th$ iteration of $\lambda f_{[g]}.U_1(f, \mathcal{O})$, which yields the result in the Case 1. Furthermore, in the Case 2, $g_2$ and $g_3$ are recursive in $\left(\omega^{\omega^{\tau}} + 2\right)th$ iteration of $\lambda f_{[g]}.U_1(f, \mathcal{O})$ according to the above estimate of $\mathfrak{M}^{\sigma}_{\mathrm{ML}}(f)$, and hence $g_4$ and $g_5 = \mathfrak{R}(\sigma, \tau, g)$ are recursive in $\left(\omega^{\omega^{\tau}} + 3\right)th$ iteration of $\lambda f_{[g]}.U_1(f, \mathcal{O})$ according to our estimate of $\mathfrak{M}_{\mathrm{MC}}(f)$. This completes the proof. ■

In the second part of the proof, i.e. in the proof of Theorem 20, we argue in $\Pi^1_1$-$\mathbf{CA}_0$ extended by $\Pi^1_1$-transfinite induction along $\mathcal{O}$ up to $\omega^{\omega^{\nu}} + 3$ (cf. previous lemma). Since $\Pi^1_1$-comprehension is obviously included in $\mathbf{ITR}_0$, this shows that both the transfinite induction and all nested numerical inductions involved can be replaced by the corresponding restricted variants whose all second-order parameters are definable by $\left(\omega^{\omega^{\nu}} + 3\right)th$ iteration of $\lambda f_{[g]}.U_1(f, \mathcal{O})$. These restricted inductions are obviously derivable in $\mathbf{ITR}_0$. Hence the whole proof is also derivable in $\mathbf{ITR}_0$. This shows that $\mathbf{ITR}_0$ proves both Theorem 18 and Proposition **C**. Lemma 8 completes the proof of Main Theorem.

## 5 Proof of Proposition D

**Definition 22** Let $\mathcal{Q} = \langle Q, \preceq_{\mathrm{Q}} \rangle$, $Q \subseteq \mathbb{N}$, be any *wqo* and $\mathcal{O} = \langle P, \preceq \rangle$ a well-order, as above. *Generalized elt* (abbr.: *gelt*) and *generalized celt* (abbr.: *gcelt*) relative to $\langle \mathcal{O}, \mathcal{Q} \rangle$ are structures $\langle T, \ell, \varrho \rangle = \langle S, \trianglelefteq, \ell, \varrho \rangle$ and $\langle T, M, \ell, \varrho \rangle = \langle S, M, \trianglelefteq, \ell, \varrho \rangle$ that extend given *elt* $E = \langle T, \ell \rangle$ and *celt* $C = \langle T, M, \ell \rangle$, respectively, by a new *wqo-labeling* function $\varrho : \mathrm{E}_0(T) \to Q$, where $\mathrm{E}_0(T) = \{\mathrm{e}(x) : x \in \mathrm{L}(T)\}$ is the set of end-edges in $T$. The generalized relations $f : \langle T_1, \ell_1, \varrho_1 \rangle \leq_{\mathrm{GE}} \langle T_2, \ell_2, \varrho_2 \rangle$ and $f : \langle T_1, M_1, \ell_1, \varrho_1 \rangle \leq_{\mathrm{GC}} \langle T_2, M_2, \ell_2, \varrho_2 \rangle$ on *gelt*'s and *gcelt*'s, respectively, arise by extending $f : \langle T_1, \ell_1 \rangle \leq_{\mathrm{E}} \langle T_2, \ell_2 \rangle$ and $f : \langle T_1, M_1, \ell_1 \rangle \leq_{\mathrm{C}} \langle T_2, M_2, \ell_2 \rangle$ by a new condition

- If $x \in \mathrm{L}(T_1)$ then $\varrho_1(\mathrm{e}(x)) \preceq_{\mathrm{Q}} \varrho_2(\mathrm{e}(f(x)))$.

Let **E** and **F** abbreviate the generalized propositions about *gelt*'s and *gcelt*'s that are obtained by replacing in **C** and **D**, respectively, $\leq_{\mathrm{E}}$ by $\leq_{\mathrm{GE}}$ and $\leq_{\mathrm{C}}$ by $\leq_{\mathrm{GC}}$. Obviously **E** implies **C** and **F** implies **D**, provably in $\mathbf{ACA}_0$.

**Theorem 23** *$\boldsymbol{ITR}_0$ proves **E**. Hence Proposition **E** has the same proof-theoretic strength as each of Propositions **A**, **B** and **C** (namely, that of $\boldsymbol{ITR}_0$).*

*Proof* By straightforward modification of the proof of Main Theorem. ■

Below, additional superscripts "G" refer to the $\leq_{\mathrm{GC}}$ -counterparts of previous Notations 12, Definitions 13, 14 and Lemmas 15, 16 (with respect to $\mathcal{C}^{\nu}[\mathcal{O}]^{+}$, instead of $\mathcal{R}^{\nu}_{\sigma}[\mathcal{O}]^{+}$). Thus in particular Lemma $16^{\mathrm{G}}$ deals with minimal $\leq_{\mathrm{GC}}$-bad sequences that arise by applying Definitions $13^{\mathrm{G}}$, $14^{\mathrm{G}}$ which, in turn, are obtained by substituting *gcelt*'s for *relt*'s and $\leq_{\mathrm{GC}}$ for $\leq_{\mathrm{C}}$ everywhere in Definitions 13, 14. It is readily seen that Lemmas $15^{\mathrm{G}}$ and $16^{\mathrm{G}}$ are provable analogously to Lemmas 15 and 16, respectively.

**Theorem 24** *$\boldsymbol{ITR}_0$ proves **F**. Hence Proposition **F** has the same proof-theoretic strength as each of Propositions **A**, **B**, **C**, **D** and **F**.*

*Proof* Denote by $\mathcal{G}[\mathcal{O},\mathcal{Q}]$ the set of *gcelt*'s relative to $\langle\mathcal{O},\mathcal{Q}\rangle$ and let $\mathcal{G}[\mathcal{O},\mathcal{Q}]^{+} := \mathcal{G}[\mathcal{O},\mathcal{Q}]\cup\{\emptyset\}$. Suppose $f:\mathbb{N}\to\mathcal{G}[\mathcal{O},\mathcal{Q}]^{+}$ is normal and $\leq_{\mathrm{GC}}$-bad. Let $f_0 := f^{0}_{\mathrm{ML}}:\mathbb{N}\to\mathcal{G}[\mathcal{O},\mathcal{Q}]^{+}$. By Lemma $16^{\mathrm{G}}$, $f_0$ is normal and $\leq_{\mathrm{GC}}$-bad, whereas $\mathcal{S}_0 = \{C : (\exists n\in\mathbb{N})C \subsetneq_{\mathrm{CL}} f_0(n)\}$ is $\leq_{\mathrm{GC}}$-wqo. Moreover by Lemma $15^{\mathrm{G}}$, every infinite sequence of nopen *gcelt*'s $f_0(n)$ must include an infinite subsequence consisting of closed *gcelt*'s $f_0(n')$. A desired contradiction now follows by cases.

**Case 1**. Assume that $f_0$ includes an infinite subsequence $f_1$ consisting of open *gcelt*'s $f_1(n) = f_0(\psi(n))$ for a strictly increasing $\psi:\mathbb{N}\to\mathbb{N}$. Consider a wqo $\mathcal{Q}_0 = \langle\mathcal{S}_0,\leq_{\mathrm{GC}}\rangle$. Furthermore, with every *gcelt* $f_1(n)$ we associate a *gelt* $g(n)$ relative to $\langle\mathcal{O},\mathcal{Q}_0\rangle$ that is obtained by replacing, in $f_1(n)$, every $C \subsetneq_{\mathrm{CL}} f_1(n)$ by its root-edge $e$ supplied with the wqo-label $\varrho(e) := C\in\mathcal{S}_0$ along with the old ordinal-label $\ell(e)\in P$; thus $e$ is an end-edge in $g(n)$. Note that no mark occurs in $g(n)$ anymore. It is readily seen that for any $i<j\in\mathbb{N}$, $g(i)\leq_{\mathrm{GE}} g(j)$ implies $f_1(i)\leq_{\mathrm{GC}} f_1(j)$ relative to $\langle\mathcal{O},\mathcal{Q}_0\rangle$ and $\langle\mathcal{O},\mathcal{Q}\rangle$, respectively. Now by Theorem 23, there exist $i<j\in\mathbb{N}$ with $g(i)\leq_{\mathrm{GE}} g(j)$, and hence $f_1$ is not $\leq_{\mathrm{GC}}$-bad—a contradiction.

**Case 2**. Assume that the assumption of Case 1 fails. Hence by Lemma $15^{\mathrm{G}}$, $f_0$ includes an infinite subsequence $f_1$ consisting of closed *gcelt*'s $f_1(n) = f_0(\psi(n))$ for a strictly increasing $\psi:\mathbb{N}\to\mathbb{N}$. Moreover, since $\mathcal{O}$ is a well-order, we may just as well assume that the root-ordinals in $f_1(n)$, $n\in\mathbb{N}$, are weakly increasing. Now for any $n\in\mathbb{N}$ let $f_2(n)$ arise by deleting the lowermost mark (that is assigned to the root-neighbor) of $f_1(n)$; thus $f_2(n)$ is open. By the monotonicity of the root-ordinals, for any $i<j\in\mathbb{N}$, $f_2(i)\leq_{\mathrm{GC}} f_2(j)$ implies $f_1(i)\leq_{\mathrm{GC}} f_1(j)$ relative to $\langle\mathcal{O},\mathcal{Q}\rangle$, respectively. Hence $f_2$ is $\leq_{\mathrm{GC}}$-bad, as so is $f_1$. The rest of the proof follows by Case 1 with respect to $f_2$, instead of $f_1$. ■

*Remark 25* 1. One can further generalize *elt*'s by also applying the *wqo*-labeling functions $\varrho$ on the whole vertex domain $\mathrm{V}(T)$, while assuming that the corresponding generalized homeomorphisms $f : \langle T_1, M_1, \ell_1, \varrho_1 \rangle \leq_{\mathrm{GC}} \langle T_2, M_2, \ell_2, \varrho_2 \rangle$ also preserve $\preceq_{\mathrm{Q}}$, i.e. $(\forall x \in \mathrm{V}(T_1))\, \varrho_1(x) \preceq_{\mathrm{Q}} \varrho_2(f(x))$. Denote by **G** the resulting strengthening of Proposition **F**. Then by a slight modification of previous arguments we conclude that $\mathbf{ITR}_0$ also proves Proposition **G**. Thus, in the sense of ordinal provability, **G** is not stronger that **F**, and hence not stronger than **A**.

2. Recall that so far we considered vertex-labeled homeomorphic embeddability with the symmetric gap-condition

$(s)$ If $f(\mathrm{p}_1(y)) \lhd_2 u \lhd_2 f(y)$ then $\min\{\ell_1(\mathrm{p}_1(y)), \ell_1(y)\} \preceq \ell_2(u)$

and the asymmetric gap-condition

$(a1)$ If $f(\mathrm{p}_1(y)) \lhd_2 u \lhd_2 f(y)$ then $\ell_1(y) \preceq \ell_2(u)$.

Now consider another asymmetric gap-condition

$(a2)$ If $f(\mathrm{p}_1(y)) \lhd_2 u \lhd_2 f(y)$ then $\ell_1(\mathrm{p}_1(y)) \preceq \ell_2(u)$.

Obviously both $(a1)$ and $(a2)$ imply $(s)$. Recall that Propositions **A** and **B** deal with *vlt*-embeddability with gap-conditions $(s)$ and $(a1)$, respectively, which by [6] and Main Theorem are both provable in $\mathbf{ITR}_0$. Moreover, a modified Proposition **B** dealing gap-condition $(a2)$, instead of $(a1)$, is still provable in $\mathbf{ITR}_0$. This follows from Theorem 23 with $\mathcal{Q} := \mathcal{O}$ and the corresponding modification of Lemma 8 where $\ell_n^2(\mathrm{e}(y))$ is $\ell_n^1(\mathrm{p}(y))$, instead of $\ell_n^1(y)$. Summing up, all sound variants of ordinal (vertex- and/or edge-) labeled generalizations of Friedman's tree theorem are equivalent to Proposition **A** modulo ordinal provability over $\mathbf{ACA}_0$, while having proof-theoretic strength of $\mathbf{ITR}_0$ (cf. [6, 7]).

However, the following "unsound" vertex-labeled symmetric gap-condition

$(s)^\dagger$ If $f(\mathrm{p}_1(y)) \lhd_2 u \lhd_2 f(y)$ then $\max\{\ell_1(\mathrm{p}_1(y)), \ell_1(y)\} \preceq \ell_2(u)$

fails to provide a *wqo*. An easy counter-example is as follows. For any $n \in \mathbb{N}$, let $\mathcal{I}(n) = [0 \lhd 1 \lhd \cdots \lhd n+2]$ be an interval (i.e. a non-branching tree) whose vertices are labeled by $\ell(0) = \ell(n+2) := 2$, $\ell(2i) := 0$ and $\ell(2i+1) := 1$, for all $i < \left\lceil \frac{1}{2} n \right\rceil + 1$ and $j \leq \left\lceil \frac{1}{2}(n+1) \right\rceil$. Thus all labels are ordinals $< 3$. Obviously there are no $n \neq m$ such that $\mathcal{I}(n)$ is embeddable into $\mathcal{I}(m)$ with gap-condition $(s)^\dagger$.

## References


1. Dershowitz, N. (1987). Termination of rewriting. *Journal of Symbolic Computation*, *3*(1–2), 69–116.
2. Friedman, H., McAloon, K., & Simpson, S. (1982). A finite combinatorial principle which is equivalent to the 1 consistency of predicative analysis. In G. Metakides (Ed.), *Logic Symposium I (Patras 1980)* (pp. 197–230). Amsterdam: North-Holland.
3. Gallier, J. H. (1991). What's so special about Kruskal's theorem and the ordinal $\Gamma_0$? A survey of some results in proof theory. *APAL*, *53*, 199–260.
4. Gordeev, L. (1989). Systems of iterated projective ordinal notations and combinatorial statements about binary labeled trees. *Archive for Mathematical Logic*, *29*, 29–46.

5. Gordeev, L. (1989). Generalizations of the one-dimensional version of the Kruskal-Friedman theorems. *The Journal of Symbolic Logic*, *54*(1), 100–121.
6. Gordeev, L. (1990). Generalizations of the Kruskal-Friedman theorems. *The Journal of Symbolic Logic*, *55*(1), 157–181.
7. Gordeev, L. (1993). Quasi-ordinals and proof theory. *Contemporary Mathematics*, *147*, 485–494.
8. Higman, G. (1952). Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society*, *2*, 326–336.
9. Křiž, I. (1989). Well-quasiordering finite trees with gap-condition. *Proof of Harvey Friedman's conjecture Annals of Mathematics*, *130*, 215–226.
10. Kruskal, J. B. (1960). Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture. *Transactions of the American Mathematical Society*, *95*, 210–225.
11. Pohlers, W. (1990). *Proof Theory: An introduction*, Lecture Notes in Math (Vol. 1407). Springer.
12. Rathjen, M. (2010) Investigations of subsystems of second order arithmetic and set theory in strength between $\Pi^1_1$-CA and $\Delta^1_2$-CA+BI: Part I. In R. Schindler (Ed.), *Ways of proof theory* (pp. 363–439). Ontos Verlag
13. Rathjen, M., & Weiermann, A. (1993). Proof-theoretic investigations on Kruskal's theorem. *Annals of Pure and applied Logic*, *61*(1), 49–88.
14. Simpson, S. G. (1985). Nonprovability of certain combinatorial properties of finite trees. *Harvey Friedman's research on the foundation of mathematics* (pp. 87–117). Amsterdam: North-Holland.
15. Schütte, K., & Simpson, S. G. (1985). Ein in der reinen Zahlentheorie unbeweisbarer Satz über endliche Folgen von natürlichen Zahlen. *Archive for Mathematical Logic*, *25*, 75–89.

# Well Quasi-orderings and Roots of Polynomials in a Hahn Field

**Julia F. Knight and Karen Lange**

**Abstract** Let $G$ be a divisible ordered Abelian group, and let $K$ be a field. The *Hahn field* $K((G))$ is a field of formal power series, with terms corresponding to elements in a well ordered subset of $G$ and the coefficients coming from $K$. Ideas going back to Newton show that if $K$ is either algebraically closed of characteristic 0, or real closed, then the same is true for $K((G))$. Results of Mourgues and Ressayre [11] led us to look for bounds on the lengths of roots of a polynomial, in terms of the lengths of the coefficients [5, 6]. In the present paper, we give an introduction to Hahn fields, we indicate how well quasi-orderings arise when we try to bound the lengths of sums and products, and we re-work, in a more general way, a technical theorem from [6] that gives information on the root-taking process.

## 1 Introduction

Recall that $Th(\mathbb{C}, \cdot, 0, 1)$ is the theory of algebraically closed fields of characteristic 0. Similarly, $Th(\mathbb{R}, \cdot, 0, 1, <)$ is the theory of real closed ordered fields. In a real closed ordered field, the odd-degree polynomials over the field have roots, and, in addition, the non-negative elements of the field have square roots. A real closed ordered field $R$ is *Archimedean* if there are no infinite elements; i.e., for each positive $r \in R$, there is some positive integer $n$ such that $r < n$. Equivalently, we may say that there are no infinitesimal elements; i.e., for each positive $r \in R$, there is some positive integer $n$ such that $\frac{1}{n} < r$.

There are non-Archimedean real closed ordered fields. We can apply Compactness to produce such fields. There are also natural examples, such as the field of "Puiseux

J. F. Knight (✉)
Department of Mathematics, University of Notre Dame, 255 Hurley Hall, Notre Dame, IN 46556, USA
e-mail: knight.1@nd.edu

K. Lange
Department of Mathematics, Wellesley College, 106 Central St., Wellesley, MA 02482, USA
e-mail: karen.lange@wellesley.edu



P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_5

series" over $\mathbb{R}$. Puiseux series were introduced by Newton [13], and then studied by Puiseux [15, 16]. For a given field $K$, the *Puiseux series* over $K$ have the form $\sum_{k\leq z\in\mathbb{Z}} a_z t^{\frac{z}{n}}$, where $n$ is a positive integer, and $k$ is an integer.

**Theorem 1.1** (Newton-Puiseux) *1. If $K$ is an algebraically closed field of characteristic 0, then the set of Puiseux series over $K$ is algebraically closed.*
*2. If $K$ is a real closed ordered field, then the set of Puiseux series over $K$ is also real closed.*

There are familiar examples showing that the theorem fails for $K$ of positive characteristic. Hahn fields, introduced by Hahn [4], generalize the fields of Puiseux series. While Puiseux series have length at most $\omega$, the elements of a Hahn field $K((G))$ have more general ordinal lengths. MacLane [9] extended the Newton-Puiseux Theorem to Hahn fields. Mal'tsev [10] and Neumann [12] generalized Hahn fields further, allowing coefficients in a division ring rather than a field.

In the present paper, we consider Hahn fields $K((G))$, where $K$ is a field, either algebraically closed of characteristic 0, or real closed. Maclane's work shows that various polynomials have roots. Our goal is to understand the lengths of these roots in terms of the lengths of the coefficients. Results on well quasi-orderings let us bound the lengths of sums and products. In [6], there is a technical theorem, bounding the lengths of roots of a polynomial over a Hahn field $K((G))$, in terms of the lengths of the coefficients. The technical theorem in [6] applies to all fields $K$ that are real closed or algebraically closed of characteristic 0. However, we require that the group $G$ be "Archimedean", where this means that $G$ is isomorphic to a subgroup of the additive group of reals. We shall re-work the technical theorem from [6] to give a more general result, without the requirement that $G$ be Archimedean. We obtain the result from [6] as a corollary.

Our interest in bounding the lengths of roots of polynomials over a Hahn field grew out of work of Mourgues and Ressayre on "integer parts" for real closed fields. For a real closed ordered field $R$, an *integer part* is a discrete ordered subring $Z$ that is appropriate for the range of a floor function; i.e., for each $r \in R$, there is a unique $z \in Z$ such that $z \leq r < z + 1$. Mourgues and Ressayre [11] showed that every real closed ordered field has an integer part. The construction is quite complicated. It involves embedding the given real closed field in a Hahn field, following a very definite procedure. We wanted to measure the complexity of this procedure. To do so, we needed to bound the lengths of elements in the range of the embedding.

We now indicate what is in each of the remaining sections. In Sect. 2, we say what is a Hahn field. In Sect. 3, using results on well quasi-orderings, we bound the lengths of the sum and product of two elements in $K((G))$ in terms of their individual lengths. In Sect. 4, we re-work the result from [6] on roots of polynomials over a Hahn field. Finally, in Sect. 5, we discuss the Mourgues and Ressayre procedure for constructing integer parts for real closed fields. We indicate what results on lengths tell us about the complexity of this procedure.

## 2 Hahn Fields

We now formally define the Hahn field $K((G))$, as in [4].

**Definition 2.1** (*Hahn*) Let $G$ be an ordered Abelian group, and let $K$ be a field.

1. The Hahn field $K((G))$ consists of formal sums $s = \sum_{g \in S} a_g t^g$ in an indeterminate $t$, where $S \subset G$ is well ordered and $a_g \in K^{\neq 0}$. We write 0 for the empty sum.
2. The *support* of $s$, denoted by $Supp(s)$, is $S$, and the *length* of $s$ is the order type of $S$ under the ordering of $G$.
3. The operations on $K((G))$ are as for ordinary power series.

   (a) For the sum $s + s'$, the coefficient of $t^g$ is the sum of the coefficients of $t^g$ in $s$ and $s'$. Hence, $Supp(s + s')$ is the set of elements $g \in Supp(s) \cup Supp(s')$ such that the coefficient in the sum is nonzero. So,

   $$Supp(s + s') \subseteq Supp(s) \cup Supp(s').$$

   (b) For the product $s \cdot s'$, the coefficient of $t^g$ is the sum of the products $a_h \cdot b_{h'}$ where $(h, h') \in Supp(s) \times Supp(s')$, $a_h$ and $b_{h'}$ are the coefficients of $t^h$ in $s$ and $t^{h'}$ in $s'$ respectively, and $h + h' = g$. So,

   $$Supp(s \cdot s') \subseteq \{h + h' : h \in Supp(s) \,\&\, h' \in Supp(s')\}.$$

4. If $K$ is an ordered field, then $K((G))$ is also ordered. An element $s$ of $K((G))$ is *positive* if the coefficient of the leading term is positive.
5. Whether or not $K$ is ordered, the ordering on $G$ provides a valuation on $K((G))$. The *natural valuation* is the function $w : K((G)) \to G \cup \{\infty\}$ such that $w(s) = \min Supp(s)$ if $s \neq 0$, or $\infty$ if $s = 0$.

Maclane [9] extended the Newton-Puiseux Theorem (Theorem 1.1) to Hahn fields.

**Theorem 2.2** (Generalized Newton-Puiseux Theorem) *Let $G$ be a divisible ordered Abelian group, and let $K$ be a field.*

1. *If $K$ is algebraically closed of characteristic 0, then $K((G))$ is algebraically closed.*
2. *If $K$ is real closed, then $K((G))$ is also real closed.*

For the remainder of the paper, we consider Hahn fields $K((G))$ where $G$ is divisible ordered abelian group and $K$ is a real closed ordered field or algebraically closed field of characteristic 0. We will describe results on the following problem.

**Problem 2.3** *Let $p(x) = A_0 + A_1 x + \cdots + A_n x^n$ be a polynomial over a Hahn field $K((G))$, and let $r$ be a root of $p(x)$. Describe the support of $r$ in terms of the supports of the coefficients $A_i$. In particular, bound the length of $r$ in terms of the lengths of the $A_i$.*

# 3 Well Quasi-orderings and Hahn Fields

In this section, we recall some results on well quasi-orderings. We see how these results can be used to bound the lengths of sums and products in a Hahn field.

## 3.1 Background on Well Quasi-orderings

We begin with basic definitions.

**Definition 3.1** 1. A *quasi-ordering*, or *pre-ordering*, is a set with a binary relation that is reflexive and transitive.
2. A *well quasi-ordering* is a quasi-ordering $\mathcal{A} = (A, \leq)$ in which any infinite sequence $(x_i)_{i\in\omega}$ contains an increasing pair $x_i \leq x_j$ where $i < j$.
3. A *partial ordering* is a quasi-ordering that is anti-symmetric.
4. A *well partial ordering* is a well quasi-ordering that is anti-symmetric. This means that there is no infinite strictly decreasing sequence and there is no infinite anti-chain.

**Definition 3.2** A *linearization* of a well partial-ordering $(A, \leq)$ is a total ordering $(A, \leq^*)$ such that $\leq^*$ contains $\leq$.

Every partial ordering has a linearization.

**Proposition 3.3** *Any linearization of a well partial ordering is a well ordering.*

*Proof* Let $(A, \leq)$ be a well partial ordering, and let $(A, \leq^*)$ be a linearization. We show that there is no infinite $<^*$-decreasing sequence. Consider a sequence $(a_n)_{n\in\omega}$ of distinct elements. We consider the atomic formulas satisfied by the pairs $(a_m, a_n)$ in $(A, \leq)$. By Ramsey's Theorem, there is an infinite set $I \subseteq \omega$ such that one of the following holds for all pairs $m < n$ in $I$.

1. $a_m \leq a_n$ and $a_n \not\leq a_m$,
2. $a_n \leq a_m$ and $a_m \not\leq a_n$,
3. $a_n \leq a_m$ and $a_m \leq a_n$,
4. $a_m \not\leq a_n$ and $a_n \not\leq a_m$.

We cannot have (3) just since $(A, \leq)$ is a partial ordering. We cannot have (2) or (4) since $(A, \leq)$ is a well partial ordering. Therefore, we must have (1). Then the sequence $(a_n)_{n\in\omega}$ is not $<^*$-decreasing. □

**Notation** For a well partial ordering $\mathcal{A} = (A, \leq)$, we let $o(\mathcal{A})$ be the supremum of the ordinals that are the order types of linearizations of $\mathcal{A}$.

A priori, it might seem possible that $o(\mathcal{A})$ is the first ordinal greater than the order types of all linearizations of $\mathcal{A}$. In fact, de Jongh and Parikh [3] and Schmidt [20] showed that $o(\mathcal{A})$ is the maximum of the order types of linearizations (see also [8, 19, 21]).

**Theorem 3.4** *If $\mathcal{A}$ is a well partial ordering, then $o(\mathcal{A})$ is the order type of some linearization of $\mathcal{A}$.*

**Notation** For a linear ordering (or even a partial ordering) $<$ on a set $A$, and for $a \in A$, we write $<\text{-}pred(a)$ for $\{b \in A : b < a\}$. When there is no possibility of confusion about the ordering, we may simply write $pred(a)$.

Schmidt asked whether, for a computable well partial ordering $\mathcal{A}$, the ordinal $o(\mathcal{A})$ is computable. Montalbán [10] gave a positive answer, giving further information useful in reverse mathematics. Below, we sketch a quite different, softer, proof.

**Theorem 3.5** (Montalbán) *If $\mathcal{A}$ is a computable well partial ordering, then $o(\mathcal{A})$ is a computable ordinal.*

*Proof* (Proof sketch) We suppose that $o(\mathcal{A})$ is not computable, expecting a contradiction. Since $\mathcal{A}$ is computable, it is "hyperarithmetically" saturated. For a discussion of this notion, see [1]. By results of Ressayre [17, 18], for any $\Pi^1_1$ set $\Gamma$ of "computable infinitary" sentences in an expanded language, with symbols from the language of $\mathcal{A}$ plus finitely many new symbols, if the consequences of $\Gamma$ are true in $\mathcal{A}$, then there is an expansion $\mathcal{A}^*$ of $\mathcal{A}$ satisfying $\Gamma$. Moreover, we can take $\mathcal{A}^*$ to be hyperarithmetically saturated.

Let $\Gamma$ consist of sentences saying, about a new symbol $\leq^*$, the following:

1. $\leq^*$ is a linear ordering of the universe,
2. $\leq^*$ extends $\leq$,
3. for all computable ordinals $\alpha$, there is a $\leq^*$-initial segment of order type $\alpha$.

Together, the sentences in $\Gamma$ assure that $\leq^*$ is a linearization of $\leq$ of order type greater than any computable ordinal. If $o(\mathcal{A})$ is not a computable ordinal, then for any hyperarithmetical set $\Gamma' \subseteq \Gamma$, there is an expansion of $\mathcal{A}$ satisfying $\Gamma'$. Then by Ressayre's result, there is a hyperarithmetically saturated expansion $\mathcal{A}^* = (A, \leq, \leq^*)$ satisfying all of $\Gamma$. Using the fact that $\mathcal{A}^*$ is hyperarithmetically saturated, we can show that there is a $<^*$-decreasing sequence $(a_n)_{n\in\omega}$. There is a $\Pi^1_1$ set $\Lambda_0(x)$ of computable infinitary formulas saying that for all computable ordinals $\alpha$, $<^* -pred(x)$ has an initial segment of order type $\alpha$. Every hyperarithmetical subset of $\Lambda_0(x)$ is satisfied in $\mathcal{A}^*$, so by hyperarithmetical saturation, some element $a_0$ satisfies all of $\Lambda_0(x)$. There is a $\Pi^1_1$ set $\Lambda_1(a_0, x)$, consisting of the formulas of $\Lambda_0(x)$, plus $x < a_0$. Again, every hyperarithmetical subset is satisfied, so the whole set is satisfied by some $a_1$. Continuing, we get $a_0 > a_1 > a_2 > \dots$. This means that our linearization is not well ordered, a contradiction. Therefore, $o(\mathcal{A})$ must be a computable ordinal. □

### 3.2 Sums and Products

There is a natural way to take sums and products of well partial orderings.

**Definition 3.6** Let $\mathcal{A} = (A, \leq_A)$ and $\mathcal{B} = (B, \leq_B)$ be well partial orderings.

1. The *sum of* $\mathcal{A}$ *and* $\mathcal{B}$, denoted by $\mathcal{A} \sqcup \mathcal{B}$, is $(A \sqcup B, \leq)$, where $A \sqcup B$ is the union of universes that have been made disjoint and $\leq$ is the union of $\leq_A$ and $\leq_B$; i.e., $x \leq y$ iff either $x, y \in A$ and $x \leq_A y$ or $x, y \in B$ and $x \leq_B y$.
2. The *product of* $\mathcal{A}$ *and* $\mathcal{B}$, denoted by $\mathcal{A} \times \mathcal{B}$, is $(A \times B, \leq)$, where $(x, y) \leq (x', y')$ iff $x \leq_A x'$ and $y \leq_B y'$.

It is easy to see that the sum and product are both well partial orderings. Moreover, we can bound the maximum order type of the linearizations. For a precise result, we need the following definitions.

**Definition 3.7** (*Commutative Sum and Product*) 1. For ordinals $\alpha$ and $\beta$, the *commutative sum*, denoted by $\alpha \oplus \beta$, is the ordinal whose Cantor normal form is obtained by expressing $\alpha$ and $\beta$ in Cantor normal form and summing the coefficients of like terms.
2. For ordinals $\alpha$ and $\beta$, the *commutative product*, denoted by $\alpha \otimes \beta$, is the ordinal whose Cantor normal form is obtained by expressing $\alpha$ and $\beta$ in Cantor normal form, multiplying as for polynomials, using commutative sum in the exponents, and then combining like terms.

**Theorem 3.8** (de Jongh-Parikh, Schmidt) *Let* $\mathcal{A}$ *and* $\mathcal{B}$ *be well partial orderings. Then*

1. $o(\mathcal{A} \sqcup \mathcal{B}) = o(\mathcal{A}) \oplus o(\mathcal{B})$
2. $o(\mathcal{A} \times \mathcal{B}) = o(\mathcal{A}) \otimes o(\mathcal{B})$.

De Jongh and Parikh proved the theorem in their 1977 paper [3], and Schmidt also proved it in work toward her 1979 Habilitation [19, 20]. The special case where $\mathcal{A}$ and $\mathcal{B}$ are well-orderings was proved much earlier, in a 1942 paper of Carruth [2].

**Theorem 3.9** (Carruth) *For well orderings* $\mathcal{A}$ *of type* $\alpha$ *and* $\mathcal{B}$ *of type* $\beta$*,*

1. $o(\mathcal{A} \sqcup \mathcal{B}) = \alpha \oplus \beta$,
2. $o(\mathcal{A} \otimes \mathcal{B}) = \alpha \otimes \beta$.

In fact, Carruth was interested, just as we are, in well ordered subsets of an ordered Abelian group.

**Corollary 3.10** (Carruth) *Let* $G$ *be an ordered Abelian group, and let* $A$ *and* $B$ *be well ordered subsets of* $G$.

1. *The order type of $A \cup B$ is bounded by the commutative sum of the order types of $A$ and $B$.*
2. *The order type of the set $A + B := \{a + b : a \in A \ \&\ b \in B\}$ is bounded by the commutative product of the order types of $A$ and $B$.*

*Proof* Each statement in Corollary 3.10 can be obtained from the corresponding statement in Theorem 3.9. □

Carruth also considered the semi-group $[A]$, consisting of finite sums of elements from a well ordered set $A \subseteq G^{\geq 0}$. Of course, if $A$ contained a negative element, then $[A]$ would not be well ordered.

**Theorem 3.11** (Carruth) *Let $G$ be an ordered Abelian group. If $A \subseteq G^{\geq 0}$ is well ordered, then $[A]$ is well ordered. Moreover, if $A$ has order type $\alpha + n$, where $\alpha$ is either 0 or a limit ordinal, then $[A]$ has order type at most $\omega^{\alpha+n} \otimes \alpha^{\alpha}$.*

We now return to Hahn fields. Corollary 3.10 immediately gives bounds on lengths of sums and products in this setting.

**Corollary 3.12** *Let $s, s' \in K((G))$, where $s$ has length $\alpha$ and $s'$ has length $\beta$. Then*

1. *$s + s'$ has length at most $\alpha \oplus \beta$,*
2. *$s \cdot s'$ has length at most $\alpha \otimes \beta$.*

*Proof* We let $A = Supp(s)$, $B = Supp(s')$ and apply Corollary 3.10. □

## 4 Roots of Polynomials

In [6], there is a technical theorem bounding the length of a root of a polynomial over $K((G))$ in terms of the lengths of the coefficients, under the special assumption that the group $G$ is "Archimedean" (we define this below). In this section, we prove a new technical theorem. The new theorem has a different form from the one in [6]. It says that we can pass from the combined support of a polynomial to a set that includes the support of a root, in finitely many steps of prescribed forms. The new theorem is more general than the old. In the new theorem, we make no special assumptions about the group. The old technical theorem follows easily from the new one.

**Definition 4.1** (*Archimedean ordered Abelian group*) We call an ordered Abelian group $G$ *Archimedean* if for all $a, b \in G^{>0}$, there exist natural numbers $m, n$ such that $ma > b$ and $nb > a$.

We note that if the group $G$ is Archimedean and nontrivial, then the Hahn field $\mathbb{R}((G))$ is not Archimedean. If $g \in G^{>0}$, then $t^g$ is an infinitesimal in the field.

**Definition 4.2** (*Indecomposable ordinals*) Let $\alpha$ be an ordinal.

1. $\alpha$ is *additively indecomposable* if $\beta + \gamma < \alpha$ for any $\beta, \gamma < \alpha$.
2. $\alpha$ is *multiplicatively indecomposable* if $\beta \cdot \gamma < \alpha$ for any $\beta, \gamma < \alpha$.

The next result is well-known—see, for example [14].

**Proposition 4.3** *1. The additively indecomposable ordinals, apart from* 0*, are those of the form* $\omega^\gamma$.

*2. The multiplicatively indecomposable ordinals, apart from* 0, 1*, are those of the form* $\omega^{\omega^\gamma}$.

*Moreover, the additively and multiplicatively indecomposable ordinals are* $\oplus$*-indecomposable and* $\otimes$*-indecomposable respectively.*

**Notation 4.4** If $p(x) = A_0 + A_1x + \cdots + A_nx^n$, then we write $Supp(p)$ for $\bigcup_i Supp(A_i)$.

We want to bound the lengths of roots of a polynomial $p(x)$ in terms of the lengths of the coefficients. In what follows, we will focus on infinitesimal roots. We note that for an arbitrary polynomial $p(x)$, $r$ is a root of $p(x)$ iff $r_1 = t^g r$ is a root of $q(x) = p(t^{-g}x)$. Choosing an appropriate negative $g$, we have $q$ with all roots infinitesimal. Clearly, $r$ and $t^g r$ have the same length. Moreover, $q(x) = B_0 + B_1x + \cdots + B_nx^n$, where $B_i = A_it^{-ig}$, so $A_i$ and $B_i$ have the same length.

We will often assume that our polynomials have support in $G^{\geq 0}$. For an arbitrary polynomial $p(x) = A_0 + A_1x + \cdots + A_nx^n$ then $q(x) = t^g p(x)$ has the same roots, with coefficients $B_i = t^g A_i$. So, $w(B_i) = w(A_i) + g$. For an appropriate positive $g$, we have $w(B_i) \geq 0$ for all $i$. There is no change in the roots, and the lengths of $B_i$ and $A_i$ are the same.

Here is the technical theorem from [6] (it is given there as Theorem 3.2).

**Theorem 4.5** *Let $G$ be an Archimedean divisible ordered Abelian group, and let $K$ be a field that is algebraically closed of characteristic* 0*, or real closed. Let $p(x)$ be a polynomial over $K((G))$, where $Supp(p) \subseteq G^{\geq 0}$, and let $r$ be a root of $p$, with $w(r) > 0$. Let $\alpha$ be an infinite multiplicatively indecomposable ordinal. Finally, suppose one of the following two conditions holds:*

1. *the order type of $Supp(p)$ is less than $\alpha$, or*
2. *the order type of $Supp(p)$ equals $\alpha$ and $Supp(p)$ is co-final in $G^{\geq 0}$.*

*Then $r$ has length at most $\alpha$.*

*Remark 4.6* It is enough to prove the result for the case where $K$ is algebraically closed of characteristic 0. As a corollary, we get the statement for an arbitrary field of characteristic 0.

### 4.1 Support of a Root

Our goal in this subsection is to prove a new technical theorem. This theorem says that, for a polynomial $p(x)$ over $K((G))$ and root $r$, where $Supp(p) \subseteq G^{\geq 0}$ and $w(r) > 0$, there is a finite sequence of sets that starts with $Supp(p)$ and ends with a superset of $Supp(r)$, with successive steps obtained by transformations of prescribed kinds. The group $G$ need not be Archimedean. All of the transformations, when applied to well-ordered subsets of $G^{\geq 0}$, yield further well-ordered subsets of $G^{\geq 0}$.

One kind of transformation is a "shift".

**Definition 4.6** (*Shift*) Let $S \subseteq G^{\geq 0}$ and let $g \in G$. A *shift* of $S$ is a set of the form $S' + g = \{h + g : h \in S'\}$, where $S' \subseteq S$ and $S' + g \subseteq G^{\geq 0}$.

*Remark 4.7* Note that we require that $S' + g \subseteq G^{\geq 0}$. Thus, for certain negative $g$ and certain $S' \subseteq S$ (made up of sufficiently large positive elements), $S' - g$ is a shift, even though $S - g$ is not.

**Definition 4.7** An *allowable sequence from $X$ to $Y$* is a finite sequence $(S_i)_{i \leq M}$ of subsets of $G^{\geq 0}$ such that $S_0 = X$, $S_M = Y$, and for each $i$ with $0 < i \leq M$, one of the following holds:

1. *(Semigroup)* there is some $j < i$ such that $S_i = [S_j]$,
2. *(Shift)* there is some $j < i$ such that $S_i$ is a shift of $S_j$,
3. *(Union)* there exist $j, k < i$ such that $S_i = S_j \cup S_k$.

We can now state the new technical theorem.

**Theorem 4.8** (Supports of roots) *Let $G$ be a divisible ordered Abelian group, and let $K$ be a field that is algebraically closed of characteristic* 0. *Let $p(x)$ be a polynomial over $K((G))$ such that $Supp(p) \subseteq G^{\geq 0}$, and suppose $r$ is a root of $p(x)$ with $w(r) >$* 0. *Then there is an allowable sequence from $Supp(p)$ to a superset of $Supp(r)$.*

*Remark 4.9* The definition of allowable sequence involves the group $G$. For a given polynomial $p(x)$ over $K((G))$, the elements of $Supp(p)$ generate (as rational linear combinations) a subgroup $H$ of $G$. We may, if we like, replace $G$ by $H$. The sets leading from $Supp(p)$ to a superset of $Supp(r)$ will all be subsets of $H^{\geq 0}$.

Below, we gather together some definitions and lemmas needed for the proof of Theorem 4.8. These are all given in [6]. Our source was some unpublished notes of Starchenko [22], although there are other proofs available. The terminology that we use in [6] and here is standard, taken from Starchenko's notes.

**Definition 4.9** Let $p(x) = A_0 + A_1 x + \cdots + A_n x^n$ be a polynomial over $K((G))$, and let $\nu \in G$.

1. $p(x)$ is *$k$-semi-regular* if for all $i < k$, $w(A_i) > w(A_k)$, and for all $i > k$, $w(A_i) \geq w(A_k)$.
2. $p(x)$ is *$k$-regular* if $p(x)$ is $k$-semi-regular and $w(A_k) = 0$.

The following is Remark 4.6 of [6].

*Remark 4.10* If $p(x)$ is 0-regular, then $p(x)$ has no infinitesimal roots.

Note that if $w(r) > 0$, then $w(p(r)) = w(A_0)$, so $p(r) \neq 0$.

**Definition 4.11** Let $p(x) = A_0 + A_1x + \cdots + A_nx^n$ be a polynomial over $K((G))$, and let $\nu \in G$.

1. For $A \in K((G))$, the *$\nu$-degree* of the monomial $Ax^i$, denoted by $deg_\nu(Ax^i)$, is $w(A) + i\nu$.
2. The *$\nu$-degree of* $p(x)$, denoted by $deg_\nu(p(x))$, is the minimum of the $\nu$-degrees of the monomials $A_ix^i$ in $p(x)$.
3. The *carrier of* $\nu$ in $p(x)$, denoted by $\Delta_\nu$, is the set of all $i \leq n$ such that $deg_\nu(A_ix^i) = deg_\nu(p(x))$.

The next lemma combines Lemmas 4.2 and 4.16 of [6].

**Lemma 4.12** *Let $p(x) = A_0 + A_1x + \cdots + A_nx^n$ be a nonzero polynomial over $K((G))$.*

1. *The polynomial $p(x)$ is $k$-semi-regular for some $k \leq n$. If $g = w(A_k)$ for this $k$, then $q(x) = t^{-g}p(x)$ is $k$-regular, and $q(x)$ has the same roots as $p(x)$.*
2. *Suppose $p(x)$ is $k$-semi-regular. Let $\nu \in G$, where $\nu > 0$. If $i > k$, then $deg_\nu(A_ix^i) > deg_\nu(A_kx^k)$. Hence, $\Delta_\nu \subseteq \{0, \ldots, k\}$.*

The next theorem is Lemma 4.7 and Theorem 4.18 in [6].

**Theorem 4.13** *Let $G$ be a divisible ordered abelian group, let $K$ be an algebraically closed field of characteristic* 0*, and let $p(x)$ be a nonzero polynomial over $K((G))$. If $p(x)$ is $k$-regular, then $p(x)$ has exactly $k$ roots with positive valuation. Moreover, if $p(x)$ is* 1*-regular, and $r$ is the unique root with $w(r) > 0$, then $Supp(r) \subseteq [Supp(p)]$.*

**Definition 4.14** Let $s \in K((G))$ such that $w(s) \geq 0$. Then $\hat{s}$, the *residue* of $s$, is defined by cases as follows.

- If $w(s) = 0$, then $\hat{s}$ is the coefficient $a \in K$ of the monomial $at^0$ in $s$.
- If $w(s) > 0$, then $\hat{s} = 0$.

**Notation** If $p(x) = A_0 + A_1x + \cdots A_nx^n \in K((G))[x]$ with $Supp(p) \subseteq G^{\geq 0}$, then we write $\hat{p}(x)$ for the element of $K[x]$ given by $\hat{A}_0 + \hat{A}_1x + \cdots + \hat{A}_nx^n$.

*Remark 4.14* A polynomial $p(x)$ with $Supp(p) \subseteq G^{\geq 0}$ is $k$-regular if and only if the degree $k$ monomial in $\hat{p}(x)$ has a non-zero coefficient and, for all $i < k$, the degree $i$ monomial has coefficient 0.

We are ready to prove Theorem 4.8. Recall the statement.
Theorem 4.8. Let $G$ be a divisible ordered Abelian group, and let $K$ be an algebraically closed field of characteristic 0. Let $p(x)$ be a polynomial over $K((G))$ such that $Supp(p) \subseteq G^{\geq 0}$, and suppose $r$ is a root of $p(x)$ with $w(r) > 0$. Then there is an allowable sequence from $Supp(p)$ to a superset of $Supp(r)$.

*Proof* By Lemma 4.12(1), every polynomial over the Hahn field is $k$-semi-regular for some $k$. To prove the theorem, we show by induction on $k$, that if $p(x)$ is a $k$-semi-regular polynomial and $r$ is a root such that $w(r) > 0$, then there is an allowable sequence from $Supp(p)$ to a superset of $Supp(r)$. Take a polynomial $p(x)$ and a root $r$ satisfying the hypotheses of the theorem. If $r = 0$, then $Supp(r) = \emptyset \subseteq Supp(p)$, so we have an allowable sequence with just the one term $Supp(p)$. We suppose that $r \neq 0$. Let $g$ be the least element of $Supp(p)$. By Lemma 4.12 (1), $r$ is also a root of the $k$-regular polynomial $q(x) = t^{-g}p(x)$. Note that $Supp(q) = Supp(p) - g$ is a shift of $Supp(p)$, and $Supp(q) \subseteq G^{\geq 0}$. We put $Supp(q)$ into our sequence right after $Supp(p)$. Hence, it suffices to find an allowable sequence from $Supp(q)$ to a superset of $Supp(r)$. Since $w(r) > 0$, we must have $k \geq 1$, by Remark 4.10. If $k = 1$, then $Supp(r) \subseteq [Supp(q)]$, by Theorem 4.13, so we can complete our sequence with $[Supp(q)]$.

Now, suppose $k > 1$. Let $q(x) = B_0 + B_1x + \cdots + B_nx^n$. The inductive step will proceed by cases.

**Case A**. Suppose $B_0 = 0$. Then $q(x)$ has the form $xq_0(x)$, where $r$ is a root of $q_0(x)$. Note that $Supp(q_0) = Supp(q)$, and $q_0(x)$ is $(k-1)$-regular. By the Induction Hypothesis, there is an allowable sequence leading from $Supp(q_0)$ to a superset of $Supp(r)$.

**Case B**. Suppose $B_0 \neq 0$. Let $w(r) = \nu$. By hypothesis, $\nu > 0$. Since $r$ is nonzero with positive valuation, $Supp(q)$ cannot be $\{0\}$. Let $\delta$ be the $\nu$-degree of $q(x)$. Let $\Delta_\nu$ be the carrier of $\nu$ in $q(x)$. By Lemma 4.12 (2), $\Delta_\nu \subseteq \{0, \ldots, k\}$.

Below, we will transform both the polynomial and the root.

**Transformation via** $\nu$-degree Let $p_1(x) = t^{-\delta}q(t^\nu x)$. This means that $p_1(x) = C_0 + C_1x + \cdots + C_nx^n$, where $C_i = t^{-\delta}B_it^{i\nu}$. Note that $w(C_i) \geq 0$ for all $i$, and $w(C_i) = 0$ just in case $i \in \Delta_\nu$. Thus, $p_1(x)$ is $i$-regular, where $i$ is the first element of $\Delta_\nu$. We have $Supp(C_i) = Supp(B_i) - (\delta - i\nu)$. Since $Supp(B_i) \subseteq Supp(q)$, $Supp(C_i)$ is a shift of $Supp(q)$. By definition, $Supp(p_1)$ is the union of the sets $Supp(C_i)$. Hence, $Supp(p_1)$ is the union of finitely many shifts of $Supp(q)$. Let $r_1 = t^{-\nu}r$. Then $r_1$ is a root of $p_1(x)$ such that $Supp(r_1) \subseteq G^{\geq 0}$ and $w(r_1) = 0$. Since $Supp(r) = Supp(r_1) + \nu$, a shift, it suffices to find an allowable sequence from $Supp(p_1)$ to $Supp(r_1)$. There are sub-cases. In the first sub-case, we are lucky—there is a short sequence.

**Case B (1)**. Suppose $Supp(r_1) = \{0\}$, or $Supp(p_1) = \{0\}$. We claim that the sequence $Supp(p_1)$, $Supp(r_1)$ suffices. Suppose $Supp(r_1) = \{0\}$. Taking $g \in Supp(p_1)$, we get $\{0\} = \{g\} - g = \{0\}$ as a shift of $Supp(p_1)$. Now, suppose $Supp(p_1) = \{0\}$. Then $p_1$ is a polynomial over $K$. The root $r_1$ must be in $K$, with support $\{0\}$.

We might not be so lucky.

**Case B (2)**. Suppose $Supp(r_1) \neq \{0\}$ and $Supp(p_1) \neq \{0\}$. We express $p_1(x)$ as a sum $q_1(x) + s(x)$, where $q_1(x) = \sum_{i \in \Delta_\nu} C_ix^i$ and $s(x) = \sum_{i \notin \Delta_\nu} C_ix^i$. Since $\Delta_\nu \subseteq \{0, \ldots, k\}$, the polynomial $q_1(x)$ has degree at most $k$. Note that for $i \in \Delta_\nu$, $w(C_i) =$

0, while for $i \notin \Delta_\nu$, $w(C_i) > 0$. Thus, $\hat{s}(x) = 0$ and $\hat{p}_1(x) = \hat{q}_1(x)$, where this is a polynomial over $K$. Since $w(r_1) = 0$, we can write $r_1 = a + r_2$, where $a = \hat{r}_1 \in K$, and $0 < w(r_2) < \infty$. Since $p_1(r_1) = 0$, we have that $\hat{p}_1(\hat{r}_1) = 0$. This means that $a$ is a root of $\hat{q}_1(x)$. Finally, we have $Supp(r_1) = \{0\} \cup Supp(r_2)$. We saw in Case B (1) that $\{0\}$ is a shift of any nonempty set. Since $Supp(r_2)$ is nonempty, $Supp(r_2)$, $\{0\}$, $Supp(r_1)$ is an allowable sequence from $Supp(r_2)$ to $Supp(r_1)$.

**Taylor Transformation** Let $p_2(x) = p_1(a + x) = q_1(a + x) + s(a + x)$. Then $r_2$ is a root of $p_2(x)$, and $Supp(p_2) \subseteq Supp(p_1)$, since $a$ is a constant. By the latter fact, and our earlier comments on $Supp(r_2)$, it is enough to find an allowable sequence from $Supp(p_2)$ to $Supp(r_2)$. We know that $a$ is a root of $\hat{q}_1(x)$. Let $m$ be the multiplicity. Since $q_1(x)$ has degree at most $k$, so does $\hat{q}_1$. Therefore, $m \leq k$. We have $\hat{q}_1(x) = (x - a)^m b(x)$, where $b(x)$ is a polynomial over $K$ with $b(a) \neq 0$, and $\hat{p}_1(x) = (x - a)^m b(x)$. Therefore, $\hat{p}_2(x) = \hat{p}_1(a + x) = x^m b(a + x)$, and $x$ does not divide $b(a + x)$. It follows that $p_2(x)$ is $m$-regular. We consider sub-sub-cases.

**Case B (2) (i)**. Suppose $m < k$. By the Induction Hypothesis, there is an allowable sequence from $Supp(p_2)$ to a superset of $Supp(r_2)$.

**Case B (2) (ii)**. Suppose $m = k$. We have $\hat{p}_2(x) = \hat{q}_1(a + x) = bx^k$, where $b$ is an element of $K$.

**Perturbation Transformation**. We consider the derivative $p_2'(x)$. Since $p_2(x)$ is $k$-regular, $p_2'(x)$ is $(k - 1)$-regular. Clearly, $Supp(p_2') \subseteq Supp(p_2)$. By Theorem 4.13, $p_2'(x)$ has a root $c$ with $w(c) > 0$. We form $p_3(x) = p_2(c + x)$. By the Induction Hypothesis, there is an allowable sequence from $Supp(p_2')$ to a superset of $Supp(c)$. Let $r_3 = r_2 - c$. Then $r_3$ is a root of $p_3(x) = p_2(c + x)$. Since $Supp(p_3) \subseteq Supp(p_2) \cup Supp(c)$ and $Supp(r_2) \subseteq Supp(c) \cup Supp(r_3)$, our search is reduced to finding an allowable sequence from $Supp(p_3)$ to a superset of $Supp(r_3)$. Since $p_2(x)$ is $k$-regular and $c$ has positive valuation, $p_3$ is also $k$-regular, by Taylor's formula. Since $p_2'(c) = 0$, the coefficient of $x$ in $p_3(x)$ is 0. Hence, no carrier $\Delta_\mu$ relative to $p_3(x)$ contains the index 1.

We now run the whole argument with the polynomial $p_3(x)$ replacing $q(x)$ and $r_3$ replacing $r$. Since $\Delta_{w(r_3)}$ does not contain 1, Case B (2) (ii) cannot occur. In particular, the root $a$ of polynomial $\hat{q}_1$ in this argument cannot have multiplicity $k$. If it did, $q_1(x)$ would have a nonzero degree 1 term, contradicting the definition of $q_1(x)$. (Recall that $w(r_1) = 0$, and $a$ is the coefficient of the constant term in $r_1$, so $a \neq 0$.) Therefore, we obtain a sequence from $Supp(p_3)$ to a superset of $Supp(r_3)$, as required. □

The referee suggested the following problem.

**Problem 4.15** *Let G be a divisible ordered Abelian group and let K be an algebraically closed field of characteristic* 0*. For polynomials $p(x)$ over $K((G))$, give a sharp bound (in terms of the degree, or possibly the regularity number) on the number of semigroup steps needed for an allowable sequence from $Supp(p)$ to a superset of $Supp(r)$.*

## 4.2 Lengths of Roots

In this subsection, we apply the new technical theorem, Theorem 4.8, to prove the old one, Theorem 4.5. In the proof, we use the following definition.

**Definition 4.16** Let $S \subseteq G^{\geq 0}$. We say that $S$ is $\alpha$-*good* if $S$ is a well ordered subset of $G$, and for all $b \in G^{>0}$, $pred(b) \cap S$ has order type less than $\alpha$.

Recall the statement we want to prove.
Theorem 4.5. Suppose $G$ is Archimedean. Let $p(x)$ be a polynomial over $K((G))$ with $Supp(p) \subseteq G^{\geq 0}$, and let $r$ be a root of $p(x)$ with $w(r) > 0$. Let $\alpha$ be an infinite multiplicatively indecomposable ordinal. Finally, suppose that $Supp(p)$ has order type at most $\alpha$, and if the order type is $\alpha$, then it is co-final in $G$. Then $r$ has length at most $\alpha$.

*Proof* The hypotheses of the theorem say that $Supp(p)$ is $\alpha$-good. To obtain the conclusion, it is enough to show that $Supp(r)$ has a superset that is $\alpha$-good. By Theorem 4.8, there is an allowable sequence $(S_i)_{i \leq M}$ leading from $S_0 = Supp(p)$ to $S_M$, a superset of $Supp(r)$. For all $i \leq M$, $S_i \subseteq G^{\geq 0}$, and one of the following holds:

1. (Semigroup) there is some $j < i$ such that $S_i = [S_j]$,
2. (Shift) there is some $j < i$ such that $S_i$ is a shift of $S_j$,
3. (Union) there exist $j, k < i$ such that $S_i = S_j \cup S_k$.

It is enough to show that all of the sets $S_i$ are $\alpha$-good. We proceed by induction. The set $S_0 = Supp(p)$ is $\alpha$-good, by hypothesis. The next lemma is given in [6] (see Lemma 2.9).

**Lemma 4.17** *Assuming that $G$ is Archimedean, if $S \subseteq G^{\geq 0}$ is $\alpha$-good, then so is $[S]$.*

*Proof* By Theorem 3.11, $[S]$ is well ordered. Take $b \in G^{>0}$. We show that $pred(b) \cap [S]$ has order type less than $\alpha$. If $S$ has no positive elements, this is trivially true. So, suppose $d$ is the first positive element of $S$. Since $G$ is Archimedean, there is some $n$ such that $nd \geq b$. Then any positive element of $pred(b) \cap [S]$ is the sum of at most $n$ positive elements of $S$, all less than $b$. If $D = pred(b) \cap S$, then $pred(b) \cap [S] \subseteq \cup_{k \leq n} D_k$. Since $S$ is $\alpha$-good, $D$ has order type $\beta$ for some $\beta < \alpha$. For each $k \leq n$, let $D_k$ be the set of sums of $k$ elements of $D$. We take the empty sum to be 0, so $D_0 = \{0\}$. Now, we apply Corollary 3.10. Part 1 of the corrolary says that $D_k$ has order type at most $\beta_k = \underbrace{\beta \otimes \cdots \otimes \beta}_{k}$ (of course, $\beta_0 = 1$). Part 2 of the corollary says that $\cup_{k \leq n} D_k$ has order type at most $\beta^* = \underbrace{\beta_0 \oplus \cdots \oplus \beta_n}_{n}$. Since $\alpha$ is multiplicatively indecomposable, all $\beta_k$ are less than $\alpha$. Since $\alpha$ is additively indecomposable, $\beta^* < \alpha$. This shows that $[S]$ is $\alpha$-good. □

**Lemma 4.18** *Suppose $A$ is a shift of $S$. If $S$ is $\alpha$-good, then so is $A$.*

*Proof* Let $A = S' + g$, where $S' \subseteq S$. The group element $g$ may be either positive or negative, but all elements of $S'$, and of $S' + g$, are non-negative. We must show that for $b \in G^{>0}$, $pred(b) \cap A$ has order type less than $\alpha$. It is enough to show this for $b > g$. For $h \in S'$, we have $h + g < b$ iff $h < b - g$. Since $S$ is $\alpha$-good, $pred(b - g) \cap S$ has order type less than $\alpha$, so $pred(b - g) \cap S'$ has order type less than $\alpha$. The order type of $pred(b) \cap A$ is the same. This argument works for both positive and negative $g$. □

**Lemma 4.19** *Suppose $S$, $S'$ are $\alpha$-good. Then $S \cup S'$ is also $\alpha$-good.*

*Proof* Let $b \in G^{>0}$. Say $A = pred(b) \cap S$ has order type $\beta$ and $B = pred(b) \cap S'$ has order type $\beta'$. By our hypothesis, both $\beta$ and $\beta'$ are less than $\alpha$. By Corollary 3.10, $A \cup B$ has order type at most $\beta \oplus \beta'$. Since $\alpha$ is additively indecomposable, this is less than $\alpha$.

By induction, we conclude that $S_M$ is $\alpha$-good. Since $Supp(r) \subseteq S_M$, we see that $Supp(r)$ has order type at most $\alpha$, completing the proof of Theorem 4.5.

# 5 More on Our Original Motivation

There are no new results proved in this section. In the introduction, it was mentioned that the authors' interest in lengths of roots of polynomials in Hahn fields grew out of work on "integer parts" for real closed fields. The authors set out, in [5], to measure the complexity of a procedure, due to Mourgues and Ressayre [11], for finding integer parts. In [5], there is a conjecture on lengths of elements of certain subfields of a Hahn field, which, if true, would yield a bound on the complexity of the procedure of Mourgues and Ressayre. This conjecture was proved in [6], using Theorem 4.5. In this section, we further describe our original motivations, for the interested reader. First, we discuss integer parts and the work of Mourgues and Ressayre. Next, we state results from [6] bounding lengths of elements. Finally, we state the resulting bound on the complexity of the Mourgues and Ressayre procedure. Although the material in this section could all have gone into the introduction, we felt it would distract the main subject of the paper: finding roots of polynomials over a Hahn field.

## 5.1 *Integer Parts and the Work of Mourgues and Ressayre*

**Definition 5.1** (*Integer part*) Let $R$ be a real closed ordered field. An *integer part* is a discrete ordered subring $I$ such that for all $r \in R$, there exists $i \in I$ such that $i \leq r < i + 1$.

Mourgues and Ressayre [11] proved the following.

**Theorem 5.2** (Mourgues and Ressayre) *Every real closed ordered field has an integer part.*

To prove Theorem 5.2, Mourgues and Ressayre provided an explicit construction that involves embedding the given real closed field $R$ into a Hahn field. Given a well ordering $\preceq$ and a "residue field section" $k$ of $R$, the construction yields a unique "value group section" $G$, and an embedding $d : R \hookrightarrow k((G))$ whose image is "truncation-closed". The authors set out to measure the complexity of the Mourgues and Ressayre procedure, locating it in the hyperarithmetical hierarchy.

We explain the terminology. Throughout this subsection, $R$ is a real closed ordered field. The definitions below are standard.

**Definition 5.3** (*Value group, residue field*)

(i) For $x, y \in R^{>0}$, $x \sim y$ if there exist natural numbers $m, n$ such that $mx > y$ and $ny > x$.
(ii) The *natural value group* $G$ is the collection of $\sim$-classes, under the operation inherited from $(R^{>0}, \cdot)$. The ordering on $G$ is the reverse of the ordering inherited from $R$.
(iii) The *natural valuation* $w : R \to G \cup \{\infty\}$ is the function that takes each $r \in R^{\neq 0}$ to the $\sim$-class of $|r|$. We let $w(0) = \infty$, which we consider to be greater than any $g \in G$.
(iv) The *residue field* $k$ of $R$ is the quotient of the ring of *finite* elements, $R_{fin} := \{r \in R : w(r) \geq 0\}$, by the unique maximal ideal of *infinitesimals*, $\mu_R := \{r \in R : w(r) > 0\}$.

*Remark 5.4* If $R$ is a real closed field, then the natural value group $G$ is a divisible ordered abelian group. The residue field $k$ is an Archimedean real closed field representing the real numbers present in $R$.

**Definition 5.5** (*Sections*) A *value group section* of $R$ is an ordered subgroup of $(R^{>0}, \cdot, >)$ that is isomorphic to the value group under the valuation map $w$. Similarly, a *residue field section* is a subfield of $R$ that is isomorphic to the residue field under the map that takes each element of $R_{fin}$ to the corresponding element of $k = R_{fin}/\mu_R$.

**Definition 5.6** (*Truncation-closed*) A subset $F \subseteq k((G))$ is *truncation-closed* if it is closed under initial segments; i.e., for all $s = \sum_{g \in S} a_g t^g \in k((G))$ and $h \in G$, if $s \in F$, then

$$s_{<h} = \sum_{g \in S \,\&\, g < h} a_g t^g \in F \, .$$

Mourgues and Ressayre [11] observed that if $F$ is a truncation-closed subfield of $k((G))$, then $F$ has a natural integer part

$$I_F = \{s + zt^0 \mid s \in F \,\&\, Supp(s) \subset G^{<0} \,\&\, z \in \mathbb{Z}\},$$

where $G^{<0} = \{g \in G \mid g < 0\}$. They then proved the following.

**Theorem 5.7** (Mourgues and Ressayre) *If $R$ is a real closed ordered field, with a residue field section $k$, then there is a value group section $G$ and an embedding $d : R \hookrightarrow k((G))$ such that*

- *for all $a \in k$, $d(a) = at^0$,*
- *for all $g \in G$, $d(g) = t^g$, and*
- *$d(R)$ is truncation-closed.*

If $d$ is as in Theorem 5.7, and $I_{d(R)}$ is the natural integer part for $d(R)$, then $\{r \in R \mid d(r) \in I_{d(R)}\}$ is an integer part for $R$, completing the proof of Theorem 5.2.

For a given real closed ordered field $R$, a residue field section $k$, and a well ordering $\preceq$ of $R$, the Mourgues and Ressayre construction of the value group section $G$ and the embedding $d : R \hookrightarrow k((G))$ is completely determined. The construction produces a special kind of transcendence basis $(r_\beta)_{\beta<\alpha}$ for $d(R)$ over $k$, with an associated chain of truncation-closed subfields of $k((G))$, $(R_\beta)_{\beta\leq\alpha}$ such that $d(R) = R_\alpha$.

Suppose $R$ is a countable real closed field with universe $\omega$. Fix a residue field section $k$. As a well ordering $\preceq$, we may take the usual ordering on $\omega$. Running the Mourgues and Ressayre construction on these inputs, we obtain a value-group section $G$ and an embedding $d$ of $R$ into $k((G))$ such that the special transcendence basis for $d(R)$ over $k$ has length $\omega$. We have the associated chain of truncation-closed subfields of $k((G))$ $R_n$ with union $d(R)$. In [5], it was conjectured that for $n \geq 1$, the elements of $R_n$ have length at most $\omega^{\omega^{n-1}}$, so the elements of $d(R)$ all have length less than $\omega^{\omega^\omega}$.

## 5.2 *Bounds on Lengths of Elements of Special Subfields*

In this subsection, $K$ is either algebraically closed of characteristic 0 or real closed. The definition below appears in [6], and again in [7].

**Definition 5.8** (i) A *truncation-closed independent sequence* in $K((G))$ is a sequence $(r_\beta)_{\beta<\alpha}$, algebraically independent over $K$, such that for each $\beta < \alpha$, either $r_\beta$ has the form $t^g$ for $g \in G$ or else $r_\beta$ has limit length, and all proper initial segments of $r_\beta$ are algebraic over $K \cup \{r_{\beta'} : \beta' < \beta\}$.

(ii) For a truncation-closed independent sequence $(r_\beta)_{\beta<\alpha}$, the associated *canonical sequence* is the sequence $(R_\beta)_{\beta\leq\alpha}$ such that for each $\beta$, the set $R_\beta$ consists of the elements of $K((G))$ algebraic over $K \cup \{r_\gamma : \gamma < \beta\}$. We call $(r_\beta)_{\beta<\alpha}$ a *truncation-closed basis* for $R_\alpha$.

The definition of canonical sequence isolates the important properties of the sequence $(R_\beta)_{\beta\leq\alpha}$ of subfields of $k((G))$ from the procedure of Mourgues and Ressayre. A key part of the proof of Theorem 5.7 is demonstrating that the $R_\beta$ in any canonical sequence are truncation-closed subfields of $K((G))$. Here are the main results from [6].

**Theorem 5.9** *Let $R$ be a truncation-closed subfield of $K((G))$.*

1. *If $R$ has a truncation-closed basis of length $n \geq 1$, then the elements of $R$ have length at most $\omega^{\omega^{(n-1)}}$.*
2. *If $R$ has a truncation-closed basis of length $\omega$, then the elements of $R$ have length less than $\omega^{\omega^{\omega}}$.*
3. *If $G$ is Archimedean, and $R$ has a truncation-closed basis of length $\alpha \geq \omega$, then the elements of $R$ have length at most $\omega^{\omega^{\alpha}}$.*

Part (2) follows immediately from Part (1). Parts (1) and (3) were proved using Theorem 5.9. In Part (1), the group $G$ need not be Archimedean. To prove it, in the case where $G$ is not Archimedean, we think of elements of $R$ both as elements of $K((G))$, and as elements of $K^*((G^*))$, where $G^*$ is Archimedean and $K^*$ is a truncation-closed subfield of $K((G))$, with a truncation-closed basis of length less than $n$.

In [5, 6], there are examples showing that the bounds in Theorem 5.9 are sharp.

**Theorem 5.10** *There are examples of canonical sequences $(R_\alpha)_{\alpha \leq \gamma}$ such that for each finite $n \geq 1$, $R_n$ has an element of length $\omega^{\omega^{(n-1)}}$, and for successor ordinals $\alpha > \omega$, $R_\alpha$ has an element of length $\omega^{\omega^{\alpha}}$.*

In [7], there are results on lengths of elements of $R \subseteq K((G))$ in the case where $R$ has a truncation-closed basis of arbitrary countable length.

## *5.3 Complexity*

Finally, we say something about the complexity of the Mourgues and Ressayre construction. In [5], it is shown that for a countable real closed field $R$ with universe $\omega$, there is a residue field section $k$ that is $\Pi^0_2(R)$. The usual well ordering $\preceq$ of $\omega$ is computable. Running the Mourgues and Ressayre construction on the inputs $R$, $k$, and $\preceq$, we obtain a value-group section $G$ and an embedding $d$ of $R$ into $k((G))$ such that $d(R)$ has a truncation-closed basis of length at most $\omega$. By Theorem 5.9, the elements of $d(R)$ all have length less than $\omega^{\omega^{\omega}}$. From this, it is possible to show the following.

**Corollary 5.11** *The embedding $d : R \hookrightarrow k((G))$ obtained in the Mourgues and Ressayre construction is $\Delta^0_{\omega^{\omega}}(R)$.*

## References


1. Ash, C. J., & Knight, J. F. (2000). *Computable structures and the hyperarithmetical hierarchy*. Elsevier.
2. Carruth, P. (1942). Arithmetic of ordinals with application to the theory of ordered abelian groups. *Bulletin of the American Mathematical Society*, *48*, 262–271.
3. deJongh, D., & Parikh, R. (1977). Well partial orderings and hierarchies. Proceedings Koninklijke Nederlandse Akademie Sci. *Series A*, *80*, 195–207.
4. Hahn, H. (1995). Über die nichtarchimedischen Größensysteme. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften*, Wien, Mathematisch-Naturwissenschaftliche Klasse (Wien. Ber.) vol. 116(1907), pp. 601–655, reprinted in H. Hahn, *Gesammelte Abhandlungen I*. Springer.
5. Knight, J. F., & Lange, K. (2013). Complexity of structures associated with real closed fields. *Proceedings of the London Mathematical Society*, *107*, 177–197.
6. Knight, J. F., & Lange, K. (2019). Lengths of developments in $K((G))$. *Selecta Mathematica*, *25*.
7. Knight, J. F., & Lange, K., Truncation-closed subfields of a Hahn field. pre-print.
8. Křížš, I., & Thomas, R., Ordinal types in Ramsey Theory and well-partial-ordering theory. In *Mathematics of Ramsey Theory*, ed. by J. Nešetřil & V. Rödl, pp. 57–95.
9. MacLane, S. (1939). The universality of formal power series fields. *Bulletin of the American Mathematical Society*, *45*, 888–890.
10. Mal'tsev, A. L. (1948). On embedding of group algebras in a division algebra. (Russian) *Doklady Akademii Nauk SSSR (N.S.)*, *60*, 1499–1501.
11. Mourgues, M. H., & Ressayre, J.-P. (1993). Every real closed field has an integer part. *Journal of Symbolic Logic*, *58*, 641–647.
12. Neumann, B. H. (1949). On ordered division rings. *Transactions of the American Mathematical Society*, *66*, 202–252.
13. Newton, I. (1960). *Letter to Oldenburg dated 1676 Oct 24* (pp. 126–127). The Correspondence of Isaac Newton II: Cambridge University Press.
14. Pohlers, W. (1980). *Proof theory: An introduction*. Berlin: Springer.
15. Puiseux, V. A. (1850). Recherches sur les fonctions algébriques. *Journal de Mathématiques Pures et Appliquées*, *15*, 365–480.
16. Puiseux, V. A. (1851). Nouvelles recherches sur les fonctions algébriques. *Journal de Mathématiques Pures et Appliquées*, *16*, 228–240.
17. Ressayre, J.-P. (1973). Boolean models and infinitary first order languages. *Annals of Mathematical Logic*, *6*, 41–92.
18. Ressayre, J.-P. (1977). Models with compactness properties relative to an admissible language. *Annals of Mathematical Logic*, *11*, 31–55.
19. Schmidt, D. (1978). Associative ordinal functions, well partial orderings, and a problem of Skolem. *Z. Math. Logik Grundlagen Math.*, *24*, 297–302.
20. Schmidt, D. (1979). *Well partial orderings and their maximal order types*. Heidelberg: Habilitationsschrift.
21. Schmidt, D. (1981). The relation between the height of a well-founded partial ordering and the order types of its chains and antichains. *Journal of Combinatorial Theory, Series B*, *31*, 183–189.
22. Starchenko, S., Notes on roots of polynomials.

# Upper Bounds on the Graph Minor Theorem

**Martin Krombholz and Michael Rathjen**

**Abstract** Lower bounds on the proof-theoretic strength of the graph minor theorem were found over 30 years ago by Friedman, Robertson and Seymour (Metamathematics of the graph minor theorem, pp 229–261, [4]), but upper bounds have always been elusive. We present recently found upper bounds on the graph minor theorem and other theorems appearing in the Graph Minors series. Further, we give some ideas as to how the lower bounds on some of these theorems might be improved.

## 1 Introduction

Graph theory supplies many well-quasi-ordering theorems for proof theory to study. The best known of these is Kruskal's theorem, which as discovered independently by Schmidt [13] and Friedman (published by Simpson [14]) possesses an unusually high proof-theoretic strength that lies above that of $ATR_0$. This result was then extended by Friedman to extended Kruskal's theorem, a form of Kruskal's theorem that uses labelled trees for which the embedding has to obey a certain gap-condition, which was shown to have proof-theoretic strength just above even the theory of $\Pi^1_1$–$CA_0$, the strongest of the five main theories considered in the research program known as reverse mathematics.

Reverse mathematics (RM) strives to classify the strength of particular theorems, or bodies of theorems, of "ordinary" mathematics by means of isolating the essential set existence principles used to prove them, mainly in the framework of subsystems of second order arithmetic. The program is often summarized by saying that there are just five systems, known as the "Big Five", that are sufficient for this classification. The picture of RM that we currently see, though, is more complicated:

M. Krombholz · M. Rathjen (✉)
School of Mathematics, University of Leeds, Leeds LS2 9JT, UK
e-mail: M.Rathjen@leeds.ac.uk

M. Krombholz
e-mail: martin.krombholz@protonmail.com



P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_6

1. Those parts of mathematics that have been analyzed in RM, are mostly results from the 19th century and the early 20th century with rather short proofs (varying from half a page to a few pages in length). By contrast, e.g., the large edifice of mathematics that Wiles' proof of *Fermat's Last Theorem* utilizes has not been analyzed in detail.
2. By now there are quite a number of theorems that do not fit the mold of the Big Five. For instance, Ramsey's theorem for pairs, Kruskal's theorem and the graph minor theorem do not equate to any of them. For several others, such as Hindman's theorem, this is still an open question.
3. There are areas of mathematics where complicated double, triple and more times nested transfinite inductions play a central role. Such proof strategies are particularly frequent in set theory (e.g. in fine structure theory and combinatorial theorems pertaining to $L$) and in higher proof theory (e.g. in the second predicative cut elimination theorem and the impredicative cut elimination and collapsing theorems). As RM is usually presented, one might be tempted to conclude that such transfinite proof modes are absent from or even alien to "ordinary" mathematics. However, they are used in the proof of the graph minor theorem. Are they really necessary for its proof?

In this paper we will be concerned with the proof of the graph minor theorem, which is a fairly recent result. It has a very complicated and long proof that features intricate transfinite inductions. In particular, we will be analyzing these inductions and classify them according to principles that are familiar from proof theory and the foundations of mathematics. As to the importance attributed to the graph minor theorem, let's quote from a book on Graph Theory [2], p. 249.

> Our goal [...] is a single theorem, one which dwarfs any other result in graph theory and may doubtless be counted among the deepest theorems that mathematics has to offer: *in every infinite set of graphs there are two such that one is a minor of the other*. This *graph minor theorem*, inconspicuous though it may look at first glance, has made a fundamental impact both outside graph theory and within. Its proof, due to Neil Robertson and Paul Seymour, takes well over 500 pages.

The starting point of this grand proof is the bounded graph minor theorem, i.e. the graph minor theorem restricted to those graphs of bounded "tree-width". The bounded graph minor theorem was connected to Friedman's extended Kruskal's theorem by Friedman et al. [4], and the two were even shown to be equivalent. This provided a natural example of a theorem of combinatorial mathematics that has extremely high proof-theoretic strength, and at the same time gave a lower bound on the graph minor theorem. While the precise proof-theoretic strength of the bounded graph minor theorem was established by Friedman et al. [4], the same was not the case for the full graph minor theorem, for which not even an upper bound was found, which no doubt was due to the fact that the proof spreads over 500 pages of complicated combinatorial arguments. In the following, we will thus outline how the graph minor theorem and other important theorems of the Graph Minors series, like the immersion theorem, can be proved in $\Pi^1_1$–$\mathrm{CA}_0$ with the additional principles of $\Pi^1_3$-induction and $\Pi^1_2$-bar induction.

## 2 Well-Quasi-ordering Theorems of the Graph Minors Series

The relations of minor and immersion can be understood as finding a certain expansion of one graph $G_1$ in another graph $G_2$. All graphs in this paper are finite and without loops unless noted otherwise, and we denote the vertex set of a graph $G$ by $V(G)$ and its edge set by $E(G)$. For the minor relation, define a minor-expansion of $G_1$ to be a function $f : G_1 \longrightarrow G_2$ so that $v \in V(G_1)$ gets mapped to a connected subgraph $f(v) \subseteq G_2$ so that $f(v) \cap f(u) = \emptyset$ if $u \neq v$, and each edge $e \in E(G_1)$ gets mapped injectively to an edge $f(e) \in E(G_2)$ so that if the endpoints of $e$ are $u$ and $v$, then $f(e)$ connects vertices $u' \in f(u)$ and $v' \in f(v)$. If an expansion of $G_1$ is a subgraph of $G_2$, $G_1$ is said to be a minor of $G_2$, denoted $G_1 \leq G_2$. An immersion relation between graphs $G_1$ and $G_2$ is similarly witnessed by an immersion-expansion $f : G_1 \longrightarrow G_2$ so that vertices of $G_1$ are mapped injectively to vertices of $G_2$, and so that an edge $e$ with endpoints $u$ and $v$ is mapped to a path $f(e)$ in $G_2$ between $f(u)$ and $f(v)$ so that for distinct edges $e_1, e_2 \in E(G_1)$ the paths $f(e_1)$ and $f(e_2)$ are edge-disjoint (but may intersect at vertices), i.e. $E(f(e_1)) \cap E(f(e_2)) = \emptyset$. The graph minor and immersion theorem are then the following theorems.

**Theorem 1** (Graph minor theorem, Robertson and Seymour [11]) *For every sequence $\langle G_i : i \in \mathbb{N}\rangle$ of graphs there are $i < j$ so that $G_i$ is a minor of $G_j$.*

**Theorem 2** (Immersion theorem, Robertson and Seymour [12]) *For every sequence $\langle G_i : i \in \mathbb{N}\rangle$ of graphs there are $i < j$ so that there is an immersion of $G_i$ into $G_j$.*

The proof of the graph minor theorem can be divided into two major steps. First, the excluded minor theorem is proved, which takes up most of the Graph Minors series. The excluded minor theorem says that if one graph $G$ does not contain another graph $H$ as a minor, then $G$ has to have a certain structure, namely that it can be decomposed into parts which are connected in a tree-like shape and can almost be embedded into a surface into which $H$ can not be embedded. This is then used as follows: In a proof of the graph minor theorem, for any sequence of graphs $\langle G_1, G_2, \ldots\rangle$ one may assume that $G_1$ is not a minor of any $G_j$, $j > 1$, as otherwise the graph minor theorem holds. Thus, it suffices to prove the graph minor theorem for any sequence of graphs possessing the structure obtained by applying the excluded minor theorem for $G_1$, for any such $G_1$. This means that it is enough to prove the graph minor theorem for graphs which consist of parts connected in a tree-like shape that are almost embeddable into some fixed surface, which is the second major step of the proof of the graph minor theorem.

The proof of the excluded minor theorem is not very complex from a metamathematical point of view. This is due to the fact that surfaces are uniquely determined by their fundamental polygons, and that graph embeddings on any surface can thus be represented by a natural number encoding a graph drawing with rational coordinates in this fundamental polygon. With this approach, the entire proof of the excluded minor theorem does not feature any infinite objects nor any infinite proof techniques, and it is straightforward to carry it out in $ACA_0$, which will be our base theory in the following. The only papers of the Graph Minors series that use more advanced proof techniques are Graph Minors IV [7], VIII [8], XVIII [9], XIX [10], XX [11] and XXIII [12].

Graph Minors IV [7] proves in a sense an early version of the graph minor theorem for graphs with a certain structure as described above, namely the graph minor theorem for graphs that have bounded tree-width, a property which is defined in terms of tree-decompositions. A tree-decomposition of a graph $G$ is essentially a decomposition of $G$ into parts that are connected in a tree-like shape, i.e. a tree-decomposition of $G$ consists of a tree $T$ and for every $t \in V(T)$ a subgraph $G_t$ of $G$ so that

- $\bigcup_{t \in V(T)} G_t = G$, and
- if an edge $e$ of $T$ has endpoints $t_1$ and $t_2$, and $T_1$ and $T_2$ are the two components of $T$ obtained by removing $e$ from $T$, then every path in $G$ from some $v \in \bigcup_{t \in V(T_1)} G_t$ to some $u \in \bigcup_{t \in V(T_2)} G_t$ has to contain a vertex of $G_{t_1} \cap G_{t_2}$.

The width of such a tree-decomposition is then defined to be $\max_{t \in V(T)} |V(G_t)| - 1$. The tree-width $tw(G)$ of $G$ is the minimum width of all its tree-decompositions, and the bounded graph minor theorem can be stated as follows.

**Theorem 3** (Bounded graph minor theorem, [7]) *Let $n$ be a natural number, then in any sequence $\langle G_i : i \in \mathbb{N}\rangle$ of graphs so that $tw(G_i) \leq n$ for every $i \in \mathbb{N}$, there are $G_i$ and $G_j$ with $i < j$ so that $G_i$ is a minor of $G_j$.*

The bounded graph minor theorem has been analyzed from a metamathematical perspective by Friedman, Robertson and Seymour [4], who determined that its proof-theoretic strength lies just above that of $\Pi^1_1$–$CA_0$. They observed that the bounded graph minor theorem can be proved for each individual tree-width in $\Pi^1_1$–$CA_0$, and since the bounded graph minor theorem is a $\Pi^1_1$-statement, that an application of $\Pi^1_1$-reflection for $\Pi^1_1$–$CA_0$ thus suffices to prove the bounded graph minor theorem. This approach circumvents a $\Pi^1_3$-induction, which is roughly used to show that some minimal bad sequence always exists under certain circumstances, and Friedman et al. [4] in turn showed that no theory of lower proof-theoretic strength than $\Pi^1_1$–$CA_0$ augmented with $\Pi^1_1$-reflection for $\Pi^1_1$–$CA_0$ can prove the bounded graph minor theorem. There is however no such proof for some theorems of Graph Minors IV [7] which are more important for the rest of the Graph Minors series, and for these theorems only the upper bound of $\Pi^1_1$–$CA_0 + \Pi^1_3$-IND is known. Friedman et al. [4] further showed that the bounded graph minor theorem is equivalent to the planar graph minor theorem, i.e. the graph minor theorem for those graphs which can be drawn (or equivalently, embedded) in the plane.

Graph Minors VIII [8] proves a generalization of the planar graph minor theorem. Define for every surface $\Sigma$ the $\Sigma$-graph minor theorem:

**Theorem 4** ($\Sigma$-graph minor theorem) *For every sequence $\langle G_i : i \in \mathbb{N} \ldots\rangle$ of graphs that can be drawn in $\Sigma$ without crossings there are $i < j$ so that $G_i \leq G_j$.*

If $S^2$ denotes the sphere, then the planar graph minor theorem is just the $S^2$-graph minor theorem, since embeddability in the sphere and drawability in the plane are equivalent. Denote by $\forall\Sigma$-GMT the statement that the $\Sigma$-graph minor theorem holds for every surface $\Sigma$. It is shown in Graph Minors VIII that the $\Sigma$-graph minor theorem and $\forall\Sigma$-GMT are indeed true, and it can further be shown that both of these theorems are equivalent to the planar and hence also the bounded graph minor theorem. This is done by extending the proof that each instance of the bounded graph minor theorem is provable in $\Pi^1_1$–$\mathrm{CA}_0$ all the way into Graph Minors VII [8], so that it can be shown that for each surface $\Sigma$, the $\Sigma$-graph minor theorem is provable in $\Pi^1_1$–$\mathrm{CA}_0$. An application of $\Pi^1_1$-reflection for $\Pi^1_1$–$\mathrm{CA}_0$ then establishes the equivalence of $\forall\Sigma$-GMT and the planar graph minor theorem, and hence also that of $\forall\Sigma$-GMT and the bounded graph minor theorem. The results of [4] can thus be extended as follows, see [5].

**Theorem 5** *The following are equivalent over $ACA_0$:*

- *The well-orderedness of the ordinal $\psi_0(\Omega_\omega)$,*
- *Friedman's extended Kruskal's theorem,*
- *the bounded graph minor theorem,*
- *the planar graph minor theorem,*
- *the $\Sigma$-graph minor theorem, for any surface $\Sigma$, and*
- *$\forall\Sigma$-GMT.*

The next use of strong infinitary proof-techniques is in Graph Minors XVIII [9] which provides another restricted form of the graph minor theorem that facilitates the proof of the version of the graph minor theorem necessary for the second major step of the proof of the graph minor theorem outlined above. The theorem of Graph Minors XVIII [9] in a sense allows one to focus on the individual pieces of the graph decomposition obtained by the excluded minor theorem, thereby avoiding the need to work with tree-decompositions. The theorem that these individual pieces of the above graph decomposition are well-quasi-ordered by the minor relation is then proved in Graph Minors XIX [10]. The proof of this version of the graph minor theorem requires a further very strong proof principle, namely that of $\Pi^1_2$-bar induction. In Graph Minors XX [11] these results are then combined to prove the full graph minor theorem. Finally, Graph Minors XXIII [12] proves the immersion theorem and a generalization of the graph minor theorem to hypergraphs in a certain sense.

This generalization to hypergraphs can be stated as follows. For a vertex set $V$ denote by $K_V$ the complete graph on $V$, i.e. the graph with vertex set $V$ in which every two distinct vertices are connected by an edge. Then a collapse $f$ of $G_2$ to $G_1$ is a function mapping vertices of $G_1$ to disjoint connected subgraphs of $K_{V(G_2)}$ and edges of $G_1$ injectively to edges of $G_2$ so that $f(e)$ is incident with a vertex of $f(v)$ whenever $e$ is incident with $v$ for all $e \in E(G_1)$ and $v \in V(G_1)$, and further that for every vertex $v$ and every edge $e_v$ of $f(v)$ with endpoints $v_1$ and $v_2$, there must be an edge of $G_2$ that has among its endpoints the vertices $v_1$ and $v_2$. Further, if $Q$ is a well-quasi-order and the edges of $G_1$ and $G_2$ are labelled via functions

$\phi_1 : E(G_1) \longrightarrow Q$, $\phi_2 : E(G_2) \longrightarrow Q$, then $f$ is also required to respect the edge labels of $G_1$ and $G_2$, in the sense that $\phi_1(e) \leq_Q \phi_2(f(e))$ has to hold for every edge $e \in E(G_1)$. Then Graph Minors XXIII [12] shows that the following generalization of the graph minor theorem holds.

**Theorem 6** *Let $Q$ be a well-quasi-order. Then in every infinite sequence $\langle G_i : i \in \mathbb{N}\rangle$ of $Q$-edge-labelled hypergraphs there are $j > i$ so that there is a collapse of $G_j$ to $G_i$ which respects the labels of $G_i$ and $G_j$.*

Further, Graph Minors XXIII [12] also proves that similar labelled versions of the graph minor and immersion theorem hold. If $Q$ is a well-quasi-order and $\phi_1 : E(G) \longrightarrow Q$, $\phi_2 : E(G) \longrightarrow Q$ are labelling functions for the edges of $G_1$ and $G_2$, then a minor relation $G_1 \leq G_2$ via an expansion $f$ is said to respect these labels if $\phi_1(e) \leq_Q \phi_2(f(e))$ for every edge $e \in G_1$. Similarly, for vertex-labelling functions $\phi_1 : V(G) \longrightarrow Q$, $\phi_2 : V(G) \longrightarrow Q$ the minor relation is said to respect the labels if for every $v \in V(G_1)$ there is a $v' \in f(v)$ so that $\phi_1(v) \leq_Q \phi_2(v')$. If $\phi_1$ and $\phi_2$ are vertex-labelling functions from a well-quasi-order $Q$ of $G_1$ and $G_2$ respectively, say that an immersion $f$ respects this labelling if $\phi_1(v) \leq_Q \phi_2(f(v))$ for every $v \in V(G)$. Then the labelled graph minor and immersion theorem are true as well.

**Theorem 7** (Labelled graph minor theorem) *Let $Q$ be a well-quasi-order and let $\langle G_i : i \in \mathbb{N}\rangle$ be a sequence of $Q$-vertex- and edge-labelled graphs. Then there are $i < j$ and a minor expansion $f : G_i \longrightarrow G_j$ that respects the labels of $G_i$ and $G_j$.*

**Theorem 8** (Labelled immersion theorem) *Let $Q$ be a well-quasi-order and let $\langle G_i : i \in \mathbb{N}\rangle$ be a sequence of $Q$-vertex-labelled graphs. Then there are $i < j$ and an immersion expansion $f : G_i \longrightarrow G_j$ that respects the labels of $G_i$ and $G_j$.*

In order to prove these theorems, Graph Minors XXIII [12] requires another $\Pi^1_2$-bar induction similar to that used in Graph Minors XIX [10]. The bar induction of Graph Minors XIX [10] is used when assuming that a certain class of graph embeddings is minimal with respect to certain properties, in order to prove that the above mentioned sequence of graphs embedded in a surface is good. As said above, the graphs themselves might not actually be completely embeddable in the surface, and so the non-embeddable parts are coded as labels from a well-quasi-order, to provide a (now labelled) graph that is completely embeddable into the surface. When assuming that the set of possible labels is a minimal well-quasi-order so that the set of corresponding graphs is a counterexample, one essentially performs a $\Pi^1_2$-bar induction on a well-quasi-order.

## 3 Bar Induction in the Graph Minors Series

More precisely, in Graph Minors XIX [10] two $\Pi^1_2$-bar inductions and three ordinary $\Pi^1_2$-inductions need to be performed. These inductions take the form of the assumption that there is no minimal bad counterexample to a version of the graph minor theorem. This version of the graph minor theorem is for graphs that are embedded in a fixed surface and have labels from well-quasi-orders on the edges. Further, the minor relation between these graphs is altered in such a way that edges incident with a cuff stay fixed on the surface under minor-expansions, and so that it respects the labels of the well-quasi-order. The minimal counterexample to the graph minor theorem for such graphs is then required to have as few handles, crosscaps, cuffs and edges around cuffs as possible, which correspond to the ordinary $\Pi^1_2$-inductions mentioned above, since the well-quasi-orders for the edges are not required to be the same for "smaller" possible counterexamples.

The $\Pi^1_2$-bar inductions then occur when requiring that the well-quasi-orders of the counterexample are also minimal with respect to the initial ideal ordering and so-called refinement relation. We present the bar induction corresponding to the initial ideal relation in greater detail to illustrate that it can deal with the induction principle actually performed in Graph Minors XIX [10]; the relation corresponding to refinement can be handled analogously. As already noted, the counterexample to our version of the graph minor theorem is required to have labels from a well-quasi-order that is minimal with regard to the initial ideal relation. A well-quasi-order $X$ is an initial ideal of another well-quasi-order $X'$, denoted $X \preceq X'$, if $X \subseteq X'$ and if $X$ is closed downward with regard to $X'$, that is if

$$\forall x \in X \forall x' \in X'(x' \leq_{X'} x \to x' \in X).$$

Assuming that the counterexample has minimal well-quasi-orders with regard to this relation then corresponds to the induction scheme

$$\forall X(WQO(X) \to (\forall X' \prec X(\forall X'' \prec X' \varphi(X'') \to \varphi(X')) \to \varphi(X))).$$

This is different from the standard bar induction scheme, which postulates that

$$\forall X(WF(X) \to \forall j(\forall i <_X j \varphi(i) \to \varphi(j)) \to \forall n \in X \varphi(n)).$$

Further, it is not clear whether the induction scheme used in Graph Minors XIX [10] is actually implied by the usual bar-induction scheme, and it does not seem to be the case that this initial ideal induction scheme has been considered before in the literature of reverse mathematics. Note also that due to the different kinds of quantifiers present in second order arithmetic, it may for instance occur that the initial ideal induction scheme quantifies over uncountably many predecessor objects while the ordinary bar induction scheme is constrained to only countably many predecessor objects. Inspecting the proofs of Graph Minors XIX [10] further, it can however be discerned that a more restricted notion of initial ideal is sufficient to carry out the proofs. In the proofs of Graph Minors XIX [10], the minimality of the counterexample with regard to this initial ideal relation is only used when a whole segment above a certain element is "cut out" of the well-quasi-ordering, that is only the relation $\preceq_1$ defined by

$$X' \prec_1 X :\Leftrightarrow \exists \langle x_1, \ldots, x_n \rangle \in X^{<\omega} \forall x'(x' \in X' \leftrightarrow x' \in X \wedge \forall i < n(x' \ngeq x_i))$$

is actually used in Graph Minors XIX [10]. Defining a relation $\leq_1$ (in other contexts known as the Smyth quasi-order) on the finite subsets $[X]^{<\omega}$ of a well-quasi-ordered set $X$ by

$$\{y_1, \ldots, y_n\} \leq_1 \{z_1, \ldots, z_m\} :\Leftrightarrow \forall j \in \{1, \ldots, m\} \exists i \in \{1, \ldots, n\} y_i \leq z_j,$$

and setting $X^{z_1, \ldots, z_n} := \{x \in X : \forall i < n(x \ngeq z_i)\}$ it can be shown that bar induction for $\leq_1$ implies initial ideal induction for $\preceq_1$:

**Lemma 9** *Assume that for every well-quasi-ordered set $X^*$ and every $\Pi^1_2$-formula $\varphi'(n)$ the ordinary bar induction scheme holds with regard to $[X^*]^{<\omega}$ and $\leq_1$, i.e. that*

$$\forall j(\forall i <_1 j\varphi'(i) \rightarrow \varphi'(j)) \rightarrow \forall n \in [X^*]^{<\omega} \varphi'(n).$$

*Then also the initial ideal induction scheme holds for every well-quasi-ordered set $X$ and every $\Pi^1_2$-formula $\varphi(Y)$ with regard to $\preceq_1$, i.e.*

$$\forall X' \prec_1 X(\forall X'' \prec_1 X'\varphi(X'') \rightarrow \varphi(X')) \rightarrow \varphi(X).$$

*Proof* Note that if $X$ is well-quasi-ordered then $\leq_1$ is well-founded on $[X]^{<\omega}$ since a bad $\leq_1$-sequence in $X$ would in particular induce a bad $\preceq$-sequence in $X$ (see e.g. [3]), which is in contradiction to the well-quasi-orderedness of $X$.

Now let $X$ be well-quasi-ordered and let $\top$ be a new element so that $\top > x$ for all $x \in X$. Define $\hat{X} := X \cup \{\top\}$. The idea for showing that the initial ideal induction scheme holds given the ordinary induction scheme is to encode the predecessors of $X$ with regard to $\preceq_1$ by finite subsets of $\hat{X}$, and to perform an ordinary bar induction on $[\hat{X}]^{<\omega}$ instead.

So assume that the usual bar induction scheme for $\Pi^1_2$-formulas with regard to $[\hat{X}]^{<\omega}$ and $\leq_1$ holds. Let $\varphi(X)$ be any $\Pi^1_2$-formula, then we need to show that $\prec_1$-initial ideal induction over $X$ holds for $\varphi$. Hence assume $\varphi$ is progressive with respect to $\prec_1$, i.e. that

$$\forall X' \prec_1 X(\forall X'' \prec_1 X'\varphi(X'') \rightarrow \varphi(X')).$$

Then we need to show that $\varphi(X)$ holds. To do this, we define a formula $\varphi'(i)$ so that $\varphi'(\{y_1, \ldots, y_n\})$ essentially emulates $\varphi(\{x \in \hat{X} : \forall j < n : x \ngeq y_j\})$, as follows:

$$\varphi'(i) := \forall Y(i = \{y_1, \ldots, y_n\} \rightarrow (\forall x(x \in Y \leftrightarrow x \in \hat{X} \wedge \forall j < n : x \ngeq_1 y_j) \rightarrow \varphi(Y))).$$

By $\Sigma^0_0$-comprehension a set $Y$ satisfying the conditions in the antecedent always exists, and so $\varphi'$ is in fact the intended statement. Note that $\varphi'(i)$ is further still a $\Pi^1_2$-formula, and that we can thus utilize our idea to employ $\Pi^1_2$-bar induction for $\varphi'$ in order to show that $\varphi'(\{\top\})$ and hence $\varphi(X)$ holds. To this end we need to prove the progressiveness of $\varphi'$. So assume (letting $i$, $j$ be codes for finite subsets of $\hat{X}$) that $\forall i <_1 j \varphi'(i)$, then we need to show $\varphi'(j)$.

For this, we first show that $\forall i <_1 j \varphi'(i)$ implies $\forall X'' \prec_1 X^j \varphi(X'')$. But if $j = \{x_1, \ldots, x_m\}$, say, then $X'' \prec_1 X^j$ means that $X'' = X^{x_1,\ldots,x_m,z_1,\ldots,z_k}$ for some $z_1, \ldots, z_k$, and trivially $\{x_1, \ldots, x_m, z_1, \ldots, z_k\} <_1 \{x_1, \ldots, x_m\}$, where the inequality must be strict since $X'' \prec_1 X^j$. Let $i = \{x_1, \ldots, x_m, z_1, \ldots, z_k\}$. Then $\varphi'(i)$ holds since we assumed $\forall i <_1 j \varphi'(i)$, and since $X^i = X''$ we can infer that $\varphi(X'')$ holds as well.

So we have shown that $\forall X'' \prec_1 X^j \varphi(X'')$. Since $\varphi$ was assumed to be progressive with regard to $\prec_1$, this gives $\varphi(X^j)$ and therefore $\varphi'(j)$. This is what we needed to show for $\varphi'$ to be progressive. Since $\varphi'$ is progressive we can apply $\Pi^1_2$-bar induction on $\varphi'$ to obtain $\forall x \in [\hat{X}]^{<\omega} \varphi'(x)$. This gives us in particular $\varphi'(\{\top\})$, which in turn implies $\varphi(X)$ and thus completes the proof. □

In the above, finite sets of elements of $X$ are used to code the appropriate subsets of $X$. For the bar induction corresponding to the refinement relation, a finite sequence of such finite sets is needed instead. The critical condition of the refinement relation says in a sense that the well-quasi-orders from which some of the edges are allowed to be labelled can be arranged in such a way that some of those well-quasi-orders are initial ideals of others, and at most identical. More precisely, a sequence $\langle X_1, \ldots, X_n \rangle$ is a refinement of a sequence $\langle X'_1, \ldots, X'_m \rangle$ if $n \geq m$ and there is a function $f : \{1, \ldots, n\} \longrightarrow \{1, \ldots, m\}$ with the property that $X_i \preceq X_{f(i)}$ for all $i \leq n$, so that additionally $X_i, X_j \prec X_{f(i)}$ whenever $f(i) = f(j)$ for $i \neq j$, and so that $X_i \prec X_{f(i)}$ for some $i$. As in the previous induction, the $\prec$-relations are not actually required in their full form and can be replaced by $\prec_1$ relations, which enables us to perform a bar-induction in order to simulate the induction corresponding to the refinement relation. We write $\langle X_1, \ldots, X_n \rangle \prec_2 \langle X'_1, \ldots, X'_m \rangle$ if $\langle X_1, \ldots, X_n \rangle$ is a refinement of $\langle X'_1, \ldots, X'_m \rangle$. To perform the bar-induction, we need a relation corresponding to $\prec_2$. As above, denote the set of finite subsets of a set $Y$ by $[Y]^{<\omega}$, and use $\rho$ and $\sigma$ as variables for such finite subsets. Define then on $([X]^{<\omega})^{<\omega}$ a relation $<_2$ by

$$\begin{aligned}\langle \rho_1, \ldots, \rho_n \rangle <_2 \langle \sigma_1, \ldots, \sigma_m \rangle :\Leftrightarrow \exists f : \{1, \ldots, n\} \longrightarrow \{1, \ldots, m\}\\ (\forall i \leq n(\rho_i \leq_1 \sigma_{f(i)}) \wedge \exists i \leq n(\rho_i <_1 \sigma_{f(i)}) \wedge\\ \forall i, j(i \neq j \wedge f(i) = f(j) \rightarrow \rho_i <_1 \sigma_{f(i)})).\end{aligned}$$

In order to be able to carry out a bar-induction along this relation, we need to show that it is well-founded. This is done in the next lemma.

**Lemma 10** *Let $X$ be a well-quasi-ordered set. Then $([X]^{<\omega})^{<\omega}$ is well-founded with regard to $\leq_2$.*

*Proof* Because $X$ is well-quasi-ordered, $[X]^{<\omega}$ is well-founded with regard to $<_1$ by the remarks in the proof of the above lemma. Our aim is to employ König's lemma in order to show that there can be no infinite descending $\leq_2$-sequence in $([X]^{<\omega})^{<\omega}$. Thus if $\langle \rho_1, \ldots, \rho_n \rangle <_2 \langle \sigma_1, \ldots, \sigma_m \rangle$ via $f$, we say that $\sigma_j$ branches into $\rho_{i_1}, \ldots, \rho_{i_{m_j}}$

if $f^{-1}(j) = \{i_1, \dots, i_{m_j}\}$ and $\rho_{i_1} <_1 \sigma_j$ (which is immediate if $f^{-1}(j)$ consists of more than one element).

Now assume that there is a sequence $s := \left\langle\left\langle\rho^i_1, \dots, \rho^i_{n_i}\right\rangle : i \in \mathbb{N}\right\rangle$ so that $s(i) >_2 s(i+1)$ for all $i$, and let $\langle f_i : \{1, \dots, n_i\} \longrightarrow \{1, \dots, n_{i-1}\}\rangle_{i\geq 2}$ be the corresponding sequence of functions witnessing the $<_2$ relations. In order to avoid confusing duplicate elements that may appear multiple times in that sequence, we interpret each $\rho^i_k$ as a term, and identify two such terms transitively if $\rho^{i+1}_k = \rho^i_l$ and $f_{i+1}(k) = l$.

We now turn toward defining the tree we want to use König's lemma on. Let $S = \{\rho^i_k : i \in \mathbb{N} \wedge k \leq n_i\}$, and for $\rho, \sigma \in S$ define $\sigma$ to be a successor of $\rho$ if at some step in $s$ an element underlying $\rho$ branches into an element underlying $\sigma$. Note that due to the definition of $<_2$ every $\rho$ can branch only once, and that it can only branch into finitely many successors. This successor relation thus defines a forest on $S$, which is infinite since $s$ is an infinite descending sequence and in which every tree is finitely branching. Since this forest consists of $n_1$ and hence finitely many trees, one of these trees must be infinite as well. We can thus apply König's Lemma to this tree to obtain an infinite, strictly decreasing $<_1$-sequence in $[X]^{<\omega}$, which is a contradiction since $[X]^{<\omega}$ is well-founded by $<_1$. □

Similarly to $\prec_1$-initial ideal induction, we can now prove a lemma that shows that ordinary bar induction for $\leq_2$ implies the induction scheme corresponding to refinement. This is made precise in the following lemma.

**Lemma 11** *Assume that for every well-quasi-ordered set $X^*$ and every $\Pi^1_2$-formula $\varphi'(n)$ the bar induction scheme holds with regard to $([X^*]^{<\omega})^{<\omega}$ and $\leq_2$, i.e. that*

$$\forall j(\forall i <_2 j\varphi'(i) \to \varphi'(j)) \to \forall n \in ([X^*]^{<\omega})^{<\omega}\varphi'(n).$$

*Then for every finite sequence of well-quasi-ordered sets $X := \langle X_1, \dots, X_n\rangle$ and every $\Pi^1_2$-formula $\varphi(Y)$ the induction scheme corresponding to refinement*

$$(\forall X' \prec_2 X(\forall X'' \prec_2 X'\varphi(X'') \to \varphi(X')) \to \varphi(X))$$

*holds as well.*

*Proof* The proof is essentially the same as the one for Lemma 9. □

This shows that the critical parts of Graph Minors XIX [10] can be dealt with by a $\Pi^1_2$-bar induction. A similar induction is performed in the proof of the immersion theorem in Graph Minors XXIII [12] that can be dealt with by the same techniques. To give an overview, based on unpublished research we have the following placements of proof-theoretic strength:

(a) $|\Pi^1_1-\mathbf{CA}_0| = \psi_0(\Omega_\omega)$.
(b) $|\Pi^1_1-\mathbf{CA}_0 + \Pi^1_2\text{-IND}| = \psi_0(\Omega_\omega{\cdot}\omega^\omega)$.
(c) $|\Pi^1_1-\mathbf{CA}| = \psi_0(\Omega_\omega{\cdot}\varepsilon_0)$.
(d) $|\Pi^1_1-\mathbf{CA}_0 + \Pi^1_2\text{-BI}| = \psi_0(\Omega^\omega_\omega)$.
(e) $|\Pi^1_1-\mathbf{CA}_0 + \Pi^1_2\text{-BI} + \Pi^1_3\text{-IND}| = \psi_0(\Omega^{\omega^\omega}_\omega)$.
(f) $\psi_0(\Omega_\omega)$ < ordinal of graph minor and immersion theorems $\leq$ $\psi_0(\Omega^{\omega^\omega}_\omega)$.

## 4 Possible Lower Bound Improvements

To narrow down the corridor in which the proof-theoretic strength of the theorems considered above lies, one might try to increase their lower bounds. The immersion theorem with well-quasi-ordered labels seems to be particularly suited for such a task, since it almost imposes an approach similar to that of Friedman's extended Kruskal's theorem $EKT$ [14]. There, a function is used to relate labelled trees ordered by embedding with gap-condition to ordinals from the ordinal notation system $OT(\Omega_\omega)$. This ordinal notation system is used for the ordinal analysis of $\Pi^1_1$–CA$_0$, which shows that $\left|\Pi^1_1\text{–CA}_0\right| = \Psi_0(\Omega_\omega)$, and derived from the set $C_0(\Omega_\omega)$ from Buchholz [1]. In Simpson [14] it is then shown that the above approach yields:

**Theorem 12** *$ACA_0 \vdash EKT \rightarrow WO(\Psi_0(\Omega_\omega))$. In particular, $EKT$ is not provable in $\Pi^1_1$–$CA_0$.*

Similar to $EKT$, a principle $GKT_\omega(Q)$, denoting generalized Kruskal's theorem with labels from $\omega$ and additional well-quasi-ordered labels from a well-quasi-order $Q$, can be defined as follows. First, the objects related to this principle are rooted trees $T$ that have two labelling functions associated with them, one function $l : V(T) \longrightarrow \omega$ and another function $l_Q : V(T) \longrightarrow Q$. They are ordered by embeddings $f : T_1 \longrightarrow T_2$ that satisfy the gap-condition

$$\forall x \in V(T_1)\forall y \in V(T_2)(y \leq f(x) \wedge \neg\exists z \in V(T_1)(z < x \wedge y \leq f(z)) \rightarrow l(y) \geq l(x)),$$

and additionally respect the labels from $Q$ in the sense that

$$\forall x \in V(T_1)(l_Q(x) \leq l_Q(f(x))).$$

For any vertex $v \neq root(T)$ in such a tree, if $w$ is the first vertex on the path from $v$ to $root(T)$, we define $T^v$ to be the component of $T \setminus w$ which includes $v$, and set $root(T^v) := v$. Then one can relate ordinals to a subset of these trees, by decreeing that the well-quasi-order $Q$ have the form $Q = W_Q \cup \{+, \omega^\cdot, \psi\}$, where $W_Q$ is a well-order and the elements of $\{+, \omega^\cdot, \psi\}$ are incomparable to all others, in the following way. First, we need an ordinal notation system $OT(\Omega_\omega \cdot W)$ from [6] which relativizes $OT(\Omega_\omega)$ by putting $\sup(W)$ many copies of $\Omega_\omega$ above $\Omega_\omega$. Interpret a well-order $W$ as an ordinal and for $w \in W$ set $\overline{w} := \Omega_\omega \cdot (1 + w)$. Define then sets $C^W_m(\alpha)$, $m \in \mathbb{N}$, and collapsing functions $\psi^W_m(\alpha)$, $m \in \mathbb{N}$ by induction on $\alpha$. Let $C^W_m(\alpha)$ be the least set $C \supseteq \Omega_m \cup \{\Omega_i : i \in \mathbb{N}\} \cup \{\overline{w} : w \in W\}$ so that:

- $C \cap \Omega_\omega$ is closed under $+$ and $\omega^\cdot$,
- $\overline{w} + \alpha \in C$ whenever $w \in W$ and $\alpha \in C \cap \Omega_\omega$, and
- $C \cap \alpha$ is closed under $\psi_n$ for all $n \in \mathbb{N}$.

Then we can define $\psi^W_m(\alpha)$ by

$$\psi^W_m(\alpha) := \min\{\xi : \xi \notin C^W_m(\alpha)\}.$$

We also write $\psi_m$ instead of $\psi_m^W$ if no confusion is possible. The proof-theoretic ordinal of $\Pi^1_1$–CA in terms of these collapsing functions is then $\psi_0(\Omega_\omega \cdot \varepsilon_0)$. Let $w' := \sup(W)$. In the following we will always assume that ordinals are in normal form with regard to the ordinal notation system $OT(\Omega_\omega \cdot\ W)$ that corresponds to $C_0\left(\overline{w'}\right)$; see [6] for details.

To define the ordinal related to a tree, we additionally assume that $W$ has a special element $w_0$ so that $w_0 < w$ for all $w \in W \setminus \{w_0\}$ (normally $w_0$ would correspond to 0, but we need it to be "less than" 0). We then define $\psi_m(w_0) := \Omega_m$, and to simplify notation, we define further $\psi_m(w + \alpha) := \psi_m(\overline{w} + \alpha)$ for all $w \in W \setminus \{w_0\}$. A tree $T$ can then be assigned an ordinal $o(T)$ from $OT(\Omega_\omega \cdot W) \cap \Omega_\omega$ as follows:

- If $l_Q(root(T)) \in W$ and $root(T)$ has no successor, then set $o(T) := \psi_n(w)$, where $n = l(root(T))$ and $w = l_Q(root(T))$.
- If $l_Q(root(T)) \in W \setminus \{w_0\}$ and $root(T)$ has one successor $v$, then set $o(T) := \psi_n(w + o(T^v))$, where $n = l(root(T))$ and $w = l_Q(root(T))$.
- If $l_Q(root(T)) = +$ and $v_1, v_2$ are the successors of $root(T)$ ordered so that $o(T^{v_1}) \geq o(T^{v_2})$, then set $o(T) := o(T^{v_1}) + o(T^{v_2})$.
- If $l_Q(root(T)) = \omega^\cdot$ and $v$ is the successor of $root(T)$, then set $o(T) := \omega^{o(T^v)}$.
- If $l_Q(root(T)) = \psi$ and $v$ is the successor of $root(T)$, then set $o(T) := \psi_n o(T^v)$, where $n = l(root(T))$.
- If none of these cases can be applied, $T$ is not assigned an ordinal.

In the following we will restrict ourselves to trees that can be assigned an ordinal as above, and well-quasi-orders suitable for labelling those trees. Then it can be shown that:

**Theorem 13** ([5]) *Let $Q$ be a well-quasi-order and $T_1$, $T_2$ be trees as above. Then $o(T_1) \leq o(T_2)$ whenever $T_1 \leq T_2$.*

*In particular, $GKT_\omega(Q)$ implies the well-orderedness of $OT(\Omega_\omega \cdot W_Q)$.*

From which, letting $GKT_\omega(\forall Q) := \forall Q(WQO(Q) \rightarrow GKT_\omega(Q))$, follows immediately:

**Theorem 14** $ACA_0 \vdash GKT_\omega(\forall Q) \rightarrow [\forall X(WO(X) \rightarrow WO(OT(\Omega_\omega \cdot X)))]$.

Then, observing that $\left|\Pi^1_1\text{–CA}\right| = \Psi_0(\Omega_\omega \cdot \varepsilon_0)$, we get stronger lower bounds on $GKT_\omega(\forall Q)$ (and in fact even $GKT_\omega(\varepsilon_0)$).

**Corollary 15** $\Pi^1_1$*–$CA_0 + GKT_\omega(\forall Q)$ proves $WO(\psi_0(\Omega_\omega \cdot \varepsilon_0))$.*

**Corollary 16** $\Pi^1_1$*–$CA \nvdash GKT_\omega(\forall Q)$.*

This idea might possibly be leveraged in the following way, by extending it to theorems of the Graph Minors series. Recall that an immersion of one graph $G_1$ into another graph $G_2$ is an injective function $f : G_1 \longrightarrow G_2$ that maps vertices injectively to vertices and edges to edge-disjoint paths (the paths may intersect at vertices however). Given a labelled tree $T$ as in the statement $GKT_\omega(Q)$ with

$Q = W_Q \cup \{+, \omega^{\cdot}, \psi\}$, one can then define a tree-like graph which under immersion expansion aims to behave like the labelled tree.

Set $Q' := Q \cup \{root\}$ where $root$ is incomparable to all other elements of $Q'$, and define $V(G) := V(T) \cup \{r\}$, where $r$ is a new vertex. Set further $l_{Q'}(v) := l_Q(v)$ if $v \in V(T)$ and set $l_{Q'}(r) := root$. Connect then vertices $v$ of $G$ to their immediate predecessors by $l(v) + 1$ parallel edges, and connect $root(T)$ to $r$ by $l(root(T)) + 1$ parallel edges. We then adopt the notation $v \leq u$ if when deleting edges in $G$ until no multiple edges remain (which results in a tree), $v$ lies on the unique path from $u$ to the vertex labelled with $root$ in $G$. We also speak of predecessors and successors in $G$ with regard to this ordering. For $v$ in $V(G)$ define then $G^v$ to be the induced subgraph of $G$ with vertex-set $\{u \in V(G) : v \leq u\} \cup \{r'\}$ where $r'$ is a new vertex labelled with $root$, and where $r'$ is connected to $v$ by as many edges as $v$ was connected to its immediate predecessor $p(v)$ in $G$. For vertices $v$ not labelled with $root$ set further $l(v) := |\{e \in E(G) : \ e \text{ connects } v \text{ and } p(v)\}| - 1$ (which is the same as $l(v)$ in $T$).

One can then relate an ordinal to $G$ in the obvious way, by definining $o(G)$ as follows:

- If the successor $v$ of $r$ is labelled from $W$ and $v$ has no successors, let $o(G) := \psi_{l(v)}(l_{Q'}(v))$.
- If the successor $v$ of $r$ is labelled from $W$ and $v$ has a successor $w$, let $o(G) := \psi_{l(v)}(l_{Q'}(v) + o(G^w))$.
- If the successor $v$ of $r$ is labelled with $+$, set $o(G) := o(G^{w_1}) + o(G^{w_2})$, where $w_1$ and $w_2$ are the successors of $v$ so that $o(G^{w_1}) \geq o(G^{w_2})$.
- If the successor $v$ of $r$ is labelled with $\omega^{\cdot}$, set $o(G) := \omega^{o(G^w)}$, where $w$ is the successor of $v$.
- If the successor $v$ of $r$ is labelled with $\psi$, set $o(G) := \psi_{l(v)} o(G^w)$, where $w$ is the successor of $v$.

One could hope that $o(G_1) \leq o(G_2)$ whenever $G_1$ can be immersed into $G_2$, but sadly this result has not been established yet. When doing the proof for labelled trees, an induction on the height of the tree with additional induction hypotheses is usually used. However, aside from mapping the vertex labelled with $root$ in $G_1$ to the vertex labelled with $root$ in $G_2$, an immersion from $G_1$ into $G_2$ does not have to respect the "tree-structure" of $G_1$, as illustrated in Fig. 1.

The induction hypotheses necessary for proving $o(G_1) \leq o(G_2)$ can not always be used in such a case, which makes the proof that this holds (if it should indeed hold) a lot harder. It should be noted that the immersion relation between two such graphs corresponds to a root preserving embedding $f$ between edge-labelled trees that is not order or infimum preserving (i.e. so that $f$ maps vertices injectively to vertices and edges to paths that do not have to be disjoint), that however satisfies a different gap-condition, namely that for $e \in E(G_2)$ it has to hold that $l(e) \geq \sum_{e' \in f^{-1}(e)} l(e')$, where $f^{-1}(e)$ denotes the set of edges $e'$ so that $e$ is an edge of $f(e')$.

While it is not clear whether this construction works with immersions due to the above, it should be noted that it does work when using directed graphs and immersions, i.e. so that edges are directed from $u$ to $v$ if $u \leq v$ and so that an immersion

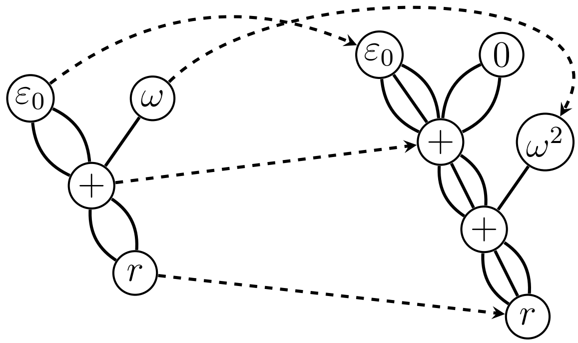



**Fig. 1** One example where a valid immersion embedding does not respect "infima" of the graphs. The labels of the vertices are drawn inside the nodes, with $r$ used instead of $root$. The vertex map of the immersion embedding is given by the dashed arrows, with the edge map implied in the obvious way

expansion maps edges to edge-disjoint directed paths. However, the immersion theorem is known to not hold for the class of all directed graphs in general, and it is currently an ongoing effort in graph theory to establish for which classes of directed graphs it does hold. Thus, it is an open question whether lower bounds like these can be established for a more natural class of directed graphs, and further whether these results can be extended to undirected immersions.

**Acknowledgements** The first author was supported by a scholarship from the University of Leeds ("University of Leeds 110th Anniversary Scholarship").
The second author was supported by a grant from the John Templeton Foundation ("A new dawn of intuitionism: mathematical and philosophical advances", ID 60842) (The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.).

## References

1. Buchholz, W. (1986). A new system of proof-theoretic ordinals. *Annals of Pure and Applied Logic*, *32*(1), 195–207.
2. Diestel, R. (2017). *Graph theory*. 5th ed. Springer.
3. Forster, T. (2003). Better-quasi-orderings and coinduction. *Theoretical Computer Science*, *309*, 111–123.
4. Friedman, H., Robertson, N., & Seymour, P. (1987). Metamathematics of the graph minor theorem. In *Logic and combinatorics*. Ed. S. Simpson. Contemporary Mathematics 65. American Mathematical Society, pp. 229–261.
5. Krombholz, M. (2018). *Proof theory of graph minors and tree embeddings*. Ph.D. thesis. University of Leeds.
6. Rathjen, M., & Thompson, I. A. (n.d.). Well-ordering principles, $\omega$- models and $\prod_1^1$-comprehension. In *The legacy of Kurt Schütte*. Ed. by R. Kahle & M. Rathjen. forthcoming. Springer.
7. Robertson, N., & Seymour, P. (1990a). Graph minors. IV. Tree-width and well-quasi-ordering. *Journal of Combinatorial Theory, Series B*, *48*(1), 227–254.
8. Robertson, N., & Seymour, P. (1990b). Graph minors. VIII. A Kuratowski theorem for general surfaces. *Journal of Combinatorial Theory, Series B 48*(1), 255–288.

9. Robertson, N., & Seymour, P. (2003). Graph minors. XVIII. Tree-decompositions and well-quasi-ordering. In *Journal of Combinatorial Theory, Series B*, *89*(1), 77–108.
10. Robertson, N., & Seymour, P. (2004a). Graph minors. XIX. Well-quasi-ordering on a surface. *Journal of Combinatorial Theory, Series B*, *90*(1), 325–385.
11. Robertson, N., & Seymour, P. (2004b). Graph minors. XX. Wagner's conjecture. *Journal of Combinatorial Theory, Series B*, *92*(1), 325–357.
12. Robertson, N., & Seymour, P. (2010). Graph minors XXIII. Nash-William's immersion conjecture. *Journal of Combinatorial Theory, Series B*, *100*(1), 181–205.
13. Schmidt, D. (1979). Well-partial orderings and their maximal order types. Habilitationsschrift.
14. Simpson, S. (1985). Nonprovability of certain combinatorial properties of finite trees. In *Harvey Friedman's Research on the Foundations of Mathematics*. Studies in Logic and the Foundations of Mathematics. Elsevier Science Publishers (North Holland), pp. 87–117.

# Recent Progress on Well-Quasi-ordering Graphs

**Chun-Hung Liu**

**Abstract** Graphs are arguably the first objects studied in the field of well-quasi-ordering. Giant successes in research on well-quasi-ordering graphs and fruitful extensions of them have been obtained since Vázsonyi proposed the conjecture about well-quasi-ordering trees by the topological minor relation in the 1940's. In this article, we survey recent development of well-quasi-ordering on graphs and directed graphs by various graph containment relations, including the relations of topological minor, minor, immersion, subgraph, and their variants.

## 1 Introduction

A *quasi-ordering* on a set $X$ is a reflexive and transitive binary relation on $X$. A quasi-ordering $\preceq$ on $X$ is a *well-quasi-ordering* if for every infinite sequence $x_1, x_2, \ldots$, there exist $i < i'$ such that $x_i \preceq x_{i'}$. The concept of well-quasi-ordering was discovered from different aspects. One problem that stimulates the development of this concept was raised by Vázsonyi in the 1940's about well-quasi-ordering graphs. Precisely, he conjectured that any infinite collection of trees contains some pair of trees such that one is homeomorphically embeddable in the other (see [29]). This conjecture together with another conjecture of Vázsonyi, which states that subcubic graphs are well-quasi-ordered by the topological minor relation, motivate the study of well-quasi-

Partially supported by the National Science Foundation under Grant No. DMS-1664593.
This material is based upon work supported by the National Science Foundation under Grant No. DMS-1664593.

C.-H. Liu (✉)
Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA
e-mail: chliu@math.tamu.edu

Department of Mathematics, Princeton University, Princeton, NJ 08544-1000, USA



P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_7

ordering on graphs. During past decades, giant successes and fruitful extensions were obtained in this direction.

One benefit of well-quasi-ordering is the existence of finite characterization of properties closed under well-quasi-orderings. Given a quasi-ordering $Q = (V(Q), \preceq_Q)$, a *Q-ideal* $\mathscr{I}$ is a set of graphs such that if $G \in \mathscr{I}$ and $H \preceq_Q G$, then $H \in \mathscr{I}$. For any graph property that is closed under $Q$, the set of graphs satisfying this property is a $Q$-ideal. Let $\mathscr{F}$ be the family of graphs consisting of the minimal graphs that do not belong to $\mathscr{I}$ with respect to $Q$. Then a graph $G$ belongs to $\mathscr{I}$ if and only if $H \not\preceq_Q G$ for every $H \in \mathscr{F}$. Hence, to describe $\mathscr{I}$, it is sufficient to describe $\mathscr{F}$. Since $\mathscr{F}$ is an antichain with respect to $Q$, $\mathscr{F}$ must be finite if $Q$ is a well-quasi-ordering. If for any fixed graph $H \in \mathscr{F}$, one can test whether any input graph $G$ satisfies $H \preceq_Q G$ or not in time polynomial in $|V(G)|$, then $\mathscr{I}$ and hence the corresponding graph property can be tested in polynomial time.

The purpose of this article is to survey recent development on well-quasi-ordering on graphs. Though some notions and results mentioned in this article were extended to other combinatorial objects, such as matroids, permutations or words, we focus on results about graphs only for the simplicity.

This paper is organized as follows. We will discuss the topological minor relation, which is the relation stated in Vázsonyi's conjectures, in Sect. 2. Then we will discuss minor and immersion relations, which are two graph containments closely related to the topological minor relation and attract wide attention, in Sects. 3 and 4, respectively. Finally, we will discuss the subgraph relation, which is the most natural containment on graphs, in Sect. 5.

We start with some formal definitions about graphs. Graphs are finite and possibly have parallel edges and loops in this article. That is, an *(undirected) graph G* consists of a finite set $V(G)$ of vertices and a finite multiset $E(G)$ of 2-element multisubsets of $V(G)$. Each member $e$ of $E(G)$ is called an *edge* of $G$, and its two elements are called the *ends* of $e$. Any edge with no two distinct ends is called a *loop*. Two distinct edges are *parallel* if they have the same ends. Loops and parallel edges are considered as cycles of length 1 and 2, respectively. Two vertices are *adjacent* if they are the ends of the same edge. A vertex is *incident* with an edge if it is an end of this edge. The *degree* of a vertex is the number of edges incident with it, where any loop is counted twice. A graph is *subcubic* if every vertex has degree at most three.

A graph is *simple* if it does not contain any loop or parallel edges. A *directed graph* is a graph equipped with an orientation of its edges. Formally, a directed graph consists of a finite set $V(G)$ of vertices and a finite multiset $E(G)$ of ordered pairs of vertices. If $(x, y) \in E(G)$, then we say that $x$ is the *tail* of $(x, y)$ and $y$ is the *head* of $(x, y)$. The *in-degree* (or *out-degree*, respectively) of a vertex $v$ is the number of edges with head (or tail, respectively) $v$. The *underlying graph* of a directed graph $D$ is a graph obtained from $D$ by removing the direction of the edges. That is, replacing each ordered pair in the edge-set by a 2-element mulitset.

For any positive integer $n$, $[n]$ denotes the set $\{1, 2, ..., n\}$. The *complete graph* on $n$ vertices, denoted by $K_n$, is the simple graph on $n$ vertices with vertices pairwise adjacent. We also call $K_3$ a *triangle*. And $K_n^-$ denotes the simple graph obtained from $K_n$ by deleting an edge. A *stable set* of a graph is a subset of pairwise non-adjacent vertices. A graph is *bipartite* if it is simple and its vertex-set can be partitioned into two stable sets, and we call this partition a *bipartition*. The *complete bipartite graph*, denoted by $K_{m,n}$ for some positive integers $m, n$, is the simple bipartite graph with a bipartition whose one part has $m$ vertices and the other part has $n$ vertices such that any pair of vertices belonging to different parts of this bipartition is adjacent. The *4-wheel* $W_4$ is the simple graph obtained from the cycle of length four by adding a new vertex adjacent to all other vertices. The graph $W_4^-$ is the simple graph obtained from $W_4$ by deleting an edge not incident with the vertex of degree four. The path and cycle on $k$ vertices are denoted by $P_k$ and $C_k$, respectively. Given a simple graph $G$, the *complement of* $G$ is the simple graph with vertex-set $V(G)$ such that any pair of distinct vertices are adjacent if and only if they are non-adjacent in $G$. A *clique* is a set of pairwise adjacent vertices. A *split graph* is a simple graph whose vertex-set can be partitioned into a clique and a stable set. Given a collection $\mathscr{X}$ of sets, the *intersection graph* of $\mathscr{X}$ is the simple graph with vertex-set $\mathscr{X}$, and two distinct vertices $S, T \in \mathscr{X}$ are adjacent if and only if $S \cap T \neq \emptyset$. Given two graphs $G, H$, $G \cup H$ denotes the graph that is a disjoint union of a copy of $G$ and a copy of $H$.

We refer readers to [12] for other undefined standard terminologies about graphs.

## 2 Topological Minors

We focus on the topological minor relation in this section. It is the graph containment that is involved in Vázsonyi's conjectures, so it is arguably the oldest graph containment that is considered for well-quasi-ordering.

Let $G$ be a graph and $v$ a vertex of degree two in $G$. By *suppressing* $v$ we mean deleting $v$ and all its incident edges from $G$, and then adding an edge with ends $x, y$, where the two edges of $G$ incident with $v$ are $\{x, v\}$ and $\{y, v\}$, if $v$ is not incident with a loop; and we simply delete $v$, if $v$ is incident with a loop. Note that suppressing a vertex of degree two is equivalent with contracting an edge incident with it. (Edge-contraction is an operation that will be defined in Sect. 3.)

A graph $G$ contains another graph $H$ as a *topological minor* if $H$ can be obtained from $G$ by repeatedly deleting vertices and edges and suppressing vertices of degree two.

A equivalent way to define the topological minor relation is through the notion of homeomorphic embeddings. For graphs $G$ and $H$, we say that a function $\pi$ with domain $V(H) \cup E(H)$ is a homeomorphic embedding from $H$ into $G$ if the following hold.

- $\pi$ maps vertices of $H$ injectively to vertices of $G$.
- For each non-loop $e$ of $H$ with ends $x, y$, $\pi(e)$ is a path in $G$ with ends $\pi(x)$ and $\pi(y)$.
- For each loop $e$ of $H$ with end $v$, $\pi(e)$ is a cycle in $G$ containing $\pi(v)$.
- If $e_1, e_2$ are distinct edges of $H$, then $\pi(e_1) \cap \pi(e_2) \subseteq \{\pi(t) : t \in e_1 \cap e_2\}$.

It is easy to see that $G$ contains $H$ as a topological minor if and only if there exists a homeomorphic embedding from $H$ into $G$.

Vázsonyi in the 1940's conjectured that trees are well-quasi-ordered by the topological minor relation. This conjecture was proved by Kruskal [28] and independently by Tarkowski [52]. Nash-Williams [41] later introduced the "minimal bad sequence" argument to provide an elegant proof of this conjecture. The minimal bad sequence argument has had a profound impact on proving well-quasi-ordering results since then. Indeed, they proved Vázsonyi's conjecture is true even when vertices are labelled by a well-quasi-ordering.

**Theorem 1** ([28, 41, 52]) *Let $Q = (V(Q), \preceq_Q)$ be a well-quasi-ordering. For each positive integer $i$, let $T_i$ be a tree and let $\phi_i : V(T) \to V(Q)$ be a function. Then there exist $1 \leq j < j'$ and a homeomorphic embedding $\pi$ from $T_j$ into $T_{j'}$ such that $\phi_j(v) \preceq_Q \phi_{j'}(\pi(v))$ for every $v \in V(T_j)$.*

One might expect that Theorem 1 can be generalized in a way that the homeomorphic embedding mentioned in Theorem 1 also preserves the ancestor-descendant relation if we make those trees be rooted trees. But in fact, this stronger version for rooted trees is equivalent with Theorem 1 as one can add a new incomparable element into $Q$ to obtained a new well-quasi-ordering and add this new element into the labels of the roots of those trees.

Theorem 1 is a generalization of a very useful result of Higman [19], which is now known as the Higman's Lemma. Higman's Lemma states that every well-quasi-ordering $Q$ on a set $V(Q)$ can be extended to a well-quasi-ordering on the set of finite sequences over $V(Q)$ by the natural "sequence embedding" relation.

**Theorem 2** ([19]) *If $Q = (V(Q), \preceq_Q)$ is a well-quasi-ordering, then the set of finite sequences over $V(Q)$ is well-quasi-ordered by $\preceq$, where two finite sequences $a = (a_1, a_2, ..., a_m)$ and $b = (b_1, b_2, ..., b_n)$ over $V(Q)$ satisfy $a \preceq b$ if and only if there exist $1 \leq i_1 < i_2 < ... < i_m \leq n$ such that $a_j \preceq_Q b_{i_j}$ for every $j \in [m]$.*

Higman's Lemma is equivalent with the case when every tree $T_i$ is a path in Theorem 1. In addition, by using Higman's Lemma, Theorem 1 can be extended to the case that each $T_i$ is a forest.

Theorem 1 was later generalized by Mader [37] and Fellows, Hermelin and Rosamond [17] as follows.

**Theorem 3** *Let $t$ be a positive integer.*

1. *Reference [37] Graphs that do not contain $t$ disjoint cycles are well-quasi-ordered by the topological minor relation.*
2. *Reference [17] Graphs that have feedback vertex sets with size at most $t$ are well-quasi-ordered by the topological minor relation.*

A *feedback vertex set* in a graph $G$ is a subset of $V(G)$ intersecting all cycles in $G$. In fact, Statement 2 of Theorem 3 can be easily derived from the forest-version of Theorem 1 by appropriately labelling the vertices; Statement 1 of Theorem 3 is equivalent with Statement 2 due to a classical result of Erdős and Pósa [16] stating that a graph has only a bounded number of disjoint cycles if and only if it has a feedback vertex set with bounded size. We remark that Statement 1 was proved much earlier than Statement 2.

Another class of graphs that is known to be well-quasi-ordered by the topological minor relation is the set of subcubic graphs. It was originally conjectured by Vázsonyi and proved by Robertson and Seymour [49] via the Graph Minor Theorem. (The Graph Minor Theorem will be described in Sect. 3.) We remark that the proof of the Graph Minor Theorem is very difficult, and it remains unknown how to prove Vázsonyi's conjecture on subcubic graphs without using the Graph Minor Theorem.

However, the topological minor relation does not well-quasi-order graphs in general. For every positive integer $k$, let $R_k$ be the graph obtained from a path of length $k$ by doubling each edge. The *ends* of $R_k$ are the ends of the original path. We call $R_k$ the *Robertson chain* of length $k$. Let $R'_k$ be the graph obtained from $R_k$ by attaching two leaves to each end of $R_k$. It is easy to see that $\{R'_k : k \geq 1\}$ is an antichain with respect to the topological minor relation.

There is another infinite antichain. Let $R''_k$ be the graph obtained from a cycle of length $k$ by duplicating each edge. Then for any subdivision $R^*_k$ of $R''_k$, $\{R^*_k : k \geq 1\}$ is also an antichain with respect to the topological minor relation. More antichains were known, and all of them contain arbitrarily long Robertson chain as a topological minor.

Robertson in the 1980's conjectured that the Robertson chain is the only obstruction. That is, he conjectured that for every positive integer $k$, the set of graphs that do not contain $R_k$ as a topological minor is well-quasi-ordered by the topological minor relation. We remark that Robertson's conjecture is strong. Though $R_k$ has quite simple structures, the class of graphs with no $R_k$ topological minor is still broad. In particular, every subcubic graph does not contain $R_2$ as a topological minor. So Robertson's conjecture for the case $k = 2$ contains Vázsonyi's subcubic graph conjecture.

Ding [14] proved that a weakening of Robertson's conjecture is true: the set of graphs that do not contain $R_k$ as a minor is well-quasi-ordered by the topological minor relation.

Robertson's conjecture was recently completely solved by Liu and Thomas [32, 34], even when vertices are labeled.

**Theorem 4** ([32, 34]) *For every well-quasi-ordering $Q = (V(Q), \preceq_Q)$ and for every positive integer k, if for each $i \geq 1$, $G_i$ is a graph with no $R_k$ topological minor and $f_i : V(G) \to V(Q)$ is a function, then there exist $1 \leq j < j'$ and a homeomorphic embedding $\pi : G_j \to G_{j'}$ such that $f_j(v) \preceq_Q f_{j'}(\pi(v))$ for every $v \in V(G_j)$.*

Theorem 4 implies all known results about well-quasi-ordering graphs by the topological minor relation. The case $k = 1$ of Theorem 4 implies Kruskal's Tree Theorem (Theorem 1); the case $k = 2t - 1$ implies Mader's theorem for graphs with no $t$ disjoint cycles (Theorem 3) and hence for graphs having feedback vertex sets of bounded size; the case $k = 2$ implies Vázsonyi's conjecture on subcubic graphs. Theorem 4 also implies a well-known result about well-quasi-ordering bounded diamenter graphs by the subgraph relation (see Theorem 24 in Sect. 5).

The proof of Theorem 4 is long and difficult. The first step is to prove the case for graphs with bounded treewidth. It turns out to be harder than expected for graphs with bounded treewidth, which is a case that is not very hard to deal with in the proofs of the Graph Minor Theorem and several algorithmic results in the literature. We remark that the bounded treewidth case of Theorem 4 implies Ding's result, since graphs that do not contain $R_k$ as a minor do not contain a $3 \times (k+1)$ grid as a minor and hence have bounded treewidth. Though the proof of the bounded treewidth case is not simple, the proof is self-contained and does not require the Graph Minor Theorem. One key ingredient is a technique to convert vertex-cuts realized by bags in the tree-decomposition into edge-cuts.

The second step of the proof of Theorem 4 is to study the structure of graphs with large treewidth but with no $R_k$ topological minor. Liu and Thomas [32, 34] prove that such graphs are "nearly subcubic" by extensively applying techniques developed in Robertson and Seymour's Graph Minors series and earlier work of Liu and Thomas [33]. (The formal description of nearly subcubic graphs is complicated, so we skip the details in this paper.)

The third step for proving Theorem 4 is to prove that nearly subcubic graphs are well-quasi-ordered by the topological minor relation. This step requires non-trivial applications of the Graph Minor Theorem, and it is the only step that uses the Graph Minor Theorem in the entire proof.

We remark that Theorem 4 is best possible as long as the vertices are labelled by a well-quasi-ordered set $Q$ with $|V(Q)| \geq 2$. Let $a$ be a maximal element of $Q$, and let $b \in V(Q) - \{a\}$. If we label the ends of $R_k$ by $a$ and label other vertices by $b$, then $\{R_k : k \geq 1\}$ is an antichain with respect to the topological minor relation that preserves ordering on the labels of the vertices. This shows that the converse of Theorem 4 is also true.

However, Robertson's conjecture can be strengthened if the vertices are unlabelled (or equivalently, labelled by $Q$ with $|V(Q)| = 1$), since $\{R_k : k \geq 1\}$ is not an antichain if the vertices are unlabelled. Liu and Thomas [34] also provide a complete characterization for the family of unlabelled graphs that are well-quasi-ordered by the topological minor relation. Such a characterization involves a notion of Robertson family that is defined as follows.

For a positive integer $k$ and an end $v$ of $R_k$, by *planting* on $v$ we mean the operation that either adds a new vertex adjacent to $v$, or adds a new loop incident with $v$; a *thickening* on $v$ is the operation that adds a new edge incident with $v$ and its neighbor; a *strong planting* on $v$ is the operation that either applies planting on $v$ twice, or applies thickening on $v$ once. Let $k$ be a positive integer, the *Robertson cycle* of length $k$ is the graph that can be obtained from the cycle of length $k$ by duplicating each edge.

For each positive integer $k \geq 3$, the *Robertson family* of length $k$ is the set of graphs consisting of the Robertson cycle of length $k$ and the graphs that can be obtained from $R_k$ by either

- strong planting on each end of $R_k$ once, or
- planting on each end of $R_k$ once and adding an edge incident with both ends of $R_k$, or

- planting on one end of $R_k$ once, thickening on the other end once, and adding an edge incident with both ends of $R_k$.

So for each $k \geq 3$, the Robertson family of length $k$ consists of 16 non-isomorphic graphs. Note that the graph $R'_k$ mentioned earlier in the infinite antichain $\{R'_k : k \geq 1\}$ can be obtained from $R_k$ by strong planting on each end of $R_k$. Clearly, the union of the Robertson families of length $k$ over all integers $k$ can be partitioned into 16 infinite antichains with respect to the topological minor relation.

Liu and Thomas [34] prove that Robertson's conjecture can be strengthened to graphs with no topological minor isomorphic to members of Robertson families.

**Theorem 5** ([34]) *For every positive integer $k \geq 3$, if $G_1, G_2, \ldots$ are graphs that do not contain any member of the Robertson family of length at least k as a topological minor, then there exist $1 \leq j < j'$ such that $G_{j'}$ contains $G_j$ as a topological minor.*

Theorem 5 is best possible since to obtain a well-quasi-ordered set, we can only allow finitely many members in each of the 16 infinite disjoint antichains whose union is the union of Robertson families of all lengths. This theorem also provides a characterization of well-quasi-ordered topological minor ideals.

A family $\mathscr{I}$ of graphs is a *topological minor ideal* if every topological minor of any member of $\mathscr{I}$ belongs to $\mathscr{I}$.

**Theorem 6** ([34]) *Let $\mathscr{I}$ be a topological minor ideal. Let $\mathscr{R}$ be the union of the Robertson family of length k over all positive integers $k \geq 3$. Then $\mathscr{I}$ is well-quasi-ordered by the topological minor relation if and only if $\mathscr{I}$ contains only finitely many members of $\mathscr{R}$.*

We remark that Theorems 4 and 6 show a significant difference between well-quasi-ordered topological minor ideals for labelled graphs and for unlabelled graphs. Furthermore, if a topological minor ideal is well-quasi-ordered with a set of two labels, then it cannot contain arbitrarily long Robertson chain, so Theorem 4 shows that it is also well-quasi-ordered with a set of labels of any cardinality. Hence, the cardinality of the set of labels does not affect whether a topological minor ideal is well-quasi-ordered or not, as long as at least two labels are allowed. This fact could be viewed as a possible support for a conjecture of Pazout (Conjecture 7) about a similar situation for the induced subgraph relation, though it could also be viewed as a support for a similar but false conjecture of Kříž and Thomas [27] on QO-categories disproved by Kříž and Sgall [26].

## 2.1 Directed Graphs

Now we consider topological minors for directed graphs. The notion of homeomorphic embedding of undirected graphs naturally extends to directed graphs. A function $\pi$ is a *homeomorphic embedding* from a directed graph $H$ into a directed graph $G$ if the following hold.

- $\pi$ maps vertices of $H$ injectively to vertices of $G$.
- For each non-loop $e$ of $H$ with tail $x$ and head $y$, $\pi(e)$ is a directed path in $G$ from $\pi(x)$ to $\pi(y)$.
- For each loop $e$ of $H$ with end $v$, $\pi(e)$ is a directed cycle in $G$ containing $\pi(v)$.
- If $e_1, e_2$ are distinct edges of $H$, then $\pi(e_1) \cap \pi(e_2) \subseteq \{\pi(t) : t \in e_1 \cap e_2\}$.

We say that a directed graph $G$ contains another directed graph $H$ as a *topological minor* if there exists a homeomorphic embedding from $H$ into $G$.

It is easy to see that the topological minor relation does not well-quasi-order directed graphs, as any orientation of the graphs in $\{R'_k : k \geq 1\}$ form an infinite antichain. Indeed, it is still not a well-quasi-ordering even if we restrict the problem to a specific kind of directed graphs.

A directed graph $G$ is a *tournament* if its underlying graph is a simple graph, and for every pair $u, v$ of distinct vertices of $G$, exactly one of $(u, v)$ and $(v, u)$ belongs to $E(G)$. It seems well-known that tournaments are not well-quasi-ordered by the topological minor relation, but we were not able to find any example of an infinite antichain in the literature. So we provide an example of an infinite antichain here.

For any positive integer $n$, we say a tournament is a *transitive tournament on* $[n]$ if its vertex-set is $[n]$ and every edge is of the form $(i, j)$ with $1 \leq i < j \leq n$. For any positive integer $k$, let $G_k$ be the tournament obtained from the transitive tournament on $[2k + 13]$ by reversing the direction of the edges in $\{(1, 2), (3, 4), (5, 6), (2, 7), (4, 7), (6, 7), (2k + 7, 2k + 8), (2k + 8, 2k + 9), (2k + 7, 2k + 10), (2k + 10, 2k + 11), (2k + 7, 2k + 12), (2k + 12, 2k + 13)\} \cup \{(2i + 5, 2i + 6), (2i + 6, 2i + 7), (2i + 5, 2i + 7) : i \in [k]\}$. Note that the undirected graph formed by the reversed edges is the simple graph obtained from $R_k$ by attaching three leaves to each end of $R_k$ and then subdividing all except one edge in each pair of parallel edges once.

**Theorem 7** $\{G_k : k \geq 1\}$ *is an antichain of tournaments with respect to the topological minor relation.*

*Proof* Suppose to the contrary that there exist $1 < i < j$ and a homeomorphic embedding $\pi$ from $G_i$ to $G_j$. Let $u_1, u_2, ..., u_{2i+13}$ be the vertices $1, 2, ..., 2i + 13$ of $G_i$, respectively; let $v_1, v_2, ..., v_{2j+13}$ be the vertices $1, 2, ..., 2j + 13$ of $G_j$, respectively.

We first show that $\pi(u_7) = v_7$. For $t \in [3]$, let $H_t$ be the directed cycles $\pi((u_7, u_{8-2t})) \cup \pi((u_{8-2t}, u_{7-2t})) \cup \pi((u_{7-2t}, u_7))$ in $G_j$. Suppose that $\pi(u_7) = v_{2r+7}$ for some $r \in [j]$. Since there exists no edge from $\{v_\ell : \ell > 2r + 7\}$ to $\{v_\ell : \ell < 2r + 7\}$ in $G_j$, at most one of $H_1, H_2, H_3$, say $H_1$, contains an edge of the form $(v_{2r+7}, v_x)$ with $x > 2r + 7$. Since $(v_{2r+7}, v_{2r+5})$ and $(v_{2r+7}, v_{2r+6})$ are the only two edges of the form $(v_{2r+7}, v_y)$ with $y < 2r + 7$, one of $H_2, H_3$ contains $(v_{2r+7}, v_{2r+5})$ and the other contains $(v_{2r+7}, v_{2r+6})$. But then $H_2, H_3$ must share $v_{2r+5}$, a contradiction. A similar argument shows that $\pi(u_7) \notin \{v_\ell : \ell \in [6]\} \cup \{v_{2\ell+6} : \ell \in [j]\}$. Since the out-degree of $u_7$ equals $2i + 7$, which is greater than the out-degree of any vertex in $\{v_\ell : 2j + 8 \leq \ell \leq 2j + 13\}$. Hence $\pi(u_7) = v_7$. Similarly, $\pi(u_{2i+7}) = v_{2j+7}$.

Since $u_7$ has in-degree five in $G_i$, in order to accommodate $H_1$, $H_2$, $H_3$, $\pi((u_8, u_7))$ and $\pi((u_9, u_7))$, we have that $\{\pi(u_\ell) : \ell \in [7]\} = \{v_\ell : \ell \in [7]\}$. Since $u_{2i+5}$ has out-degree six in $G_i$, $\pi(u_{2i+5}) \notin \{v_\ell : 2j+8 \leq \ell \leq 2j+13\}$. Then we have $\{\pi(u_\ell) : 2i+8 \leq \ell \leq 2i+13\} = \{v_\ell : 2j+8 \leq \ell \leq 2j+13\}$. Then it is easy to show that $\pi(u_{2\ell+7}) = v_{2\ell+7}$ for each $\ell \in [i] \cup \{0\}$ by induction on $\ell$. In particular, $\pi(u_{2i+7}) = v_{2i+7}$. So $j = i$, a contradiction. This proves the theorem. □

## 3 Minors

Let $G$ be a graph, and $e$ be an edge with ends $x, y$. By *contracting* $e$ we mean deleting $x, y$ from $V(G)$ and adding a new element $w$ into $V(G)$, and deleting $e$ from $E(G)$ and replacing any appearance of $x$ or $y$ in edges by $w$. Note that contracting an edge contained in a triangle will create parallel edges; contracting an edge in a pair of parallel edges will create loops. We say that $G$ contains a graph $H$ as a *minor* if $H$ can be obtained from $G$ by repeatedly deleting vertices and edges and contracting edges.

Wagner [54] conjectured that the minor relation is a well-quasi-ordering. Note that Wagner's conjecture contains Vázsonyi's conjecture on subcubic graphs since the minor relation and the topological minor relation are equivalent for subcubic graphs. Robertson and Seymour [49] proved Wagner's conjecture and hence Vázsonyi's conjecture on subcubic graphs. Note that deriving from Wagner's conjecture is the only currently known proof of Vázsonyi's conjecture on subcubic graphs.

**Theorem 8** ([49]) *Graphs are well-quasi-ordered by the minor relation.*

Theorem 8 is now known as the Graph Minor Theorem. The Graph Minor Theorem is one of the most difficult theorems in graph theory. It is proved in the 20th paper of the famous Graph Minors series by extensively applying the structural theorems developed in other papers of the same series. Robertson and Seymour's groundbreaking work in this series of paper not only solves well-quasi-ordering problems but also opens a new research field in structural graph theory.

Indeed, Robertson and Seymour proved that Theorem 8 is true even when the edges of the graphs are labelled. Formal descriptions for the version of labelled graphs are involved, so we omit the details. We refer interested readers to [49, 50]. A sketch of a proof of Theorem 8 can be found in [12].

A *minor ideal* $\mathscr{I}$ is a set of graphs such that every minor of a member of $\mathscr{I}$ belongs to $\mathscr{I}$. Theorem 8 implies that for every minor ideal $\mathscr{I}$, there exists a finite set of graphs $\mathscr{F}$ such that any graph belongs to $\mathscr{I}$ if and only if it does not contain any member of $\mathscr{F}$ as a minor. In other words, any minor ideal (or any minor closed property) can be characterized by finitely many graphs. Since minor testing is fixed-parameter tractable [48], any minor closed property can be tested in polynomial time.

### *3.1 Directed Minors*

There are different notions of minors for directed graphs, and it is unclear which one is the better than others.

One possible way to define minors for directed graphs is the same as for undirected graphs: just deleting vertices, edges or contracting edges. Robertson and Seymour [49] also showed that the Graph Minor Theorem is true for this notion of minors for directed graphs.

**Theorem 9** ([49]) *Given infinitely many directed graphs $G_1, G_2, ...,$ there exist $1 \leq i < j$ such that $G_i$ can be obtained from $G_j$ by deleting vertices and edges and contracting edges.*

One drawback for allowing contracting any edges in directed graphs is about an issue of connectivity. Observe that contracting edges in undirected graphs does not create new connected components. A natural analog of the connectivity for directed graphs is strong connectivity. A directed graph is *strongly connected* if for any pair of vertices $u, v$, there exist a directed path from $u$ to $v$ and a directed path from $v$ to $u$. A *strong connected component* in a directed graph is a maximal strongly connected subdigraph. Note that contracting edges in directed graph might create new strongly connected components. Hence, people seek notions of minors for directed graphs that preserve strong connectivity. In this subsection we discuss two such notions.

The first one is called the butterfly minor. A directed graph $G$ contains another directed graph $H$ as a *butterfly minor* if $H$ can be obtained from a subdigraph of $G$ by repeatedly contracting edges $e$ satisfying the property that either the tail of $e$ has out-degree one, or the head of $e$ has in-degree one. Note that contracting such edges will not create new strongly connected components.

However, the butterfly minor relation is not a well-quasi-ordering on directed graphs. For any positive integer $k$, let $G_k$ be the directed graph obtained from a cycle of length $2k$ by orienting the edges clockwise or counterclockwise alternately. Hence every vertex of $G_k$ has either in-degree 0 and out-degree 2, or in-degree 2 and out degree 0. So no edge of $G_k$ can be contracted according to the requirement for butterfly minors. Therefore, $\{G_k : k \geq 1\}$ is an antichain with respect to the butterfly minor relation.

Another antichain with respect to the butterfly minor relation is as follows. For any positive integer $k$, let $G_k$ be the directed graph obtained by a path of length $2k$ by orienting edges alternately such that the ends of the path have in-degree 0, and attaching two leaves to each end of the original path and direct the edges such that the ends of the original paths have out-degree 3. Similarly as the previous example, no edge in $G_k$ can be contracted. Therefore, $\{G_k : k \geq 1\}$ is an antichain with respect to the butterfly minor relation.

Each of these two antichains contains arbitrarily long paths with edges oriented alternately. Chudnovsky, Muzi, Oum, Seymour and Wollan (see [39]) proved that such long alternating paths are the only obstructions for butterfly minor ideals being well-quasi-ordered by the butterfly minor relation.

A set of directed graphs $\mathscr{I}$ is called a *butterfly minor ideal* if any butterfly minor of any member of $\mathscr{I}$ belongs to $\mathscr{I}$. An *alternating path* of length $k$ is a directed graph that is obtained from a path of length $k$ by orienting edges such that no directed subpath has length two.

**Theorem 10** ([39]) *Let $\mathscr{I}$ be a butterfly minor ideal. If there exists a positive integer $k$ such that $\mathscr{I}$ does not contain any alternating path of length $k$, then $\mathscr{I}$ is well-quasi-ordered by the butterfly minor relation.*

Now we discuss another notion of minors for directed graphs. An equivalent way to define minors for undirected graphs is by contracting connected subgraphs instead of contracting edges. Here we consider such an analog for directed graphs. More precisely, this containment allows vertex-deletions, edge-deletions and contracting directed cycles. Note that as contracting special edges for butterfly minors, contracting directed cycles does not create new strongly connected components, either. We are not aware of any formal term in the literature describing this type of minor containment besides of simply calling it "minors". But to avoid confusion, we do not call it minors in this paper.

Note that this new containment is incomparable with the butterfly minor relation. There exist directed graphs $G_1, G_2, H$ such that $G_1$ contains $H$ as a butterfly minor, but $H$ cannot be obtained from a subdigraph of $G_1$ by contracting directed cycles, and $G_2$ does not contain $H$ as a butterfly minor, but $H$ can be obtained from a subdigraph of $G_2$ by contracting directed cycles.

Clearly, directed graphs are not well-quasi-ordered by this containment relation since the set of directed cycles is an antichain with respect to this containment. But Kim and Seymour [22] proved that the set of semi-complete directed graphs are well-quasi-ordered by this containment. A directed graph $D$ is *semi-complete* if $E(D)$ is a set of ordered pairs of distinct vertices, and for any distinct vertices $u, v$ of $D$, at least one of $(u, v)$ and $(v, u)$ belongs to $E(D)$.

**Theorem 11** ([22]) *If $G_1, G_2, \ldots$ are semi-complete directed graphs, then there exist $1 \leq i < j$ such that $G_i$ can be obtained from a subdigraph of $G_j$ by repeatedly contracting directed cycles.*

## 3.2 Induced Minors

In this subsection we consider minors where edge-deletions are not allowed. This notion is a combination of the minor relation and the induced subgraph relation. We remark that most of the statements in this subsection address simple graphs. One reason is that for any graph $H$, the set $\{H_i : i \geq 1\}$ is an infinite antichain with respect to the induced minor relation or the induced subgraph relation, where $H_i$ is the graph obtained from $H$ by duplicating each edge $i$ times. Hence, to keep graphs simple, we have to delete all resulting loops and parallel edges when we contract an edge.

Formally, we say a simple graph $G$ contains another simple graph $H$ as an *induced minor* if $H$ can be obtained from $G$ by repeatedly deleting vertices, contracting edges, and deleting resulting loops and parallel edges.

Several infinite antichains with respect to the induced minor relation are known in the literature.

**Theorem 12** *The following sets are antichains with respect to the induced minor relation.*

1. *Reference [53] The set of "alternating double wheels".*
2. *Reference [38] A specific set of simple graphs of maximum degree at most eight with no $K_5^-$ minor.*
3. *Reference [15] A specific set of interval graphs.*
4. *Reference [5] The set of anti-holes with length at least six.*
5. *Reference [31] A specific set of simple graphs that do not contain $W_4$ or $K_5^-$ as an induced minor.*

An *alternating double wheel* is the simple graph obtained from a cycle $v_1 v_2 ... v_{2k} v_1$ of even length with $k \geq 6$ by adding two non-adjacent vertices $x$, $y$ and adding the edges $\{x, v_{2i}\}$, $\{y, v_{2i-1}\}$ for $i \in [k]$. Recall that $K_n^-$ denotes the graph obtained from $K_n$ by deleting an edge, and $W_4$ is the simple graph obtained from the cycle of length four by adding a new vertex adjacent to all other vertices. A graph is an *interval graph* if it is the intersection graph of intervals of $\mathbb{R}$. A graph is an *anti-hole* of length $k$ if it is the complement of a cycle of length $k$.

Since simple graphs are not well-quasi-ordered by the induced minor relation, questions about graphs in more restricted sets were proposed. Thomas [53] first proved the following.

**Theorem 13** ([53]) *The set of simple series-parallel graphs are well-quasi-ordered by the induced minor relation.*

A graph is *series-parallel* if it does not contain $K_4$ as a minor. Note that a simple graph contains $K_4$ as a minor if and only if it contains $K_4$ as an induced minor. Thomas [53] also asked whether Theorem 13 can be generalized to the set of simple graphs with no $K_5^-$ minor. Matoušek et al. [38] and Lewchalermvongs [31] provided negative answers of this question as indicated in Statements 2 and 5 of Theorem 12.

Even though Theorem 13 cannot be generalized to graphs with no $K_5^-$ minor, people keep looking for specific classes of simple graphs that are well-quasi-ordered by the induced minor relation.

For any set $\mathscr{F}$ of graphs, define $\text{Forb}^s_{im}(\mathscr{F})$ to be the set of simple graphs that do not contain any member of $\mathscr{F}$ as an induced minor. When the set $\mathscr{F}$ consists of only one graph, say $H$, we write $\text{Forb}^s_{im}(\mathscr{F})$ as $\text{Forb}^s_{im}(H)$. Well-quasi-ordered $\text{Forb}^s_{im}(\mathscr{F})$ are characterized by Błasiok, Kamiński, Raymond and Trunck, when $|\mathscr{F}| = 1$.

**Theorem 14** ([5]) *Let $H$ be a simple graph. Then* $\text{Forb}^s_{im}(H)$ *is well-quasi-ordered by the induced minor relation if and only if $H$ is $\hat{K}_4$ or $W_4^-$.*

Here $\hat{K}_4$ is the simple graph obtained from $K_4$ by adding a new vertex $v$ adjacent to exactly two vertices of $K_4$; $W_4^-$ is the graph obtained from $W_4$ by deleting an edge not incident with the vertex of degree four.

Furthermore, Lewchalermvongs [31] characterizes all induced minor ideals $\mathscr{I}$ that are contained in $\text{Forb}^s_{im}(\{W_4, K_5^-\})$ and well-quasi-ordered by the induced minor relation. Formal descriptions of this result are involved, so we omit the details.

Another result about induced minors was proved by Ding [15] as follows. (A graph is *chordal* if it does not contain any cycle of length at least four as an induced subgraph.)

**Theorem 15** ([15]) *If $t$ is a positive integer, then simple chordal graphs with no clique of size $t + 1$ are well-quasi-ordered by the induced minor relation.*

Other classes of simple graphs are also concerned. Lozin and Mayhill [35] proposed the following conjecture. (A *unit interval graph* is an intersection graph of a collection of intervals of $\mathbb{R}$ of length one; a *permutation graph* is a simple graph such that its vertex-set is $\{v_1, v_2, ..., v_n\}$ for some positive integer $n$, and there exists a permutation $\sigma$ on $[n]$ such that $v_i$ is adjacent to $v_j$ if and only if $(i - j)(\sigma(i) - \sigma(j)) < 0$.)

**Conjecture 1** ([35]) Unit interval graphs and bipartite permutation graphs are well-quasi-ordered by the induced minor relation.

Note that the set of interval graphs is not well-quasi-ordered by the induced minor relation by Statement 3 in Theorem 12.

Another positive result about induced minors is proved by Fellows, Hermelin and Rosamond [17].

**Theorem 16** ([17]) *If $k$ is a positive integer, then the set of simple graphs with no cycle of length greater than $k$ is well-quasi-ordered by the induced minor relation.*

In the rest of the subsection, we consider containment relations that only allow edge-contractions. Clearly, graphs with different number of components form an antichain if only edge-contractions are allowed. Hence one should limit the number of components when considering this containment.

We say that a simple graph (or loopless graph, respectively) $G$ contains another simple graph (or loopless graph, respectively) $H$ as a *simple-contraction* (or *loopless-contraction*, respectively) if $H$ can be obtained from $G$ by contracting edges and deleting resulting loops and parallel edges (or deleting resulting loops, respectively). For every positive integer $k$, define $\Theta_k$ to be the 2-vertex loopless graph with $k$ parallel edges. It is easy to see that $\{\Theta_k : k \geq 1\}$ is an antichain with respect to loopless-contraction.

For a positive integer $p$ and a family of graphs $\mathscr{F}$, let $\text{Forb}^{s,p}_{sc}(\mathscr{F})$ (or $\text{Forb}^{\ell,p}_{lc}(\mathscr{F})$, respectively) be the set of simple (or loopless, respectively) graphs with at most $p$ components containing no member of $\mathscr{F}$ as a simple-contraction (or loopless-contraction, respectively). The following are proved by Kamiński et al. [20, 21].

**Theorem 17** *Let $k, p$ be positive integers.*

1. *Reference [21] Let $H$ be a simple graph. Then* $\mathrm{Forb}_{sc}^{s,1}(\{H\})$ *is well-quasi-ordered by the simple-contraction relation if and only if $K_4^-$ contains $H$ as a simple-contraction.*
2. *Reference [20]* $\mathrm{Forb}_{lc}^{\ell,p}(\{\Theta_i : i \geq k\})$ *is well-quasi-ordered by the loopless-contraction relation.*

## 3.3 Vertex-Minors and Pivot-Minors

Let $G$ be a simple graph. The simple graph obtained from $G$ by applying *local complementation* on a vertex $v$ of $G$ is the simple graph $G * v$ with vertex-set $V(G)$ and two distinct vertices $x, y$ are adjacent in $G * v$ if and only if either $v$ is adjacent in $G$ to both $x, y$ and $\{x, y\} \notin E(G)$, or at least one of $x, y$ is not adjacent in $G$ to $v$ and $\{x, y\} \in E(G)$. A simple graph $H$ is a *vertex-minor* of $G$ if $H$ can be obtained from $G$ by repeatedly deleting vertices and applying local complementations.

It is straightforward to verify that for any edge $\{x, y\}$ of a simple graph $G$, $G * x * y * x = G * y * x * y$. The simple graph, denoted by $G \wedge \{u, v\}$, obtained from $G$ by applying *pivoting* an edge $\{u, v\}$ of $G$ is the graph $G * u * v * u$. A simple graph $H$ is a *pivot-minor* of $G$ if $H$ can be obtained from $G$ by repeatedly deleting vertices and applying pivotings.

Clearly, if $H$ is a pivot-minor of $G$, then $H$ is a vertex-minor of $G$. Oum [44] asks whether the pivot-minor relation is a well-quasi-ordering on simple graphs or not.

*Question 1* Are simple graphs well-quasi-ordered by the pivot-minor relation?

Proving a positive answer of Question 1 is expected to be very difficult, since even a positive answer of this question on bipartite graphs implies the Graph Minor Theorem.

Now we discuss the relationship between pivot-minors and minors. Note that if $G$ is a graph and $T$ is a spanning forest in $G$, then for every edge $e \in E(G) - E(T)$, there uniquely exists a cycle in $T + e$ containing $e$. This cycle is called the *fundamental cycle for $e$ with respect to $T$*. For a graph $G$ and a spanning forest $T$ of $G$, the *fundamental graph of $G$ with respect to $T$*, denoted by $F(G; T)$, is a simple bipartite graph with (ordered) bipartition $(E(T), E(G) - E(T))$ such that for any $e \in E(T)$ and $f \in E(G) - E(T)$, $e$ is adjacent to $f$ in $F(G; T)$ if and only if $e$ belongs to the fundamental cycle for $f$ with respect to $T$.

Deleting vertices from $F(G; T)$ corresponds to deleting or contracting edges of $G$. Let $e \in V(F(G; T))$. It is straightforward to see that if $e \in E(T)$, then deleting $e$ from $F(G; T)$ results in the graph $F(G/e; T/e)$, where $G/e$ and $T/e$ denote the graphs obtained from $G$ and $T$ by contracting $e$, respectively; if $e \notin E(T)$, then deleting $e$ from $F(G; T)$ results in the graph $F(G - e; T)$.

Pivoting an edge in $F(G; T)$ corresponds to switching to a new spanning forest. Let $\{e, f\} \in E(F(G; T))$, where $e \in E(T)$ and $f \in E(G) - E(T)$. Then $F(G; T) \wedge \{e, f\} = F(G; (T - e) + f)$.

Therefore, if $G_1, G_2$ are graphs and $T_1, T_2$ are spanning forests in $G_1, G_2$, respectively, such that $F(G_1; T_1)$ is a pivot-minor of $F(G_2; T_2)$, then $G_1$ is a minor of $G_2$. This shows that if simple bipartite graphs are well-quasi-ordered by the pivot-minor relation, then graphs are well-quasi-ordered by the minor relation, which is what the Graph Minor Theorem states.

Oum [43] proved that Question 1 has a positive answer for simple graphs with bounded "rank-width". Rank-width is a graph parameter that does not increase by taking vertex-minors or pivot-minors, which is an analog of the relationship between treewidth and the minor containment. Oum's theorem can be viewed as a step toward a potential answer of Question 1 as proving the bounded treewidth case serves the first step of the proofs of the Graph Minor Theorem and Robertson's conjecture. We omit the formal definition of rank-width in this article.

**Theorem 18** ([43]) *For every positive integer k, simple graphs with rank-width at most k are well-quasi-ordered by the pivot-minor relation.*

One can ask whether a weakening of Question 1 for the vertex-minor relation holds.

**Conjecture 2** Simple graphs are well-quasi-ordered by the vertex-minor relation.

The vertex-minor relation is a weakening of the induced topological minor relation. We say that a simple graph $H$ is an *induced topological minor* of another simple graph $G$ if $H$ can be obtained from $G$ by repeatedly deleting vertices and suppressing vertices of degree two not contained in triangles. Note that we only allow suppressing vertices not contained in triangles since we focus on simple graphs here. Furthermore, one can also define induced topological minors that allow suppressing any vertex of degree two and deleting parallel edges. It is equivalent with the earlier definition since suppressing a vertex of degree two contained in a triangle and deleting resulting parallel edges is equivalent with the operation that simply deletes this degree two vertex. It is easy to see that the simple graph obtained from a simple graph $G$ by suppressing a vertex $v$ of degree two not contained in a triangle can be obtained from $G * v$ by deleting $v$. Therefore, if a simple graph $G$ contains another simple graph $H$ as an induced topological minor, then $G$ contains $H$ as a vertex-minor. It is easy to see that the topological minor relation and the induced topological minor relation are the same for trees. Hence Theorem 1 indeed shows that trees are well-quasi-ordered by the induced topological minor relation.

Conjecture 2 is known to be true for circle graphs. A simple graph is a *circle graph* if it is the intersection graph of a set of chords of a circle. Bouchet [6] proved that the following theorem follows from Theorem 20 on 4-regular graphs.

**Theorem 19** ([6]) *Circle graphs are well-quasi-ordered by the vertex-minor relation.*

# 4 Immersions

Immersions are graph containments that are closely related to the topological minor relation. A *weak immersion* of a graph $H$ in another graph $G$ is a function $\pi$ with domain $V(H) \cup E(H)$ such that the following hold.

- $\pi$ maps vertices of $H$ to vertices of $G$ injectively.
- For each non-loop $e$ of $H$ with ends $x, y$, $\pi(e)$ is a path in $G$ with ends $\pi(x)$ and $\pi(y)$.
- For each loop $e$ of $H$ with end $v$, $\pi(e)$ is a cycle in $G$ containing $\pi(v)$.
- If $e_1, e_2$ are distinct edges of $H$, then $E(\pi(e_1) \cap \pi(e_2)) = \emptyset$.

A *strong immersion* of $H$ in $G$ is a weak immersion $\pi$ of $H$ in $G$ such that for every $e \in E(H)$ and vertex $v$ of $H$ not incident with $e$, $\pi(v) \notin V(\pi(e))$. We say that $G$ contains $H$ as a *weak immersion* (or *strong immersion*, respectively) if there exists a weak (or strong, respectively) immersion of $H$ in $G$.

Clearly, any homeomorphic embedding from $H$ into $G$ is a strong immersion of $H$ in $G$, and every strong immersion of $H$ in $G$ is a weak immersion of $H$ in $G$. Hence if $G$ contains $H$ as a topological minor, then $G$ contains $H$ as a strong immersion and a weak immersion. However, the immersion relations and minor relation are incomparable. There exist graphs $G, H$ such that $G$ contains $H$ as a minor, but $G$ does not contain $H$ as a weak immersion; there exist graphs $G', H'$ such that $G'$ contains $H'$ as a strong immersion, but $G'$ does not contain $H'$ as a minor. It is worthwhile mentioning that the minor relation, topological minor relation and weak and strong immersion relations are equivalent for subcubic graphs.

Nash-Williams in the 1960's conjectured that the weak immersion relation [40] and the strong immersion relation [42] are well-quasi-ordering. The weak immersion conjecture was proved by Robertson and Seymour [50] in the currently last paper in their Graph Minors Series. Indeed, they proved that it is true even when graphs are labelled.

**Theorem 20** ([50]) *Let $Q = (V(Q), \preceq_Q)$ be a well-quasi-ordering. For each positive integer $i$, let $G_i$ be a graph and $\phi_i : V(G_i) \to V(Q)$ be a function. Then there exist $1 \le j < j'$ and a weak immersion $\pi$ of $G_j$ in $G_{j'}$ such that $\phi_j(v) \preceq_Q \phi_{j'}(\pi(v))$ for every $v \in V(G_j)$.*

The strong immersion conjecture remains open. Robertson and Seymour believe that they had a proof of the strong immersion conjecture at one time, but even if it was correct, it was very complicated, and it is unlikely that they will write it down (see [50]).

**Conjecture 3** ([42]) Graphs are well-quasi-ordered by the strong immersion relation.

It is not hard to prove Conjecture 3 for graphs with bounded maximum degree by using Theorem 20.

**Theorem 21** *Let $k$ be a nonnegative integer, and let $Q = (V(Q), \preceq_Q)$ be a well-quasi-ordering. For each positive integer $i$, let $G_i$ be a graph with maximum degree at most $k$, and let $\phi_i : V(G_i) \to V(Q)$. Then there exist $1 \leq j < j'$ and a strong immersion $\pi$ of $G_j$ in $G_{j'}$ such that $\phi_j(v) \preceq_Q \phi_{j'}(\pi(v))$ for every $v \in V(G_j)$.*

*Proof* Define $Q'$ to be the well-quasi-ordering $(V(Q) \times ([k] \cup \{0\}), \preceq)$, where for any $(x_1, y_1), (x_2, y_2) \in V(Q) \times ([k] \cup \{0\})$, $(x_1, y_1) \preceq (x_2, y_2)$ if and only if $x_1 \preceq_Q x_2$ and $y_1 = y_2$. For each $i \geq 1$, define $f_i : V(G_i) \to V(Q) \times ([k] \cup \{0\})$ to be the function such that $f_i(v) = (\phi_i(v), d(v))$ for each $v \in V(G_i)$, where $d(v)$ is the degree of $v$ in $G_i$. By Theorem 20, there exist $1 \leq j < j'$ and a weak immersion $\pi$ of $G_j$ in $G_{j'}$ such that $f_j(v) \preceq f_{j'}(\pi(v))$ for every $v \in V(G'_j)$. In particular, for every $v \in V(G_j)$, the degree of $\pi(v)$ in $G_{j'}$ equals the degree of $v$ in $G_j$. So for each $v \in V(G_j)$, all edges of $G_{j'}$ incident with $\pi(v)$ are contained in $\bigcup \pi(e)$, where the union is over all edges $e$ of $G_j$ incident with $v$. Hence $\pi$ is a strong immersion of $G_j$ in $G_{j'}$. This proves the theorem. □

Andreae [2] made some progress on Conjecture 3.

**Theorem 22** ([2]) *The following classes of simple graphs are well-quasi-ordered by the strong immersion relation.*

1. *Simple graphs that do not contain $K_{2,3}$ as a strong immersion.*
2. *Simple graphs whose blocks are either complete graphs, cycles, or balanced complete bipartite graphs.*

## *4.1 Directed Graphs*

The notion of weak immersion and strong immersion naturally extend to directed graphs. A *weak immersion* of a directed graph $H$ in another directed graph $G$ is a function $\pi$ with domain $V(H) \cup E(H)$ such that the following hold.

- $\pi$ maps vertices of $H$ to vertices of $G$ injectively.
- For each non-loop $e$ of $H$ with head $x$ and tail $y$, $\pi(e)$ is a directed path in $G$ with from $\pi(x)$ to $\pi(y)$.
- For each loop $e$ of $H$ with end $v$, $\pi(e)$ is a directed cycle in $G$ containing $\pi(v)$.
- If $e_1, e_2$ are distinct edges of $H$, then $E(\pi(e_1) \cap \pi(e_2)) = \emptyset$.

A *strong immersion* of $H$ in $G$ is a weak immersion $\pi$ of $H$ in $G$ such that for every $e \in E(H)$ and vertex $v$ of $H$ not incident with $e$, $\pi(v) \notin V(\pi(e))$.

Directed graphs are not well-quasi-ordered by the immersion relations, even for weak immersion. Consider the cycles of length $2k$ with edges oriented clockwise and counterclockwise alternately. It is easy to see that these orientated cycles form an infinite antichain with respect to weak immersion.

But Chudnovsky and Seymour [7] proved that tournaments are well-quasi-ordered by the strong immersion relation. Recall that tournaments are not well-quasi-ordered by the topological minor relation (Theorem 7).

**Theorem 23** ([7]) *Tournaments are well-quasi-ordered by the strong immersion relation.*

## 5 Subgraphs

In this section we discuss the subgraph relation. A graph $H$ is a *subgraph* of another graph $G$ if $H$ can be obtained from $G$ by deleting vertices and edges. A*subgraph embedding* from $H$ into $G$ is an injective function $f : V(H) \cup E(H) \to E(H) \cup E(G)$ such that $f(V(H)) \subseteq V(G)$, $f(E(H)) \subseteq E(G)$, and for any edge $\{x, y\}$ of $H$, $f(e) = \{f(x), f(y)\}$. Clearly, $H$ is a subgraph of $G$ if and only if there exists a subgraph embedding from $H$ into $G$.

The subgraph relation does not well-quasi-order graphs. The set of all cycles is an infinite antichain with respect to the subgraph relation. There is another antichain. For every positive integer $k$, the *fork* of length $k$, denoted by $F_k$, is the simple graph obtained from a path of length $k$ by attaching two leaves to each end of the original path. Clearly, the set of all forks is an infinite antichain with respect to the subgraph relation. This situation is similar with the topological minor case. Indeed, Ding [13] proved an analog of Robertson's conjecture with respect to the subgraph relation.

**Theorem 24** ([13]) *Let $k$ be a positive integer, and let $Q = (V(Q), \preceq)$ be a well-quasi-ordering. For any positive integer $i$, let $G_i$ be a graph that does not contain a path of length $k$ as a subgraph, and let $\phi_i : V(G_i) \to V(Q)$. Then there exist $1 \leq j < j'$ and a subgraph embedding $\phi$ from $G_j$ into $G_{j'}$ such that $\phi_j(v) \preceq \phi_{j'}(\pi(v))$ for every $v \in V(G_j)$.*

Ding's proof of Theorem 24 is nice and short based on the simple fact that every connected graph that does not contain a path of length $k$ as a subgraph can be modified into a graph that does not contain a path of length $k-1$ as a subgraph by deleting at most $k$ vertices. Theorem 24 can also be derived from Theorem 4. Let $G'_i$ be the graph obtained from $G_i$ by subdividing every edge once and then duplicating all edges. Define a new well-quasi-ordering $Q'$ by adding a new element into $Q$ incomparable to all other elements of $Q$. Further label all vertices of $G'_i$ obtained by subdividing edges of $G_i$ by this new element. Then $G'_{j'}$ contains $G'_j$ as a topological minor with respect to the labelling if and only if $G_{j'}$ contains $G_j$ as a subgraph with respect to the labelling. And it is easy to see that if $G_i$ does not contain a path of length $k$ as a subgraph, then $G'_i$ does not contain $R_{2k+2}$ as a topological minor. So Theorem 24 follows from Theorem 4.

A set $\mathscr{I}$ of graphs is a *subgraph ideal* if every subgraph of a member of $\mathscr{I}$ belongs to $\mathscr{I}$. Ding [13] characterized all well-quasi-ordered subgraph ideals of simple graphs.

**Theorem 25** ([13]) *Let $\mathscr{I}$ be a subgraph ideal of simple graphs. Then the following are equivalent.*

1. *$\mathscr{I}$ is well-quasi-ordered by the subgraph relation.*
2. *$\mathscr{I}$ is well-quasi-ordered by the induced subgraph relation.*
3. *There exists a positive integer $k$ such that $\mathscr{I}$ does not contain any cycle or fork of length at least $k$.*

### 5.1 Subdigraphs

Now we discuss the subdigraph relation on directed graphs. As shown in Sect. 3, there exists an infinite antichain of directed graphs with respect to the butterfly minor relation. This antichain is also an antichain with respect to the subdigraph relation. Note that directed graphs in this antichain do not contain a directed path of length at least two. So Theorem 24 does not extend to directed graphs with no long directed paths. But Ding [13] points out that his proof of Theorem 24 can be easily modified to prove that directed graphs whose underlying graphs do not contain a path of length $k$ are well-quasi-ordered by the subdigraph relation.

Recall that Theorem 7 shows that there exists an infinite antichain of tournaments with respect to the topological minor relation. So tournaments are not well-quasi-ordered by the subdigraph relation. More examples of infinite antichains of tournaments were proved by Latka [30]. For any positive integer $n \geq 9$, let $A_n$ be the tournament obtained from the transitive tournament on $[n]$ by reversing the edges in $\{(i, i+1), (1, 3), (n-2, n) : 1 \leq i \leq n-1\}$. For any positive integer $n \geq 4$, let $B_n$ be the tournament with $V(B_n) = \mathbb{Z}/((2n+1)\mathbb{Z})$ such that with $E(B_n) = \{(i, j) : j - i \in \{1, 2, ..., n-1, n+1\}\}$, where the computation is in $\mathbb{Z}/((2n+1)\mathbb{Z})$.

**Theorem 26** ([30]) $\{A_n : n \geq 9\}$ *and* $\{B_n : n \geq 4\}$ *are infinite antichains with respect to the subdigraph relation.*

### 5.2 Induced Subgraphs

A graph $H$ is an *induced subgraph* of $G$ if $H$ can be obtained from $G$ by deleting vertices. It is required to focus on simple graphs only when considering well-quasi-ordering by induced subgraph relation, since for any graph $G$, the set $\{G_i : i \geq 1\}$ is an infinite antichain with respect to the induced subgraph relation, where $G_i$ is obtained from $G$ by duplicating each edge $i$ times. So we only focus on simple graphs in this subsection.

Let $\mathscr{F}$ be a set of graphs. Define $\text{Forb}_s^s(\mathscr{F})$ (and $\text{Forb}_{is}^s(\mathscr{F})$, respectively) to be the set of simple graphs that do not contain any member of $\mathscr{F}$ as a subgraph (and an induced subgraph, respectively). When $\mathscr{F}$ consists of one graph $H$, we write $\text{Forb}_s^s(\{H\})$ and $\text{Forb}_{is}^s(\{H\})$ as $\text{Forb}_s^s(H)$ and $\text{Forb}_{is}^s(H)$, respectively, for short. Theorem 24 can be restated as: $\text{Forb}_s^s(P_n)$ is well-quasi-ordered by the subgraph relation. However, Damaschke [11] showed that $\mathrm{Forb}^s_{is}(P_n)$ is not well-quasi-ordered by the induced subgraph relation for $n \geq 5$, though it is true if $n \leq 4$.

For every positive integer $k$, the *k-sun*, denoted by $S_k$, is the simple graph obtained from a complete graph with vertex-set $\{x_i : 1 \leq i \leq k\}$ by adding $k$ vertices $y_1, y_2, ..., y_k$ such that $y_i$ is adjacent to $x_{i-1}$ and $x_i$ for each $i$ with $1 \leq i \leq k$, where $x_0 = x_k$. Define $2K_2$ to be the graph that consists of a disjoint union of two copies of $K_2$. Clearly, for every $k \geq 4$, $S_k$ does not contain $2K_2$ as an induced subgraph, and hence does not contain $P_5$ as an induced subgraph. Damaschke [11] showed that $\{S_{2k} : k \geq 2\}$ is an antichain with respect to the induced subgraph relation.

**Theorem 27** ([11]) *The following statements are true.*

1. *Let $H$ be a simple graph. Then $\mathrm{Forb}^s_{is}(H)$ is well-quasi-ordered by the induced subgraph relation if and only if $H$ is an induced subgraph of $P_4$.*
2. *$\{S_{2k} : k \geq 2\}$ is an infinite antichain with respect to the induced subgraph relation. In particular, for every $k \geq 5$, $\mathrm{Forb}^s_{is}(P_k)$ is not well-quasi-ordered by the induced subgraph relation.*
3. *$\mathrm{Forb}^s_{is}(\{K_3, P_5\})$ and $\mathrm{Forb}^s_{is}(\{K_3, K_2 \cup 2K_1\})$ are well-quasi-ordered by the induced subgraph relation.*

Answering a question of Damaschke, Ding [13] showed that $\mathrm{Forb}^s_{is}(\{2K_2, C_4, C_5, S_4\})$ is not well-quasi-ordered by the induced subgraph relation. Ding also proved that several other families $\mathscr{F}$ in which $\mathrm{Forb}^s_{is}(\mathscr{F})$ are well-quasi-ordered by the induced subgraph relation. We refer interested readers to [13]. In the same paper, Ding [13] proposed the following conjecture about permutation graphs.

**Conjecture 4** ([13]) For every positive integer $k \geq 5$, permutation graphs that do not contain $P_k$ or the complement of $P_k$ as a induced subgraph are well-quasi-ordered by the induced subgraph relation.

Note that Conjecture 4 concerns special classes of graphs. This special class is actually an induced subgraph ideal. There are more results concerning special classes of graphs. For example, Atminas et al. [4] determined whether permutation graphs in $\mathrm{Forb}^s_{is}(\mathscr{F})$ for some families $\mathscr{F}$ with small size are well-quasi-ordered or not.

Another special class of graphs is the set of $k$-letter graphs introduced by Petkovšek [45]. For a positive integer $k$, a simple graph $G$ is a *k-letter graph* if $V(G)$ can be partitioned into $V_1, V_2, ..., V_p$ for some $p \leq k$, where each $V_i$ is a clique or a stable set, such that there exists a linear ordering $\sigma$ of $V(G)$ such that for each pair of distinct indices $i, j \in [p]$, either every vertex in $V_i$ is adjacent to every vertex in $V_j$, or every vertex in $V_i$ is non-adjacent to every vertex in $V_j$, or for every vertex $x$ in $V_i$, its neighbors in $V_j$ are the veritces $y$ in $V_j$ with $\sigma(x) < \sigma(y)$, or for every vertex $x$ in $V_i$, its neighbors in $V_j$ are the vertices $y$ in $V_j$ with $\sigma(x) > \sigma(y)$.

**Theorem 28** ([45]) *For every positive integer k, the set of k-letter graphs is well-quasi-ordered by the induced subgraph relation.*

Using Theorem 28, Lozin and Mayhill [35] proved results related to unit interval graphs and bipartite permutation graphs. Note that the class of unit interval graphs and the class of bipartite permutation graphs are induced subgraph ideals. Recall that $F_k$ is the fork of length $k$. Every $F_k$ is a bipartite permutation graph, but $\{F_k : k \geq 1\}$ is an antichain with respect to the induced subgraph relation. The graph $F_k^+$ is defined to be the simple graph obtained from $F_k$ by adding an edge to each pair of leaves sharing a common neighbor. Every $F_k^+$ is a unit interval graph, but $\{F_k^+ : k \geq 1\}$ is an antichain with respect to the induced subgraph relation.

**Theorem 29** ([35]) *The following are true.*

1. *Let $\mathscr{I}$ be an induced subgraph ideal of unit interval graphs. Then $\mathscr{I}$ is well-quasi-ordered by the induced subgraph relation if and only if $\mathscr{I}$ contains finitely many members of $\{F_k^+ : k \geq 1\}$.*
2. *Let $\mathscr{I}$ be an induced subgraph ideal of bipartite permutation graphs. Then $\mathscr{I}$ is well-quasi-ordered by the induced subgraph relation if and only if $\mathscr{I}$ contains finitely many members of $\{F_k : k \geq 1\}$.*

Now let us consider $\text{Forb}_{is}^s(\mathscr{F})$ in terms of the size of $\mathscr{F}$. As mentioned in Theorem 27, the family $\mathscr{F}$ with size one in which $\text{Forb}_{is}^s(\mathscr{F})$ is well-quasi-ordered by the induced subgraph relation is characterized in [11]. For families $\mathscr{F}$ with $|\mathscr{F}| \geq 2$, the complete characterization for $\mathscr{F}$ such that $\text{Forb}_{is}^s(\mathscr{F})$ is well-quasi-ordered by the induced subgraph relation is not known. But numerous families with size two were studied. For example, see [3, 4, 23, 24].

Following this direction, people study what the minimal non-well-quasi-ordered sets $\mathscr{S}$ of simple graphs such that $\mathscr{S} = \text{Forb}_{im}^s(\mathscr{F})$ for some family $\mathscr{F}$ of simple graphs with $|\mathscr{F}| \leq k$ are. For every positive integer $k$, we say that a set $\mathscr{S}$ of simple graphs is *$k$-bad* if $\mathscr{S} = \text{Forb}_{is}^s(\mathscr{F})$ for some $|\mathscr{F}| = k$, $\mathscr{S}$ is not well-quasi-ordered by the induced subgraph relation, and $\mathscr{S}$ is minimal among the sets satisfying the previous two properties. Korpelainen and Lozin [23] conjectured that for every positive integer $k$, there are only finitely many $k$-bad sets. Using Theorem 27, Korpelainen et al. [25] showed that it is true when $k = 1$. Korpelainen and Lozin [23] proved the case $k = 2$. However, the case $k \geq 3$ was disproved by Korpelainen et al. [25].

**Theorem 30** *The following are true.*

1. *Reference [25] The 1-bad sets are* $\text{Forb}_{is}^s(C_3)$, $\text{Forb}_{is}^s(C_4)$, $\text{Forb}_{is}^s(C_5)$, $\text{Forb}_{is}^s(3K_1)$ *and* $\text{Forb}_{is}^s(2K_2)$.
2. *Reference [23] There are only finitely many 2-bad sets.*
3. *Reference [25] There are infinitely many $k$-bad sets for any* $k \geq 3$. *In particular, for any positive integer $t$ with $t > k$,* $\text{Forb}_{is}^s(\{K_{1,3}, C_i, C_t : 3 \leq i \leq k\})$ *is a $k$-bad set.*

Whether $\text{Forb}_{is}^s(\mathscr{F})$ is well-quasi-ordered by the induced subgraph relation has been determined for almost all families $\mathscr{F}$ with $|\mathscr{F}| = 2$. A summary can be founded in [9]. The remaining undetermined classes are the following.

*Question 2* ([9]) Let $\mathscr{F} = \{H_1, H_2\}$ for some simple graphs $H_1, H_2$. Determine whether $\text{Forb}_{is}^{s}(\mathscr{F})$ is well-quasi-ordered by the induced subgraph relation for the following cases.

1. $H_1 = K_3$ and $H_2 \in \{P_1 \cup 2P_2, P_1 \cup P_5, P_2 \cup P_4\}$.
2. $H_1 = K_4^-$ and $H_2 \in \{P_1 \cup 2P_2, P_1 \cup P_4\}$.
3. $H_1 = W_4^-$ and $H_2 \in \{P_1 \cup P_4, 2P_2, P_2 \cup P_3, P_5\}$.

"Clique width" is a well-known graph parameter that is used for measuring how "homogenous" its vertices are. So graphs with smaller clique width are less complicated. Moreover, any induced subgraph $H$ of a graph $G$ has clique width no more than $G$. Hence, for any positive integer $k$, the set of simple graphs of clique width at most $k$ is an induced subgraph ideal. However, every cycle has clique width at most four, so the set of simple graphs of bounded clique width is not well-quasi-ordered by the induced subgraph relation, even when the bound is four. On the other hand, intuitively, graphs in any induced subgraph ideal that can be well-quasi-ordered by the induced subgraph relation are expected not to be too "complicated". Daligault et al. [10] asked whether it is true that every induced subgraph ideal containing graphs with arbitrarily large clique width cannot be well-quasi-ordered by the induced subgraph relation. However, Lozin et al. [36] provide a negative answer of this question.

For every positive integer $k$, define $D_k$ to be the simple graph with $V(D_k) = [k]$ and where two vertices $i, j$ are adjacent if and only if either $|i - j| = 1$, or $q(i) = q(j)$, where for any $x \in [k]$, $q(x)$ is the largest number of the form $2^n$ (for some positive integer $n$) dividing $x$.

**Theorem 31** ([36]) *Let $\mathscr{I}$ be the set of simple graphs consisting of $\{D_k : k \geq 1\}$ and all induced subgraphs of $D_k$ for some $k$. Then $\mathscr{I}$ is well-quasi-ordered by the induced subgraph relation, but for every number $n$, there exists $n'$ such that the clique width of $D_{n'}$ is greater than $n$.*

As indicated in [9], the ideal $\mathscr{I}$ mentioned in Theorem 31 cannot be written as $\text{Forb}_{is}^{s}(\mathscr{F})$ for some finite family $\mathscr{F}$. Dabrowski et al. [9] conjecture that the finiteness of $\mathscr{F}$ can ensure a positive answer of the question of Daligault, Rao and Thomassé mentioned above. In fact, the question of Daligault, Rao and Thomassé is motivated by another weaker conjecture of theirs (see Conjecture 9 below), and the finiteness is ensured in the setting of that weaker conjecture.

**Conjecture 5** ([9]) If $\mathscr{F}$ is a finite set of simple graphs, and $\text{Forb}_{is}^{s}(\mathscr{F})$ is well-quasi-ordered by the induced subgraph relation, then there exists a number $N$ such that every graph in $\text{Forb}_{is}^{s}(\mathscr{F})$ has clique width at most $N$.

Conjecture 5 is true when $|\mathscr{F}| = 1$. It follows from the fact that $\{P_4\}$ is the only family $\mathscr{F}$ with size one with $\text{Forb}_{is}^{s}(\mathscr{F})$ well-quasi-ordered, and the fact that every graph in $\text{Forb}_{is}^{s}(P_4)$ has clique width at most three. Almost all cases for $\mathscr{F}$ with $|\mathscr{F}| = 2$ are verified (see [9]), except the following.

*Question 3* ([9]) Let $\mathscr{F} = \{H_1, H_2\}$ for some simple graphs $H_1, H_2$. Determine whether $\mathscr{F}$ satisfies Conjecture 5 or not for the following cases.

1. $H_1 = K_3$ and $H_2 = P_2 \cup P_4$.
2. $H_1 = W_4^-$ and $H_2 = P_2 \cup P_3$.

In addition, some sets that were proved to be well-quasi-ordered by the induced subgraph relation are also well-quasi-ordered even when the vertices are labelled by a well-quasi-ordering [3]. A *induced subgraph embedding* $\pi$ from a graph $H$ into a graph $G$ is a subgraph embedding such that the image of $\pi$ is an induced subgraph of $G$. When the vertices of $G$ and $H$ are labelled by a quasi-ordering $Q$, we say that $G$ contains $H$ as a *$Q$-labelled induced subgraph* if there exists an induced subgraph embedding $\pi$ from $H$ into $G$ such that the label of $v$ is less than or equal to the label of $\pi(v)$ with respect to $Q$, for every $v \in V(H)$. We say that a set of simple graphs is well-quasi-order by the *labelled induced subgraph relation* if for every well-quasi-ordering $Q$ and any infinite sequence $G_1, G_2, \ldots$ of $Q$-labelled graphs in this set, there exist $1 \leq i < j$ such that $G_j$ contains $G_i$ as a $Q$-labelled-induced subgraph. Inspired by the known examples of ideals that are well-quasi-ordered by the labelled induced subgraph relation in the literature, Atminas and Lozin conjectured the following.

**Conjecture 6** ([3]) Let $\mathscr{I}$ be an induced subgraph ideal that is well-quasi-ordered by the induced subgraph relation. Then $\mathscr{I}$ is well-quasi-ordered by the labelled induced subgraph relation if and only if $\mathscr{I} = \text{Forb}_{is}^s(\mathscr{F})$ for some finite set $\mathscr{F}$.

Conjecture 6 implies a long-standing conjecture of Pouzet [46], which we describe as follows.

Let $n$ be a positive integer, and let $Q$ be the quasi-ordering $([n], =)$. Let $G, H$ be simple graphs and let $f_G, f_H$ be functions with $f_G : V(G) \to [n]$ and $f_H : V(H) \to [n]$. We say that $(G, f_G)$ contains $(H, f_H)$ as an *$n$-induced subgraph* if there exists an induced subgraph embedding $\pi$ from $H$ into $G$ such that $f_H(v) = f_G(\pi(v))$ for every $v \in V(H)$. A set $\mathscr{S}$ of simple graphs is *$n$-well-quasi-ordered* if for any infinite sequence of simple graphs $G_1, G_2, \ldots$ in $\mathscr{S}$ and for all functions $f_i : V(G_i) \to [n]$ for all $i \geq 1$, there exist $1 \leq j < j'$ such that $(G_{j'}, f_{j'})$ contains $(G_j, f_j)$ as an $n$-induced subgraph.

Clearly, being 1-well-quasi-ordered is equivalent to being well-quasi-ordered by the induced subgraph relation. But 2-well-quasi-ordering is very different from 1-well-quasi-ordering. One evidence is that any 2-well-quasi-ordered induced subgraph ideal of simple graphs cannot contain arbitrarily long paths, but some 1-well-quasi-ordered induced subgraph ideals can. Another evidence is shown by Daligault, Rao and Thomassé [10], that every 2-well-quasi-ordered induced subgraph ideal of simple graphs can be expressed as $\text{Forb}_{is}^s(\mathscr{F})$ for some finite family $\mathscr{F}$. However, having more than two labels seems not different from simply having two labels. The following is conjectured by Pazout [46] and Fraïssé [18].

**Conjecture 7** ([18, 46]) Let $\mathscr{I}$ be an induced subgraph ideal of simple graphs. Then $\mathscr{I}$ is 2-well-quasi-ordered if and only if $\mathscr{I}$ is $n$-well-quasi-ordered for all positive integers $n$.

We remark that Conjecture 6 implies Conjecture 7. Since any 2-well-quasi-ordered ideal is a 1-well-quasi-ordered induced ideal which is of the form $\mathrm{Forb}^{s}_{is}(\mathscr{F})$ for some finite $\mathscr{F}$, Conjecture 6 implies that it is well-quasi-ordered by the labelled induced subgraph relation, so it is $n$-well-quasi-ordered for all $n$.

When Daligault, Rao and Thomassé [10] tried to solve Conjecture 7, they found a special kind of induced subgraph ideal, denoted by $\mathrm{NLC}_k^{\mathscr{F}}$, in which 1-well-quasi-ordering is equivalent with $n$-well-quasi-ordering for any $n$. Roughly speaking, given a positive integer $k$ and a family of functions $\mathscr{F}$ from $[k]$ to $[k]$, the class $\mathrm{NLC}_k^{\mathscr{F}}$ consists of the simple graphs that can be generated by using $k$ symbols and relabelling functions in $\mathscr{F}$. When $\mathscr{F}$ is the family that consists of all functions from $[k]$ to $[k]$, any graph in $\mathrm{NLC}_k^{\mathscr{F}}$ has "NLC-width" at most $k$. The NLC-width is equivalent with the clique width in terms of boundedness. Namely, a class of graphs has bounded NLC-width if and only if it has bounded clique width. We refer readers to [10] for formal definitions of $\mathrm{NLC}_k^{\mathscr{F}}$ and the NLC-width.

**Theorem 32** ([10]) *Let $k$ be a positive integer and let $\mathscr{F}$ be a family of functions from $[k]$ to $[k]$. Then the following are equivalent.*

1. *For any $f, g \in \mathscr{F}$, either the image of $f \circ g$ equals the image of $f$, or the image of $g \circ f$ equals the image of $g$.*
2. *$\mathrm{NLC}_k^{\mathscr{F}}$ is well-quasi-ordered.*
3. *$\mathrm{NLC}_k^{\mathscr{F}}$ is $n$-well-quasi-ordered for all positive integers $n$.*
4. *There exists $M$ such that $P_M \notin \mathrm{NLC}_k^{\mathscr{F}}$.*

Daligault et al. [10] proposed the following conjecture, which implies Conjecture 7 by using Theorem 32.

**Conjecture 8** ([10]) If $\mathscr{I}$ is a 2-well-quasi-ordered induced subgraph ideal of simple graphs, then there exist a positive integer $k$ and a family of functions from $[k]$ to $[k]$ such that $\mathscr{I} \subseteq \mathrm{NLC}_k^{\mathscr{F}}$ and $\mathrm{NLC}_k^{\mathscr{F}}$ is $n$-well-quasi-ordered for every positive integer $n$.

As a potential step to prove Conjecture 8, Daligault, Rao and Thomassé proposed a weaker conjecture in which the restriction for the relabelling functions is not concerned. (Recall that having bounded NLC-width is equivalent with having bounded clique width.)

**Conjecture 9** ([10]) Let $\mathscr{I}$ be an induced subgraph ideal of simple graphs. If $\mathscr{I}$ is 2-well-quasi-ordered, then there exists $M$ such that every graph in $\mathscr{I}$ has clique width at most $M$.

Recall that any 2-well-quasi-ordered induced subgraph ideal can be written as $\mathrm{Forb}^{s}_{is}(\mathscr{F})$ for some finite set of simple graphs $\mathscr{F}$. Hence Conjecture 5 implies Conjecture 9.

## 5.3 *Rao-Containments*

Recall that the induced subgraph relation does not well-quasi-order simple graphs. Rao proposed a way to tweak this relation to be possibly a well-quasi-ordering.

Let $n$ be a positive integer. We say a finite sequence $(a_1, a_2, ..., a_n)$ over nonnegative integers is *graphic* if there exists a simple graph $G$ with $V(G) = [n]$ such that for each $i \in [n]$, the degree of $i$ in $G$ equals $a_i$. We call such a simple graph $G$ a *realization* of $(a_1, a_2, ..., a_n)$. Rao [47] proposed the following conjecture.

**Conjecture 10** ([47]) Given infinitely many graphic sequences $s_1, s_2, ...$, there exist $1 \leq j < j'$ such that some realization of $s_{j'}$ contains some realization of $s_j$ as an induced subgraph.

Conjecture 10 was completely solved by Chudnovsky and Seymour [8] in a stronger sense (Theorem 33 below). Their proof is complicated. Altomare [1] and Sivaraman [51] gave short proofs of Conjecture 10 when there exists a number $M$ such that every entry of every sequence is at most $M$.

We say that a simple graph $G$ is *degree equivalent* with another simple graph $G'$ if $V(G) = V(G')$ and for every vertex, its degree in $G$ equals its degree in $G'$. We say that a simple graph $G$ *Rao-contains* another simple graph $H$ if $H$ is an induced subgraph of a simple graph $G'$ that is degree equivalent to $G$. Chudnovsky and Seymour [8] proved the following.

**Theorem 33** ([8]) *If* $G_1, G_2, ...$ *are simple graphs, then there exist* $1 \leq j < j'$ *such that* $G_{j'}$ *Rao-contains* $G_j$.

It is clear that Theorem 33 implies that Conjecture 10. The concept of Rao containment can be extended to directed graphs. In fact, this extension to directed graphs plays an important role in the proof of Theorem 33. In the proof of Theorem 33, Chudnovsky and Seymour [8] reduced the problem to split graphs, and then further reduced the problem to "complete bipartite directed graphs" with respect to the directed version of Rao-containment.

We say that two directed graphs $G$ and $G'$ are *degree-equivalent* if their underlying graphs are the same, and every vertex has the same out-degree in $G$ and in $G'$. A directed graph $G$ *switching-contains* another directed graph $H$ if there exists a directed graph $G'$ degree-equivalent to $G$, and $H$ is isomorphic to an induced subdigraph of $G'$.

The switching-containment is not a well-quasi-ordering on directed graphs. For example, the set of directed cycles is an infinite antichain with respect to the switching-containment relation. However, Chudnovsky and Seymour [7, 8] proved that switching-containment well-quasi-orders tournaments. It follows from Theorem 23 and the observation that if a tournament $G$ contains another tournament $H$ as a strong immersion, then $G$ switching-contains $H$. Chunnovsky and Seymour [8] also proved this for the directed graphs whose underlying graphs are complete bipartite graphs, and used this fact to prove Theorem 33.

**Theorem 34** ([8]) *Tournaments and directed graphs whose underlying graphs are complete bipartite graphs are well-quasi-ordered by the switching-containment relation.*

**Acknowledgements** The author thanks Sang-il Oum, Paul Seymour and Robin Thomas for providing some references and comments for an early draft of this article. The author also thanks the anonymous reviewer for his or her careful reading and pointing out many typos.

## References

1. Altomare, C. J. (2012). A semigroup proof of the bounded degree case of S.B. Rao's conjecture on degree sequences and a bipartite analogue. *Journal of Combinatorial Theory Series B*, *102*, 756–759 (2012).
2. Andreae, T. (1986). On well-quasi-ordering-finite graphs by immersion. *Combinatorica*, *6*, 287–298.
3. Atminas, A., & Lozin, V. V. (2015). Labelled induced subgraphs and well-quasi-ordering. *Order*, *32*, 313–328.
4. Atminas, A., Brignall, R., Korpelainen, N., Lozin, V., & Vatter, V. (2015). Well-quasi-order for permutation graphs omitting a path and a clique. *The Electronic Journal of Combinatorics*, *22*(2), 20.
5. Błasiok, J., Kamiński, M., Raymond, J.-F., & Trunck, T., Induced minors and well-quasi-ordering. arXiv:1510.07135.
6. Bouchet, A. (1994). Circle graph obstructions. *Journal of Combinatorial Theory Series B*, *60*, 107–144.
7. Chudnovsky, M., & Seymour, P. (2011). A well-quasi-order for tournaments. *Journal of Combinatorial Theory Series B*, *101*, 47–53.
8. Chudnovsky, M., & Seymour, P. (2014). Rao's degree sequence conjecture. *Journal of Combinatorial Theory Series B*, *105*, 44–92.
9. Dabrowski, K. K., Lozin, V. V., & Paulusma, D., Well-quasi-ordering versus clique-width: new results on bigenic classes, arXiv:1611.03671.
10. Daligault, J., Rao, M., & Thomassé, S. (2010). Well-quasi-order of relabel functions. *Order*, *27*, 301–315.
11. Damaschke, P. (1990). Induced subgraphs and well-quasi-ordering. *Journal of Graph Theory*, *14*, 427–435.
12. Diestel, R. (2006). *Graph theory*, Springer, 3rd Edition.
13. Ding, G. (1992). Subgraphs and well-quasi-ordering. *Journal of Graph Theory*, *16*, 489–502.
14. Ding, G. (1996). Excluding a long double path minor. *Journal of Combinatorial Theory Series B*, *66*, 11–23.
15. Ding, G. (1998). Chordal graphs, interval graphs, and wqo. *Journal of Graph Theory*, *28*, 105–114.
16. Erdős, P., & Pósa, L. (1965). On independent circuits contained in a graph. *Canadian Journal of Mathematics*, *17*, 347–352.
17. Fellows, M. R., Hermelin, D., & Rosamond, F. A. (2012). Well-quasi-orders in subclasses of bounded treewidth graphs and their algorithmic applications. *Algorithmica*, *64*, 3–18.
18. Fraïssé, R. (1986). *Theory of relations*. In *Studies in logic 118*, North Holland (1986).
19. Higman, G. (1952). Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society*, *2*, 326–336.
20. Karmiński, M., Raymond, J. -F., & Trunck, T., *Multgraphs without large bonds are wqo by contraction*, arXiv:1412.2407.
21. Karmiński, M., Raymond, J. -F., & Trunck, T., *Well-quasi-ordering $H$-contraction-free graphs*, arXiv:1602.00733.

22. Kim, I., & Seymour, P. (2015). Tournament minors. *Journal of Combinatorial Theory Series B*, *112*, 138–153.
23. Korpelainen, N., & Lozin, V. (2011). Two forbidden induced subgraphs and well-quasi-ordering. *Discrete Math*, *311*, 1813–1822.
24. Korpelainen, N., & Lozin, V. V. (2011). Bipartite induced subgraphs and well-quasi-ordering. *Journal of Graph Theory*, *67*, 235–249.
25. Korpelainen, N., Lozin, V. V., & Razgon, I. (2013). Boundary properties of well-quasi-ordered sets of graphs. *Order*, *30*, 723–735.
26. Kříž, I., & Sgall, J. (1991). Well-quasiordering depends on the labels. *Acta Scientiarum Mathematicarum*, *55*, 59–65.
27. Kříž, I., & Thomas, R. (1990). On well-quasi-ordering finite structures with labels. *Graphs and Combinatorics*, *6*, 41–49.
28. Kruskal, J. B. (1960). Well-quasi-ordering, the tree theorem, and Vászonyi's conjecture. *Transactions of the American Mathematical Society*, *95*, 210–225.
29. Kruskal, J. B. (1972). The theory of well-quasi-ordering: a frequently discovered concept. *Journal of Combinatorial Theory, Series A*, *13*, 297–305.
30. Latka, B. J. (1994). Finitely constrained classes of homogeneous directed graphs. *The Journal of Symbolic Logic*, *59*, 124–139.
31. Lewchalermvongs, C. (2015). *Well-quasi-ordering by the induced-minor relation*, Ph.D. Dissertation, Louisiana State University and Agricultural and Mechanical College.
32. Liu, C.-H. (2014). *Graph Structures and well-quasi-ordering*, Ph.D. Dissertation, Georgia Institute of Technology.
33. Liu, C.-H., & Thomas, R., *Excluding subdivisions of bounded degree graphs*, arXiv:1407.4428.
34. Liu, C.-H., & Thomas, R., *Robertson's conjecture. I–IV*, in preparation.
35. Lozin, V. V., & Mayhill, C. (2011). Canonical antichains of unit interval and bipartite permutation graphs. *Order*, *28*, 513–522.
36. Lozin, V. V., Razgon, I., & Zamaraev, V. (2016). *Well-quasi-ordering does not imply bounded clique-width*. In Mayr E. (eds) *Graph-Theoretic Concepts in Computer Science. WG 2015. Lecture Notes in Computer Science 9224*, pp. 351–359 (2016).
37. Mader, W. (1972). Wohlquasigeordnete klassen endlicher graphen. *Journal of Combinatorial Theory, Series B*, *12*, 105–122.
38. Matoušek, J., Nešetřtil, J., & Thomas, R. (1988). On polynomial time decidability of induced-minor-closed classes. *Commentationes Mathematicae Universitatis Carolinae*, *29*, 703–710.
39. Muzi, I. (2017). *Paths and topological minors in directed and undireced graphs*, Ph.D. Dissertation, Sapienza Universita Di Roma.
40. Nash-Williams, C. & St. J. A. (1964). *On well-quasi-ordering trees*. In *Theory of Graphs and Its Applications (Proc. Symp. Smolenice, 1963)*, Publ. House Czechoslovak Acad. Sci., 83–84.
41. On well-quasi-ordering finite trees. (1963). Nash-Williams, C. St. J. A. *Proceedings of the Cambridge Philosophical Society*, *59*, 833–835.
42. On well-quasi-ordering infinite trees. (1965). Nash-Williams, C. St. J. A. *Proceedings of the Cambridge Philosophical Society*, *61*, 697–720.
43. Oum, S. (2008). Rank-width and well-quasi-ordering. *SIAM Journal on Discrete Mathematics*, *22*, 666–682.
44. Oum, S., Rank-width: Algorithmic and structural results. *Discrete Applied Mathematics*, https://doi.org/10.1016/j.dam.2016.08.006.
45. Petkovšek, M. (2002). Letter graphs and well-quasi-order by induced subgraphs. *Discrete Math*, *244*, 375–388.
46. Pouzet, M. (1972). Un bel ordre d'abritement et ses rapports avec les bornes d'une multirelation. *C. R. Academy of Sciences, Paris Sér. A-B*, *274*, 1677–1680 (1972).
47. Rao, S. B. (1981). Towards a theory of forcibly hereditary P-graphic sequences. In: S. B. Rao (Ed.), *Combinatorics and Graph Theory, Proc. Int. Conf., Kolkata, India, Feb. 1980*, In *Lecture Notes in Math. 885*, Springer, pp. 441–458 (1981).
48. Robertson, N., Seymour, P. D. (1995). Graph minors. XIII. The disjoint paths problem. *Journal of Combinatorial Theory Series B*, *63*, 65–110 (1995).

49. Robertson, N., & Seymour, P. D.: Graph minors. XX. Wagner's conjecture. *Journal of Combinatorial Theory Series B*, *92* 325–357 (2004).
50. Robertson, N., & Seymour, P. D.: Graph minors. XXIII. The Nash-Williams immersion conjecture. *Journal of Combinatorial Theory, Series B*, *100*, 181–205 (2010).
51. Sivaraman, V. (2013). Two short proofs of the bounded case of S.B. Rao's degree sequence conjecture. *Discrete Mathematics*, *313*, 1500–1501 (2013).
52. Tarkowski, S. (1960). On the comparability of dendrites. *Bulletin of the Polish Academy of Sciences Seried Science Mathematics and Astronomy Physics*, *8*, 39–41.
53. Thomas, R. (1985). Graphs without $K_4$ and well-quasi-ordering. *Journal of Combinatorial Theory Series B*, *38*, 240–247.
54. Wagner, K. (1970). *Graphentheorie, 248/248a* (p. 61). Hochschultaschenbucher, Mannheim: B. J.

# The Reverse Mathematics of wqos and bqos

**Alberto Marcone**

**Abstract** In this paper we survey wqo and bqo theory from the reverse mathematics perspective. We consider both elementary results (such as the equivalence of different definitions of the concepts, and basic closure properties) and more advanced theorems. The classification from the reverse mathematics viewpoint of both kinds of results provides interesting challenges, and we cover also recent advances on some long standing open problems.

**2010 Mathematics Subject Classification:** Primary: 03B30 · Secondary: 03F35 · 06A07

This paper is an update of [39], which was written in 2000 and documented the state of the research about the reverse mathematics of statements dealing with wqos and bqos at the turn of the century. Since then, new work on the subject has been carried out and we describe it here. We however include also the results already covered by [39], attempting to cover exhaustively the topic. We also highlight some open problems in the area.

In Sect. 1 we give a brief introduction to reverse mathematics for the reader whose interest in wqos and bqos originates elsewhere. The readers familiar with this research program can safely skip this section. In Sect. 2 we compare different characterizations of wqos and study their closure under basic operations, such as subset, product and intersection. Here even seemingly trivial properties provide interesting challenges for the reverse mathematician. The study of characterizations and closure under simple

I thank Marta Fiori Carones and Emanuele Frittaion who carefully read an earlier draft of this paper. Research partially supported by PRIN 2012 Grant "Logica, Modelli e Insiemi".

A. Marcone (✉)
Dipartimento di Scienze Matematiche, Informatiche e Fisiche, Università di Udine, 33100 Udine, Italy
e-mail: alberto.marcone@uniud.it



P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_8

operations is repeated in Sect. 3 for bqos: the strength of some statements go all the way up to $\mathsf{ATR}_0$ and apparently simple statements such as "3 is bqo" have escaped classification so far. In Sect. 4 we consider the minimality arguments which are one of the main proof techniques of the subject. Section 5 looks at structural results, such as the theorem by de Jongh and Parikh asserting the existence of a maximal linear extension of a wqo. Section 6 deals with what we might call the major results of wqo and bqo theory, such as Higman's, Kruskal's and Nash-Williams' theorems, the minor graph theorem and Fraïssé's conjecture. We end the paper with a section dealing with results about a topological version of wqos.

## 1 Reverse Mathematics

Reverse mathematics is a wide ranging research program in the foundations of mathematics. The main goal of the program is to give mathematical support to statements such as "Theorem $A$ is stronger than Theorem $B$" or "Theorems $C$ and $D$ are equivalent". If taken literally the first statement does not make sense: since $A$ and $B$ are both true, they are logically equivalent. By the same token, the second statement is trivially true, and thus carries no useful information. However a clarification of these statements is possible by finding out precisely the minimal axioms needed to prove $B$ and showing that they do not suffice to prove $A$, and by showing that these minimal axioms coincide for $C$ and $D$. We are thus interested in proving equivalences between theorems and axioms, yielding equivalences and nonequivalences between different theorems, over a weak base theory.

Although we can label "reverse mathematics" any study of this kind (including the study of different forms of the axiom of choice over the base theory $\mathsf{ZF}$), the term is usually restricted to the setting of subsystems of second order arithmetic. The language $\mathcal{L}_2$ of second order arithmetic has variables for natural numbers and variables for sets of natural numbers, constant symbols 0 and 1, binary function symbols for addition and product of natural numbers, symbols for equality and the order relation on the natural numbers and for membership between a natural number and a set. A model for $\mathcal{L}_2$ consists of a first order part (an interpretation for the natural numbers $\mathfrak{N}$ equipped with $+$, $\cdot$ and $\leq$) and a second order part consisting of a collection of subsets of $\mathfrak{N}$. When the first order part is standard we speak of an $\omega$-model and we can identify the model with the subset of $\mathcal{P}(\omega)$ that constitutes its second order part.

Second order arithmetic is the $\mathcal{L}_2$-theory with classical logic consisting of the axioms stating that the natural numbers are a commutative ordered semiring with identity, the induction scheme for arbitrary formulas, and the comprehension scheme for sets of natural numbers defined by arbitrary formulas.

Weyl [66] and Hilbert and Bernays [23, 24] already noticed in their work on the foundations of mathematics that $\mathcal{L}_2$ is rich enough to express, using appropriate codings, significant parts of mathematical practice, and that many mathematical theorems are provable in (fragments of) second order arithmetic. Actually Weyl used a theory similar to what we now denote by $\mathsf{ACA}_0^+$ (a slight strengthening of $\mathsf{ACA}_0$, to be described below). Recently Dean and Walsh [5] traced the history of subsystems of second order arithmetic leading to [12], where Harvey Friedman started the systematic search for the axioms that are sufficient and necessary to prove theorems of ordinary, not set-theoretic, mathematics. One of Friedman's main early discoveries was that (in his words) "When the theorem is proved from the right axioms, the axioms can be proved from the theorem". Friedman also highlighted the role of set-existence axioms, and this soon led to restricting the induction principles allowed in the various systems. The base system $\mathsf{RCA}_0$ and the now well-known $\mathsf{WKL}_0$, $\mathsf{ACA}_0$, $\mathsf{ATR}_0$, and $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$, were introduced in [13]. Today, most of reverse mathematics research compares the strength of mathematical theorems by establishing equivalences, implications and nonimplications over $\mathsf{RCA}_0$.

To describe $\mathsf{RCA}_0$ and the other systems used in reverse mathematics let us also recall that formulas of $\mathcal{L}_2$ are classified in the usual hierarchies: those with no set quantifiers and only bounded number quantifiers are $\mathbf{\Delta}^0_0$, while counting the number of alternating unbounded number quantifiers we obtain the classification of all arithmetical (= without set quantifiers) formulas as $\mathbf{\Sigma}^0_n$ and $\mathbf{\Pi}^0_n$ formulas (one uses $\mathbf{\Sigma}$ or $\mathbf{\Pi}$ depending on the type of the first quantifier in the formula, existential in the former, universal in the latter). Formulas with set quantifiers in front of an arithmetical formula are classified by counting their alternations as $\mathbf{\Sigma}^1_n$ and $\mathbf{\Pi}^1_n$. A formula is $\mathbf{\Delta}^i_n$ in a given theory if it is equivalent in that theory both to a $\mathbf{\Sigma}^i_n$ formula and to a $\mathbf{\Pi}^i_n$ formula.

In $\mathsf{RCA}_0$ the induction scheme and the comprehension scheme of second order arithmetic are restricted respectively to $\mathbf{\Sigma}^0_1$ and $\mathbf{\Delta}^0_1$ formulas. $\mathsf{RCA}_0$ is strong enough to prove some basic results about many mathematical structures, but too weak for many others. The $\omega$-models of $\mathsf{RCA}_0$ are the Turing ideals: subsets of $\mathcal{P}(\omega)$ closed under join and Turing reducibility. The minimal $\omega$-model of $\mathsf{RCA}_0$ consists of the computable sets and is usually denoted by **REC**.

If a theorem $T$ is expressible in $\mathcal{L}_2$ but unprovable in $\mathsf{RCA}_0$, the reverse mathematician asks the question: what is the weakest axiom we can add to $\mathsf{RCA}_0$ to obtain a theory that proves $T$? In principle, we could expect that this question has a different answer for each $T$, but already Friedman noticed that this is not the case. In fact, most theorems of ordinary mathematics expressible in $\mathcal{L}_2$ are either provable in $\mathsf{RCA}_0$ or equivalent over $\mathsf{RCA}_0$ to one of the following four subsystems of second order arithmetic, listed in order of increasing strength: $\mathsf{WKL}_0$, $\mathsf{ACA}_0$, $\mathsf{ATR}_0$, and $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$. This is witnessed in Steve Simpson's monograph [62] and summarized by the *Big Five* terminology. We thus obtain a neat picture where theorems belonging to quite different areas of mathematics are classified in five levels, roughly corresponding to the mathematical principles used in their proofs. $\mathsf{RCA}_0$ corresponds to "computable mathematics", $\mathsf{WKL}_0$ embodies a compactness principle, $\mathsf{ACA}_0$ is linked to sequential compactness, $\mathsf{ATR}_0$ allows for transfinite arguments, $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$ includes impredicative principles.

To obtain $\mathsf{WKL}_0$ we add to $\mathsf{RCA}_0$ the statement of Weak König's Lemma, i.e., every infinite binary tree has a path, which is essentially the compactness of Cantor space. An equivalent statement, intuitively showing that $\mathsf{WKL}_0$ is stronger than $\mathsf{RCA}_0$

(a rigorous proof needs simple arguments from model theory and computability theory), is $\mathbf{\Sigma}^0_1$-separation: if $\varphi(n)$ and $\psi(n)$ are $\mathbf{\Sigma}^0_1$-formulas such that $\forall n \neg(\varphi(n) \wedge \psi(n))$ then there exists a set $X$ such that $\varphi(n) \implies n \in X$ and $\psi(n) \implies n \notin X$ for all $n$. $\mathsf{WKL}_0$ and $\mathsf{RCA}_0$ have the same consistency strength of Primitive Recursive Arithmetic, and are thus proof-theoretically fairly weak. Nevertheless, $\mathsf{WKL}_0$ proves (and often turns out to be equivalent to) a substantial amount of classical mathematical theorems, including many results about real-valued functions and countable rings and fields, basic Banach space facts, etc. The $\omega$-models of $\mathsf{WKL}_0$ are the Scott ideals, and their intersection consists of the computable sets.

$\mathsf{ACA}_0$ is obtained from $\mathsf{RCA}_0$ by extending the comprehension scheme to all arithmetical formulas. The statements without set variables provable in $\mathsf{ACA}_0$ coincide exactly with the theorems of Peano Arithmetic, so that in particular the consistency strength of the two theories is the same. Within $\mathsf{ACA}_0$ one can develop a fairly extensive theory of continuous functions, using the completeness of the real line as an important tool. $\mathsf{ACA}_0$ proves (and often turns out to be equivalent to) also many basic theorems about countable fields, rings, and vector spaces. For example, $\mathsf{ACA}_0$ is equivalent, over $\mathsf{RCA}_0$, to the Bolzano-Weierstrass theorem on the real line. The $\omega$-models of $\mathsf{ACA}_0$ are the Turing ideals closed under jumps, so that the minimal $\omega$-model of $\mathsf{ACA}_0$ consists of all arithmetical sets.

$\mathsf{ATR}_0$ is the strengthening of $\mathsf{RCA}_0$ (and $\mathsf{ACA}_0$) obtained by allowing to iterate arithmetical comprehension along any well-order. It can be shown [62, Theorem V.5.1] that, over $\mathsf{RCA}_0$, $\mathsf{ATR}_0$ is equivalent to $\mathbf{\Sigma}^1_1$-separation, which is exactly as $\mathbf{\Sigma}^0_1$-separation but with $\mathbf{\Sigma}^1_1$ formulas allowed. This is a theory at the outer limits of predicativism and proves (and often turns out to be equivalent to) many basic statements of descriptive set theory but also some results from advanced algebra, such as Ulm's theorem.

$\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$ is the strongest of the big five systems, and is obtained from $\mathsf{RCA}_0$ by extending the comprehension scheme to $\mathbf{\Pi}^1_1$ formulas. Also this axiom scheme is equivalent to many results, including some from descriptive set theory, Banach space theory and advanced algebra, such as the structure theorem for countable Abelian groups.

In recent years there has been a change in the reverse mathematics main focus: following Seetapun's breakthrough result that Ramsey theorem for pairs is not equivalent to any of the Big Five systems [57], a plethora of statements, mostly in countable combinatorics, have been shown to form a rich and complex web of implications and nonimplications. The first paper featuring complex and non-linear diagrams representing the relationships between statements of second order arithmetics appears to be [26]. Nowadays diagrams of this kind are a common feature of reverse mathematics papers. This leads to the zoo of reverse mathematics, a terminology coined by Damir Dzhafarov when he designed "a program to help organize relations among various mathematical principles, particularly those that fail to be equivalent to any of the big five subsystems of second-order arithmetic". Hirschfeldt's monograph [25] highlights this new focus of the reverse mathematics program.

Many elements of the zoo are connected to Ramsey theorem. By $\mathsf{RT}^k_\ell$ we denote Ramsey theorem for sets of size $k$ and $\ell$ colors: for every coloring $c : [\mathbb{N}]^k \to \ell$ (here $[X]^k$ is the set of all subsets of $X$ with exactly $k$ elements, and $\ell$ is the set $\{0, , \dots, \ell - 1\}$) there exists an infinite homogenous set $H$, i.e., such that for some $i < \ell$ we have $c(s) = i$ for every $s \in [H]^k$. $\mathsf{RT}^k_{<\infty}$ is $\forall \ell\, \mathsf{RT}^k_\ell$. A classic result is that $\mathsf{RT}^k_\ell$ is equivalent to $\mathsf{ACA}_0$ over $\mathsf{RCA}_0$ when $k \geq 3$ and $\ell \geq 2$ (see [62, Sect. III.7]). On the other hand, building on Seetapun's result with the essential new step provided by Liu [36], we now know that $\mathsf{RT}^2_2$ and $\mathsf{RT}^2_\infty$ are both incomparable with $\mathsf{WKL}_0$ (see [25, Sect. 6.2 and Appendix]). For any fixed $\ell$ the infinite pigeonhole principle for $\ell$ colors $\mathsf{RT}^1_\ell$ is provable in $\mathsf{RCA}_0$. On the other hand the full infinite pigeonhole principle $\mathsf{RT}^1_{<\infty}$ is not provable in $\mathsf{RCA}_0$ and not even in $\mathsf{WKL}_0$; in fact it is equivalent over $\mathsf{RCA}_0$ to the principle known as $\mathbf{\Sigma}^0_2$-bounding, which is intermediate in strength between $\mathbf{\Sigma}^0_1$-induction and $\mathbf{\Sigma}^0_2$-induction.

Two of the earliest examples of the zoo phenomenon play a significant role with respect to statements dealing with wqos. Both statements are fairly simple consequences of $\mathsf{RT}^2_2$. $\mathsf{CAC}$ is the statement that any infinite partial order contains either an infinite antichain or an infinite chain, while $\mathsf{ADS}$ asserts that every infinite linear order has either an infinite ascending chain or an infinite descending chain. Hirschfeldt and Shore [26] showed that $\mathsf{RT}^2_2$ is properly stronger than $\mathsf{CAC}$, which in turn implies $\mathsf{ADS}$. They also showed that none of these principles imply $\mathsf{WKL}_0$ over $\mathsf{RCA}_0$. The fact that $\mathsf{CAC}$ is properly stronger than $\mathsf{ADS}$ was first proved by Lerman et al. [35], and then given a simpler proof by Patey [52]. These results support the idea that $\mathsf{RT}^2_2$, in contrast to the big five, is not robust (Montalbán [46] informally defined a theory to be robust "if it is equivalent to small perturbations of itself").

Wqo and bqo theory represents an area of combinatorics which has always interested logicians. From the viewpoint of reverse mathematics, one of the reasons for this interest stems from the fact that some important results about wqos and bqos appear to use axioms that are within the realm of second order arithmetic, yet are much stronger than those necessary to develop other areas of ordinary mathematics (as defined in the introduction of [62]). We will see that results about wqo and bqo belong to both facets of reverse mathematics: some statements fit neatly in the big five picture, while some others provide examples of the zoo.

When dealing with wqo and bqo theory, at first sight the limitations of the expressive power of second-order arithmetic compel us to consider only quasi-orders defined on countable sets. This is actually not a big restriction because a quasi-order is wqo (resp. bqo) if and only if each of its restrictions to a countable subset of its domain is wqo (resp. bqo). The limitation mentioned above must be adhered to when we quantify over the collection of all wqos (or bqos), typically in statements of the form "for every wqo …". However we can also consider specific quasi-orders defined on uncountable sets (such as the powerset of a countable set, the collection of infinite sequences of elements of a countable set, or the set of all countable linear orders); statements about these (with a fixed quasi-order) being wqo or bqo can be expressed in a natural way in second-order arithmetic (see Definition 6.3).

We often use $\leq_{\mathbb{N}}$ for the order relation given by the symbol $\leq$ in the language of second order arithmetic. This notation helps to emphasize when we are comparing elements of a quasi-order via the quasi-order relation and when we are comparing them via the underlying structure of arithmetic. We use this notation when the distinction between these orders is not immediately clear from the context.

As usual in the reverse mathematics literature, whenever we begin a definition or statement with the name of a subsystem of second order arithmetic in parenthesis we mean that the definition is given, or the statement proved, within that subsystem.

## 2 Characterizations and Basic Properties of wqos

**Definition 2.1** ($\mathsf{RCA}_0$) A *quasi-order* is a pair $(Q, \preceq)$ such that $Q$ is a set and $\preceq$ is a transitive reflexive relation on $Q$.

When there is no danger of confusion we assume that $Q$ is always equipped with the quasi-order $\preceq$ and that $\preceq$ is always a quasi-order on the set $Q$. Thus in our statements we often mention only $\preceq$ or only $Q$.

Partial orders are natural examples of quasi-orders: a partial order is a quasi-order which also satisfies antisymmetry. We can transform a quasi-order $Q$ into a partial order using the equivalence relation defined by $x \sim y$ if and only if $x \preceq y$ and $y \preceq x$. The quotient structure $Q/\!\sim$ is naturally equipped with a partial order which can be formed using $\Delta^0_1$ comprehension in $\mathsf{RCA}_0$ (it suffices to identify an equivalence class with its least member with respect to $\leq_{\mathbb{N}}$).

Much of the standard terminology and notation for partial orders is used also when dealing with quasi-orders. For example, we write $x \perp y$ to indicate that $x$ and $y$ are incomparable under $\preceq$ and we write $x \prec y$ if $x \preceq y$ and $y \npreceq x$.

**Definition 2.2** ($\mathsf{RCA}_0$) A set $A \subseteq Q$ is an *antichain* if $x \perp y$ for all $x \neq y \in A$. A set $C \subseteq Q$ is a *chain* if $x \preceq y$ or $y \preceq x$ for all $x, y \in C$.

A set $I \subseteq Q$ is an *initial interval* if $y \in I$ whenever $y \preceq x$ for some $x \in I$. The definition of *final interval* is symmetric, with $x \preceq y$ for some $x \in I$.

**Definition 2.3** ($\mathsf{RCA}_0$) A quasi-order $(Q, \preceq)$ is *linear* if $Q$ is a chain.

If $\preceq$ is a quasi-order on $Q$ and $\preceq_L$ is a linear quasi-order on $Q$, then we say $\preceq_L$ is a *linear extension* of $\preceq$ if for all $x, y \in Q$, $x \preceq y$ implies $x \preceq_L y$ and $x \sim_L y$ implies $x \sim y$.

Notice that (provably in $\mathsf{RCA}_0$) if $Q$ is a linear quasi-order then $Q/\!\sim$ is a linear order. Moreover, if $\preceq_L$ is a linear extension of $\preceq$ then $x \sim y$ if and only if $x \sim_L y$ and therefore the linear extensions of a quasi-order $Q$ correspond exactly to the linear extensions of the partial order $Q/\!\sim$.

We can now give the official definition of wqo within $\mathsf{RCA}_0$.

**Definition 2.4** ($\mathsf{RCA}_0$) Let $\preceq$ be a quasi-order on $Q$. $(Q, \preceq)$ is *wqo* if for every map $f : \mathbb{N} \to Q$ there exist $m <_{\mathbb{N}} n$ such that $f(m) \preceq f(n)$.

**Definition 2.5** ($\mathsf{RCA}_0$) An infinite sequence of elements of $Q$ is a function $f : A \to Q$ where $A \subseteq \mathbb{N}$ is infinite.

$f$ is *ascending* if $f(n) \prec f(m)$ for all $n, m \in A$ with $n <_{\mathbb{N}} m$. Similarly, $f$ is *descending* if $f(m) \prec f(n)$ whenever $n, m \in A$ are such that $n <_{\mathbb{N}} m$.

A *well-order* is a linear quasi-order with no infinite descending sequences.

We say that $f$ is a *good sequence* (with respect to $\preceq$) if there exist $m, n \in A$ such that $m <_{\mathbb{N}} n$ and $f(m) \preceq f(n)$; if this does not happen we say that $f$ is *bad*.

The following characterization of wqo is immediate, and easy to prove within $\mathsf{RCA}_0$ using the existence of the enumeration of the elements of an infinite subset of $\mathbb{N}$ in increasing order:

**Fact 2.6** *($\mathsf{RCA}_0$) Let $(Q, \preceq)$ be a quasi-order. The following are equivalent:*

*(i) $Q$ is wqo;*
*(ii) every sequence of elements of $Q$ is good with respect to $\preceq$.*

Wqos can be characterized by several other statements about quasi-orders. The systematic investigation of the axioms needed to prove the equivalences between these characterizations was started by Cholak et al. [3].

Let us begin with the characterizations which are provable in $\mathsf{RCA}_0$.

**Lemma 2.7** ($\mathsf{RCA}_0$) *Let $(Q, \preceq)$ be a quasi-order. The following are equivalent:*

*(i) $Q$ is wqo;*
*(ii) $Q$ has the finite basis property, i.e., for every $X \subseteq Q$ there exists a finite $F \subseteq X$ such that $\forall x \in X\, \exists y \in F\; y \preceq x$;*
*(iii) there is no infinite sequence of initial segments of $Q$ which is strictly decreasing with respect to inclusion;*
*(iv) there is no infinite sequence of final segments of $Q$ which is strictly increasing with respect to inclusion.*

The equivalence between *(i)* and *(ii)* was already noticed by Simpson (see [61, Lemma 3.2], where the finite basis property is stated in terms of partial orders rather than quasi-orders: full details with the current definition are provided in [39, Lemma 4.8]). The equivalence between *(iii)* and *(iv)* is immediate by taking complements with respect to $Q$. To show that *(i)* implies *(iii)* start from an infinite sequences $\{ I_n : n \in \mathbb{N} \}$ of initial segments of $Q$ which is strictly decreasing with respect to inclusion and for every $n$ let $f(n)$ be the $\leq_{\mathbb{N}}$ minimum element of $I_n \setminus I_{n+1}$: $f$ is a bad sequence. To prove that *(iii)* implies *(i)* let $f$ be a bad sequence with domain $\mathbb{N}$ and set $I_n = \left\{ x \in Q : \forall i \leq n\; f(i) \npreceq x \right\}$: $\{ I_n : n \in \mathbb{N} \}$ is an infinite strictly decreasing sequence of initial segments of $Q$.

We now consider the characterizations of the notion of wqo which turn out to be more interesting from the reverse mathematics viewpoint.

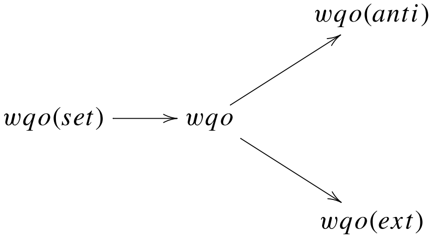
**Definition 2.8** ($\mathsf{RCA}_0$) Let $(Q, \preceq)$ be a quasi-order:

- $Q$ is *wqo(set)* if for every $f : \mathbb{N} \to Q$ there is an infinite set $A$ such that for all $n, m \in A$, $n <_{\mathbb{N}} m \to f(n) \preceq f(m)$;

- $Q$ is *wqo(anti)* if it has no infinite descending sequences and no infinite antichains;
- $Q$ is *wqo(ext)* if every linear extension of $\preceq$ is a well-order.

$\mathsf{RCA}_0$ proves quite easily some implications: every wqo(set) is wqo, and every wqo is both wqo(anti) and wqo(ext). Cholak, Marcone, and Solomon showed that all other implications between these notions are not true in the $\omega$-model **REC**, and hence are not provable within $\mathsf{RCA}_0$.

**Theorem 2.9** *The implications between the notions of wqo, wqo(set), wqo(anti) and wqo(ext) which are provable in* $\mathsf{RCA}_0$ *are exactly the ones in the transitive closure of the diagram:*




*In fact the above diagram depicts the implications which hold in* **REC**, *and thus adding induction axioms to* $\mathsf{RCA}_0$ *yields no other implications.*

To show that every wqo implies wqo(set) fails in **REC**, it suffices to recall a classical construction (due to Denisov and Tennenbaum independently: see [8]) of a computable linear order of order type $\omega + \omega^*$ which does not have any infinite computable ascending or descending sequences.

Similarly, showing that **REC** does not satisfy that every wqo(ext) is wqo means building a computable partial order $(Q, \preceq)$ such that all its computable linear extensions are computably well-ordered (i.e., do not have infinite computable descending sequences) but there is a computable $f : \mathbb{N} \to Q$ such that $f(m) \npreceq f(n)$ for all $m <_{\mathbb{N}} n$. In fact the partial order constructed in [3, Theorem 3.21] using a finite injury construction is such that $f(m) \perp f(n)$ for all $m \neq n$, thus obtaining the stronger result that **REC** does not satisfy that every wqo(ext) is wqo(anti).

To show that wqo(anti) implies wqo does not hold in **REC** one needs to find a computable partial order $(Q, \preceq)$ with no computable infinite antichains and no computable infinite descending sequences but such that there exists a computable $f : \mathbb{N} \to Q$ such that $f(m) \npreceq f(n)$ for all $m <_{\mathbb{N}} n$. The partial order built in [3, Theorem 3.9] has the additional property of having a computable linear extension with a computable infinite descending sequence (see [3, Corollary 3.10]). Hence **REC** does not satisfy that every wqo(anti) is wqo(ext).

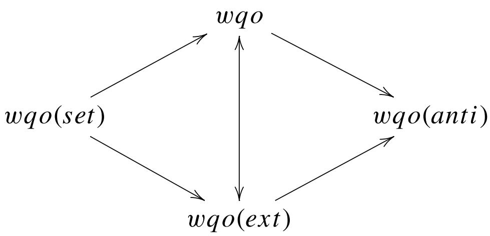
One can improve the latter construction obtaining even more information. In fact, [3, Theorem 3.11] shows that if $(X_i)_{i \in \mathbb{N}}$ is a sequence of uniformly $\Delta^0_2$, uniformly low sets there exists a computable partial order $(Q, \preceq)$, such that for all $i$ no $X_i$-computable function lists an infinite antichain or an infinite descending sequence in $Q$, but there exists a computable $f : \mathbb{N} \to Q$ such that $f(m) \npreceq f(n)$

for all $m <_{\mathbb{N}} n$. Since for an appropriate choice of $(X_i)_{i\in\mathbb{N}}$ we have that the $\omega$-model $\{\, Y : \exists i(Y \leq_T X_i)\,\}$ satisfies $\mathsf{WKL}_0$, we obtain that $\mathsf{WKL}_0$ does not prove that every wqo(anti) is wqo.

Further exploring the provability of the other implications in $\mathsf{WKL}_0$, we notice that it is fairly easy to prove in $\mathsf{RCA}_0$ that the statement that every wqo is wqo(set) implies $\mathsf{RT}^1_{<\infty}$ ([3, Lemma 3.20]), and hence is not provable in $\mathsf{WKL}_0$.

On the other hand, [3, Theorem 3.17] shows that $\mathsf{WKL}_0$ proves (using the fact, equivalent to $\mathsf{WKL}_0$, that every acyclic relation is contained in a partial order) that every wqo(ext) is wqo. Putting the information mentioned above together we obtain the following picture regarding provability in $\mathsf{WKL}_0$.

**Theorem 2.10** *The implications between the notions of wqo, wqo(set), wqo(anti) and wqo(ext) which are provable in* $\mathsf{WKL}_0$ *are exactly the ones in the transitive closure of the diagram:*

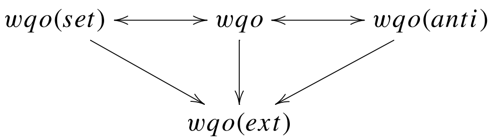



This leads to the following natural question, which has resisted any attempt so far.

**Question 2.11** Consider the statements "every wqo(ext) is wqo" and "every wqo(ext) is wqo(anti)". Are they equivalent to $\mathsf{WKL}_0$ over $\mathsf{RCA}_0$?

On the other hand, the statement "every wqo(anti) is wqo(set)" turns out to be equivalent to $\mathsf{CAC}$ over $\mathsf{RCA}_0$ ([3, Lemma 3.3]). It follows that the statements "every wqo(anti) is wqo" and "every wqo is wqo(set)" are both provable from $\mathsf{CAC}$.

**Theorem 2.12** $\mathsf{RCA}_0 + \mathsf{CAC}$ *proves the implications between the notions of wqo, wqo(set), wqo(anti) and wqo(ext) which are in the transitive closure of the diagram:*




The diagram of Theorem 2.12 is different from the ones of Theorems 2.9 and 2.10 in that it is unknown whether the missing implications can be proved in $\mathsf{RCA}_0 + \mathsf{CAC}$.

**Question 2.13** Does $\mathsf{RCA}_0 + \mathsf{CAC}$ proves "every wqo(ext) is wqo"?

Notice that a positive answer to Question 2.11 implies, since $\mathsf{RCA}_0 + \mathsf{CAC}$ does not prove $\mathsf{WKL}_0$, a negative answer to Question 2.13.

$\mathsf{RCA}_0$ easily proves that all well-orders and all finite quasi-orders are wqo (indeed for the latter fact the finite pigeonhole principle suffices). By Theorem 2.9 the same happens for wqo(anti) and wqo(ext). Regarding wqo(set) we have that, using the appropriate $\mathsf{RT}^1_\ell$, for any specific finite quasi-order $\mathsf{RCA}_0$ proves that the quasi-order is wqo(set). On the other hand, it is not difficult to see that, over $\mathsf{RCA}_0$, "every finite quasi-order is wqo(set)" is equivalent to $\mathsf{RT}^1_{<\infty}$, while "every well-order is wqo(set)" is equivalent to $\mathsf{ADS}$.

Wqos enjoy several basic closure properties. The study of these from the viewpoint of reverse mathematics was started in [3, 39].

We first consider the basic property of closure under taking subsets. The proof of the following lemma is immediate.

**Lemma 2.14** ($\mathsf{RCA}_0$) *Let $\mathcal{P}$ be any of the properties wqo, wqo(anti) or wqo(set). If $(Q, \preceq)$ satisfies $\mathcal{P}$ and $R \subseteq Q$ then the restriction of $\preceq$ to $R$ satisfies $\mathcal{P}$ as well.*

If $\mathcal{P}$ is wqo(ext) then the statement of Lemma 2.14 is slightly more difficult to prove, since the obvious proof of the reversal is based on the following fact: if $(Q, \preceq)$ is a partial order, $R \subseteq Q$, $\preceq_L$ is a linear extension of the restriction of $\preceq$ to $R$, then there exists a linear extension of the whole $\preceq$ which extends also $\preceq_L$. $\mathsf{WKL}_0$ suffices to prove this statement, because we can consider $\preceq \cup \preceq_L$, which is an acyclic relation, extend it to a partial order (here is the step using $\mathsf{WKL}_0$, see [3, Lemma 3.16]), and then to a linear order ($\mathsf{RCA}_0$ suffices for this last step).

**Question 2.15** Does $\mathsf{RCA}_0$ suffice to prove that if $(Q, \preceq)$ is wqo(ext) and $R \subseteq Q$ then the restriction of $\preceq$ to $R$ is also wqo(ext)? Is this implication equivalent to $\mathsf{WKL}_0$?

Let us now consider basic closure operations that involve two quasi-orders.

**Definition 2.16** ($\mathsf{RCA}_0$) If $\preceq_1$ and $\preceq_2$ are quasi-orders on $Q_1$ and $Q_2$ we may assume that $Q_1 \cap Q_2 = \emptyset$ (or replace each $Q_i$ by its isomorphic copy on $Q_i \times \{i\}$). We can define the *sum quasi-order* and the *disjoint union quasi-order* on $Q_1 \cup Q_2$ (denoted by $Q_1 + Q_2$ and by $Q_1 \uplus Q_2$ respectively) by

$$x \preceq_+ y \iff (x \in Q_1 \wedge y \in Q_2) \vee (x, y \in Q_1 \wedge x \preceq_1 y) \vee (x, y \in Q_2 \wedge x \preceq_2 y);$$
$$x \preceq_\uplus y \iff (x, y \in Q_1 \wedge x \preceq_1 y) \vee (x, y \in Q_2 \wedge x \preceq_2 y).$$

The *product quasi-order* on $Q_1 \times Q_2$ is defined by

$$(x_1, x_2) \preceq_\times (y_1, y_2) \iff x_1 \preceq_1 y_1 \wedge x_2 \preceq_2 y_2.$$

Moreover if $\preceq_1$ and $\preceq_2$ are quasi-orders on the same set $Q$ then the *intersection quasi-order* on $Q$ is defined by

$$x \preceq_{\cap} y \iff x \preceq_1 y \wedge x \preceq_2 y.$$

The following lemma follows easily from the provability in $\mathsf{RCA}_0$ of $\mathsf{RT}^1_2$.

**Lemma 2.17** ($\mathsf{RCA}_0$) *Let $\mathcal{P}$ be any of the properties wqo, wqo(ext), wqo(anti) or wqo(set). If $Q_1$ and $Q_2$ satisfy $\mathcal{P}$ then $Q_1 + Q_2$ and $Q_1 \cup Q_2$ satisfy $\mathcal{P}$ with respect to the sum and disjoint union quasi-orders.*

The next lemma was first noticed in [39] for wqos, and then extended to the other notions in [3].

**Lemma 2.18** ($\mathsf{RCA}_0$) *Let $\mathcal{P}$ be any of the properties wqo, wqo(anti) or wqo(set). The following are equivalent:*

*(i) if $Q$ satisfies $\mathcal{P}$ with respect to the quasi-orders $\preceq_1$ and $\preceq_2$ then $Q$ satisfies $\mathcal{P}$ with respect to the intersection quasi-order;*
*(ii) if $Q_1$ and $Q_2$ satisfy $\mathcal{P}$ then $Q_1 \times Q_2$ satisfies $\mathcal{P}$ with respect to the product quasi-order.*

The proof of *(i)* implies *(ii)* is based on the fact that products can be realized as intersections and works for wqo(ext) as well. The proof of *(ii)* implies *(i)* uses the fact that intersections can be viewed as subsets of products, and thus employs Lemma 2.14. In [3] it is claimed that Lemma 2.18 holds also when $\mathcal{P}$ is wqo(ext), but it seems that this might depend on the answer to Question 2.15.

**Question 2.19** Let $\mathcal{P}$ be wqo(ext). Does $\mathsf{RCA}_0$ suffice to prove that *(ii)* of Lemma 2.18 implies *(i)*? Is this implication equivalent to $\mathsf{WKL}_0$?

The following results are from [3].

**Lemma 2.20** ($\mathsf{RCA}_0$) *Let $\mathcal{P}$ be any of the properties wqo, wqo(ext), wqo(anti) or wqo(set).*

- *If $Q$ is wqo(set) with respect to the quasi-orders $\preceq_1$ and $\preceq_2$ then $Q$ satisfies $\mathcal{P}$ with respect to the intersection quasi-order;*
- *if $Q_1$ and $Q_2$ are wqo(set) then $Q_1 \times Q_2$ satisfies $\mathcal{P}$ with respect to the product quasi-order.*

**Theorem 2.21** *Let $\mathcal{P}_1$ be any of the properties wqo, wqo(ext) and wqo(anti). Let $\mathcal{P}_2$ be any of the properties wqo, wqo(set), wqo(ext) and wqo(anti).*

- $\mathsf{WKL}_0$ *does not prove that if $Q$ satisfies $\mathcal{P}_1$ with respect to the quasi-orders $\preceq_1$ and $\preceq_2$ then $Q$ satisfies $\mathcal{P}_2$ with respect to the intersection quasi-order;*
- $\mathsf{WKL}_0$ *does not prove that if $Q_1$ and $Q_2$ satisfy $\mathcal{P}_1$ then $Q_1 \times Q_2$ satisfies $\mathcal{P}_2$ with respect to the product quasi-order.*

All instances of Theorem 2.21 follow easily (using Lemma 2.18 and Theorem 2.10) from Theorem 4.3 of [3]. To state this theorem fix an $\omega$-model $\mathcal{M}$ of $\mathsf{WKL}_0$ which consists of the sets Turing reducible to a member of a sequence of uniformly $\Delta^0_2$, uniformly low sets. The theorem asserts the existence of computable partial orders $\preceq_0$ and $\preceq_1$ which are wqo in $\mathcal{M}$ (i.e., $\mathcal{M}$ contains no bad sequence with respect to either $\preceq_0$ or $\preceq_1$) and such that $\preceq_0 \cap \preceq_1$ is an infinite antichain (so that the intersection is not wqo(anti)). The construction of $\preceq_0$ and $\preceq_1$ is by a finite injury argument.

Theorem 2.12 and Lemma 2.20 imply that $\mathsf{RCA}_0 + \mathsf{CAC}$ proves the closure of wqos under product. On the other hand Frittaion, Marcone, and Shafer pointed out that this statement implies $\mathsf{ADS}$ and asked for a classification. Recently, Henry Towsner [65] gave a typical zoo answer to this question by proving the following theorem.

**Theorem 2.22** $\mathsf{WKL}_0$ *does not prove that the closure of wqos under product implies* $\mathsf{CAC}$*, nor that* $\mathsf{ADS}$ *implies the closure of wqos under product.*

Towsner starts by translating the statement in Ramsey-theoretic terms. Given the coloring $c : [\mathbb{N}]^2 \to \ell$ we say that color $i$ is *transitive* if $c(k_0, k_2) = i$ whenever $c(k_0, k_1) = c(k_1, k_2) = i$ for some $k_1$ satisfying $k_0 < k_1 < k_2$. Hirschfeldt and Shore [26] noticed that $\mathsf{ADS}$ is equivalent to the restriction of $\mathsf{RT}^2_2$ to colorings such that both colors are transitive, while $\mathsf{CAC}$ is equivalent to the restriction of $\mathsf{RT}^2_2$ to colorings with one transitive color. Towsner notices that the closure of wqos under product is equivalent to the following intermediate statement: if $c : [\mathbb{N}]^2 \to 3$ is such that colors 0 and 1 are transitive then there exists an infinite set $H$ such that for some $i < 2$ we have $c(s) \neq i$ for every $s \in [H]^k$ (i.e., $H$ avoids one of the transitive colors). Then he proceeds to construct Scott ideals with the appropriate properties: the first satisfies the above transitive color avoiding statement but not the restriction of $\mathsf{RT}^2_2$ to colorings with one transitive color; the second satisfies for all $\ell$ the restriction of $\mathsf{RT}^2_\ell$ to colorings such that all color are transitive, but fails to satisfy the statement equivalent to the closure of wqos under product.

Special instances of the closure of wqos under product have been studied by Simpson [61].

**Theorem 2.23** ($\mathsf{RCA}_0$) *Let* $\omega$ *denote the order* $(\mathbb{N}, \leq_{\mathbb{N}})$*. Then*

1. *the product of two copies of* $\omega$ *is wqo with respect to the product quasi-order.*
2. *the following are equivalent:*
   - *(i)* $\omega^\omega$ *is well-ordered;*
   - *(ii) for every* $k \in \mathbb{N}$ *the product of* $k$ *copies of* $\omega$ *is wqo with respect to the (obvious generalization of the) product quasi-order.*

Since $\omega^\omega$ is the proof theoretic ordinal of $\mathsf{RCA}_0$, it follows that $\mathsf{RCA}_0$ does not prove the statement *(ii)* above.

Recently Hatzikiriakou and Simpson [21] proved that another statement dealing with wqos is equivalent to the fact that $\omega^\omega$ is well-ordered. A Young diagram is a sequence of natural numbers $\langle m_0, \ldots, m_k \rangle$ such that $m_i \geq m_{i+1}$ and $m_k > 0$. We denote by $\mathcal{D}$ the set of all Young diagrams, and set $\langle m_0, \ldots, m_k \rangle \preceq_{\mathcal{D}} \langle n_0, \ldots, n_h \rangle$ if and only if $k \leq h$ and $m_i \leq n_i$ for all $i \leq k$.

**Theorem 2.24** ($\mathsf{RCA}_0$) *The following are equivalent:*

*(i)* $\omega^\omega$ *is well-ordered;*
*(ii)* $(\mathcal{D}, \preceq_{\mathcal{D}})$ *is wqo.*

Theorems 2.23 and 2.24 are both motivated by the study of results about the non-existence of infinite ascending sequences of ideals in rings.

## 3 Characterizations and Basic Properties of bqos

To give the definition of bqo we need some terminology and notation for sequences and sets (here we follow [38]). All the definitions are given in $\mathsf{RCA}_0$. Let $\mathbb{N}^{<\mathbb{N}}$ be the set of finite sequences of natural numbers. If $s \in \mathbb{N}^{<\mathbb{N}}$ we denote by $\operatorname{lh} s$ its length and, for every $i < \operatorname{lh} s$, by $s(i)$ its $(i+1)$-th element. Then we write this sequence as $s = \langle s(0), \dots, s(\operatorname{lh} s - 1)\rangle$. If $s, t \in \mathbb{N}^{<\mathbb{N}}$ we write $s \sqsubseteq t$ if $s$ is an initial segment of $t$, i.e., if $\operatorname{lh} s \leq \operatorname{lh} t$ and $\forall i < \operatorname{lh} s\, s(i) = t(i)$. We write $s \subseteq t$ if the range of $s$ is a subset of the range of $t$, i.e., if $\forall i < \operatorname{lh} s\, \exists j < \operatorname{lh} t\, s(i) = t(j)$. $s \sqsubset t$ and $s \subset t$ have the obvious meanings. We write $s^\frown t$ for the concatenation of $s$ and $t$, i.e., the sequence $u$ such that $\operatorname{lh} u = \operatorname{lh} s + \operatorname{lh} t$, $u(i) = s(i)$ for every $i < \operatorname{lh} s$, and $u(\operatorname{lh} s + i) = t(i)$ for every $i < \operatorname{lh} t$. These notations are extended to infinite sequences (i.e., functions with domain $\mathbb{N}$) as well.

If $X \subseteq \mathbb{N}$ is infinite we denote by $[X]^{<\mathbb{N}}$ the set of all finite subsets of $X$. We identify an element of $[\mathbb{N}]^{<\mathbb{N}}$ with the unique element of $\mathbb{N}^{<\mathbb{N}}$ which enumerates it in increasing order, so that we can use the notation introduced above. If $k \in \mathbb{N}$, $[X]^k$ is the subset of $[X]^{<\mathbb{N}}$ consisting of the sets with exactly $k$ elements. Similarly $[X]^{\mathbb{N}}$ stands for the collection of all infinite subsets of $X$. Note that $[X]^{\mathbb{N}}$ does not formally exist in second order arithmetic, and is only used in expressions of the form $Y \in [X]^{\mathbb{N}}$; here again we identify $Y$ with the unique sequence enumerating it in increasing order (notice that in $\mathsf{RCA}_0$ an element of $[X]^{\mathbb{N}}$ exists as a set if and only if it exists as an increasing sequence, so that this identification is harmless). For $X \in [\mathbb{N}]^{\mathbb{N}}$ let $X^- = X \setminus \{\min X\}$, i.e., $X$ with its least element removed. Similarly if $s \in [\mathbb{N}]^{<\mathbb{N}}$ is nonempty we set $s^- = s \setminus \{\min s\}$.

If $B \subseteq [\mathbb{N}]^{<\mathbb{N}}$ then $\operatorname{base}(B)$ is the set

$$\{n : \exists s \in B\, \exists i < \operatorname{lh} s\, s(i) = n\}.$$

$\mathsf{RCA}_0$ does not prove the existence of $\operatorname{base}(B)$ for arbitrary $B \subseteq [\mathbb{N}]^{<\mathbb{N}}$; indeed in [39, Lemma 1.4] it is shown that, over $\mathsf{RCA}_0$, $\mathsf{ACA}_0$ is equivalent to the assertion that $\operatorname{base}(B)$ exists as a set for every $B \subseteq [\mathbb{N}]^{<\mathbb{N}}$. However this does not affect the possibility of defining blocks and barriers within $\mathsf{RCA}_0$: e.g., "$\operatorname{base}(B)$ is infinite" (which is condition (1) in the definition of block below) can be expressed by $\forall m\, \exists n > m\, \exists s \in B\, n \in s$. Similarly, when we say $X$ is a subset of $\operatorname{base}(B)$ (for example in condition (2) of the definition of block), we mean $\forall x \in X\, \exists s \in B\, x \in s$. After giving the definitions, Lemma 3.2 will show that in fact $\mathsf{RCA}_0$ proves that $\operatorname{base}(B)$ exists whenever $B$ is a block (and, a fortiori, a barrier).

**Definition 3.1** ($\mathsf{RCA}_0$) A set $B \subseteq [\mathbb{N}]^{<\mathbb{N}}$ is a *block* if:

(1) $\text{base}(B)$ is infinite;
(2) $\forall X \in [\text{base}(B)]^{\mathbb{N}}\ \exists s \in B\ s \sqsubset X$;
(3) $\forall s, t \in B\ s \not\sqsubset t$.

$B$ is a *barrier* if it satisfies (1), (2) and

(3') $\forall s, t \in B\ s \not\subset t$.

Within $\mathsf{RCA}_0$ it is immediate that every barrier is a block and we can check that $[\mathbb{N}]^k$ (for $k > 0$), $\left\{ s \in [\mathbb{N}]^{<\mathbb{N}} : \text{lh}\, s = s(0) + 1 \right\}$ and $\left\{ s \in [\mathbb{N}]^{<\mathbb{N}} : \text{lh}\, s = s(s(0)) + 1 \right\}$ are barriers.

Notice that if $B$ is a block and $Y \in [\text{base}(B)]^{\mathbb{N}}$ then $\mathsf{RCA}_0$ proves that there exists a unique block $B' \subseteq B$ such that $\text{base}(B') = Y$: in fact $B' = \{ s \in B : s \subset Y \}$. Moreover if $B$ is a barrier then $B'$ is also a barrier and we say that $B'$ is a *subbarrier* of $B$.

The following result is Lemma 5.5 of [3].

**Lemma 3.2** ($\mathsf{RCA}_0$) *If $B$ is a block then* $\text{base}(B)$ *exists as a set and $B$ is isomorphic to a block $B'$ with* $\text{base}(B') = \mathbb{N}$.

**Definition 3.3** ($\mathsf{RCA}_0$) Let $s, t \in [\mathbb{N}]^{<\mathbb{N}}$: we write $s \lhd t$ if there exists $u \in [\mathbb{N}]^{<\mathbb{N}}$ such that $s \sqsubseteq u$ and $t \sqsubseteq u^-$.

Notice that $\langle 2, 4, 9\rangle \lhd \langle 4, 9, 10, 14\rangle \lhd \langle 9, 10, 14, 21\rangle$ and $\langle 2, 4, 9\rangle \not\lhd \langle 9, 10, 14, 21\rangle$, so that $\lhd$ is not transitive.

**Definition 3.4** ($\mathsf{RCA}_0$) Let $(Q, \preceq)$ be a quasi-order, $B$ be a block and $f : B \to Q$. We say that $f$ is *good* (with respect to $\preceq$) if there exist $s, t \in B$ such that $s \lhd t$ and $f(s) \preceq f(t)$. If $f$ is not good then we say that it is *bad*. $f$ is *perfect* if for every $s, t \in B$ such that $s \lhd t$ we have $f(s) \preceq f(t)$.

We can now give the definition of bqo:

**Definition 3.5** ($\mathsf{RCA}_0$) Let $(Q, \preceq)$ be a quasi-order.

- $Q$ is *bqo* if for every barrier $B$ every $f : B \to Q$ is good with respect to $\preceq$;
- $Q$ is *bqo(block)* if for every block $B$ every $f : B \to Q$ is good with respect to $\preceq$.

An alternative definition of bqo was given by Simpson in [59]. A block $B$ represents an infinite partition of $[\text{base}(B)]^{\mathbb{N}}$ into clopen sets with respect to the topology that $[\text{base}(B)]^{\mathbb{N}}$ inherits from $\mathbb{N}^{\mathbb{N}}$. Thus any $f : B \to Q$ represents a continuous function $F : [\text{base}(B)]^{\mathbb{N}} \to Q$ where $Q$ has the discrete topology; $f$ is good if for some $X \in [\text{base}(B)]^{\mathbb{N}}$ we have $F(X) \preceq F(X^-)$. Therefore $(Q, \preceq)$ is bqo if and only if for every continuous function $F : [\text{base}(B)]^{\mathbb{N}} \to Q$ there exists $X \in [\text{base}(B)]^{\mathbb{N}}$ such that $F(X) \preceq F(X^-)$. Moreover if we replace continuous with Borel we are still defining the same notion (this follows from the fact, originally proved by Mathias, that for every Borel function $F : [\text{base}(B)]^{\mathbb{N}} \to Q$ there exists $X \in [\text{base}(B)]^{\mathbb{N}}$ such that the restriction of $F$ to $[X]^{\mathbb{N}}$ is continuous). We are not discussing these alternative characterizations of bqo here, but they have been exploited by Montalbán in his proof of Theorem 6.28.

It is easy to see (using the barrier $[\mathbb{N}]^1$ and the fact that $\langle m \rangle \lhd \langle n \rangle$ if and only if $m < n$) that $\mathsf{RCA}_0$ proves that every bqo is wqo.

Lemma 3.2 shows that within $\mathsf{RCA}_0$ we can restrict the definition of bqo and bqo(block) to functions with domain barriers or blocks with base $\mathbb{N}$. It is also immediate that every bqo(block) is also a bqo. For the opposite implication, we have the following result [3, Theorem 5.12].

**Lemma 3.6** ($\mathsf{WKL}_0$) *Every bqo is bqo(block).*

The natural proof that every bqo is bqo(block) uses the clopen Ramsey theorem, which is equivalent to $\mathsf{ATR}_0$, to show that every block contains a barrier. The proof of Lemma 3.6 instead exploits a construction originally appeared in [37] and builds a barrier which is connected to, but in general not included in, the original block.

Lemma 3.6 leads to the following question:

**Question 3.7** Is "every bqo is bqo(block)" equivalent to $\mathsf{WKL}_0$ over $\mathsf{RCA}_0$?

Another characterization of bqos corresponds to the wqo(set) characterization of wqos.

**Definition 3.8** A quasi-order $(Q, \preceq)$ is *bqo(set)* if for every barrier $B$ and every $f : B \to Q$ there exists a subbarrier $B' \subseteq B$ such that $f$ restricted to $B'$ is perfect with respect to $\preceq$.

$\mathsf{RCA}_0$ trivially proves that every bqo(set) is bqo, while the reverse implication is known to be much stronger (see [39, Theorem 4.9], which revisits [62, Lemma V.9.5]).

**Theorem 3.9** ($\mathsf{RCA}_0$) *The following are equivalent:*

*(i)* $\mathsf{ATR}_0$*;*
*(ii) every bqo is bqo(set).*

It is easy to realize that $\mathsf{RCA}_0$ suffices to prove that every well-order is bqo, and even bqo(block) (see [39, Lemma 3.1]). Dealing with finite quasi-orders is however more problematic. Let $n$ denote the partial order consisting of $n$ mutually incomparable elements, and notice that if $n$ is bqo, or bqo(block), or bqo(set), then every quasi-order with the same number of elements has the same property. The following results are from [39, Lemma 3.2, Theorem 5.11 and Theorem 4.9].

**Theorem 3.10** *1.* $\mathsf{RCA}_0$ *proves that* 2 *is bqo and bqo(block);*
*2.* $\mathsf{ATR}_0$ *proves that* 3 *is bqo;*
*3. for any fixed $n \geq 3$,* $\mathsf{RCA}_0$ *proves that* 3 *is bqo is equivalent to $n$ is bqo;*
*4. for any fixed $n \geq 2$,* $\mathsf{RCA}_0$ *proves that $n$ is bqo(set) is equivalent to* $\mathsf{ATR}_0$.

Item *(3)* above leads to the following question, which was already stated as Problem 3.3 in [39].

**Question 3.11** What is the strength of the statement “3 is bqo”?

Over the years, the author has involved several colleagues in trying to attack this problem, but no progress has been made. We devote some time to explain the situation. The $\lhd$ relation can be viewed as defining a graph with the elements of $[\mathbb{N}]^{\mathbb{N}}$ as vertices. The assertion that $n$ is bqo amounts to state that the subgraph whose set of vertices is a barrier is not $n$-colorable. Indeed, the proof of item *(1)* of Theorem 3.10 amounts to the definition within $\mathsf{RCA}_0$ of a cycle of odd length inside any barrier or block. It is much more difficult to show that a graph is not 3-colorable, and this accounts for the increased difficulty in showing that 3 is bqo. A first step in beginning to answer Question 3.11 would be showing that the $\omega$-model **REC** does not satisfy that every barrier is 3-colorable. To this end one cannot use a computable barrier $B$: in fact being 3-colorable is an arithmetic property, and hence surely false for $B$ in **REC**. What is needed is some $B \subseteq [\mathbb{N}]^{\mathbb{N}}$ which looks like a barrier in **REC** (i.e., which satisfies *(1)* and *(3')* of Definition 3.1 and is such that for every computable $X \in [\text{base}(B)]^{\mathbb{N}}$ there exists $s \in B$ with $s \sqsubset X$), but is 3-colorable.

Moving now to the basic closure properties of bqos, we start by noticing the following obvious fact, which mirrors the results of Lemma 2.14 about wqos.

**Lemma 3.12** ($\mathsf{RCA}_0$) *Let $\mathcal{P}$ be any of the properties bqo, bqo(block) or bqo(set). If $(Q, \preceq)$ satisfies $\mathcal{P}$ and $R \subseteq Q$ then the restriction of $\preceq$ to $R$ satisfies $\mathcal{P}$ as well.*

Only part of Lemma 2.17 has an analogous for bqos.

**Lemma 3.13** ($\mathsf{RCA}_0$) *Let $\mathcal{P}$ be any of the properties bqo, bqo(block) or bqo(set). If $Q_1$ and $Q_2$ satisfy $\mathcal{P}$ then $Q_1 + Q_2$ satisfies $\mathcal{P}$ with respect to the sum quasi-order.*

When $\mathcal{P}$ is bqo this is [39, Lemma 5.14]. The proof shows that for any $f : B \to Q_1 + Q_2$ there is a subbarrier $B'$ such that the restriction of $f$ to $B'$ has range in $Q_i$ for some $i$: this yields the result also when $\mathcal{P}$ is bqo(set). Moreover the proof works also for blocks, thus taking care of the case when $\mathcal{P}$ is bqo(block).

The closure under disjoint unions of bqos is much stronger than the corresponding property for wqos. In fact we have

**Lemma 3.14** ($\mathsf{RCA}_0$) *Let $\mathcal{P}$ be any of the properties bqo, bqo(block) or bqo(set). The following are equivalent:*

*(i)* *if $Q_1$ and $Q_2$ satisfy $\mathcal{P}$ then $Q_1 \cup Q_2$ satisfies $\mathcal{P}$ with respect to the disjoint union quasi-order;*
*(ii)* *if $Q$ satisfies $\mathcal{P}$ with respect to the quasi-orders $\preceq_1$ and $\preceq_2$ then $Q$ satisfies $\mathcal{P}$ with respect to the intersection quasi-order;*
*(iii)* *if $Q_1$ and $Q_2$ satisfy $\mathcal{P}$ then $Q_1 \times Q_2$ satisfies $\mathcal{P}$ with respect to the product quasi-order.*

*All these statements are provable in* $\mathsf{ATR}_0$. *When* $\mathcal{P}$ *is bqo or bqo(block) they imply* $\mathsf{ACA}_0$*, when* $\mathcal{P}$ *is bqo(set) they are equivalent to* $\mathsf{ATR}_0$.

The equivalence between the three statements for bqo is Lemma 5.16 of [39]: the implication from *(i)* to *(iii)* uses Theorem 6.6. The same proof works also for bqo(block) and bqo(set). Provability in $\mathsf{ATR}_0$ follows easily from the clopen Ramsey theorem. The implication towards $\mathsf{ACA}_0$ is Lemma 5.17 of [39] (which uses the proof of Theorem 6.5) when we are dealing with bqos, and works also for bqo(block). The implication towards $\mathsf{ATR}_0$ is immediate from item *(4)* of Theorem 3.10 because *(i)* for bqo(set) implies that 2 is bqo(set).

**Question 3.15** What is the strength of statements *(i)–(iii)* of Lemma 3.14 when $\mathcal{P}$ is bqo or bqo(block)?

Since the statements imply $\mathsf{ACA}_0$, by Lemma 3.6 there is a single answer for bqo and bqo(block). Since *(i)* for bqo implies that 3 is bqo, Questions 3.11 and 3.15 are connected.

## 4 Minimality Arguments

One of the main tools of wqo theory is the minimal bad sequence lemma (apparently isolated for the first time in [49]). The idea is to prove that a quasi-order is wqo by showing that if there exists a bad sequence then there is one with a minimality property, and eventually reaching a contradiction from the latter assumption. To state the lemma in its general form we need the following definitions.

**Definition 4.1** ($\mathsf{RCA}_0$) Let $(Q, \preceq)$ be a quasi-order. A transitive binary relation $<'$ on $Q$ is *compatible with* $\preceq$ if for every $x, y \in Q$ we have that $x <' y$ implies $x \preceq y$. We write $x \leq' y$ for $x <' y \vee x = y$. In this situation, if $A, A' \in [\mathbb{N}]^{\mathbb{N}}$, $f : A \to Q$, and $f' : A' \to Q$ we write $f \leq' f'$ if $A \subseteq A'$ and $\forall n \in A\ f(n) \leq' f'(n)$; we write $f <' f'$ if $f \leq' f'$ and $\exists n \in A\ f(n) <' f'(n)$. $f$ is *minimal bad with respect to* $<'$ if it is bad with respect to $\preceq$ and there is no $f' <' f$ which is bad with respect to $\preceq$.

**Statement 4.2** (minimal bad sequence lemma) Let $(Q, \preceq)$ be a quasi-order and $<'$ a well-founded relation which is compatible with $\preceq$: if $A' \in [\mathbb{N}]^{\mathbb{N}}$ and $f' : A' \to Q$ is bad with respect to $\preceq$ then there exists $f : A \to Q$ such that $f \leq' f'$ and $f$ is minimal bad with respect to $<'$.

The generalization of the minimal bad sequence lemma to bqos is known as the minimal bad array lemma (the maps of Definition 3.4 are sometimes called arrays) or the forerunning technique (this method was explicitly isolated and clarified in [34]). Again, we need some preliminary definitions.

**Definition 4.3** ($\mathsf{RCA}_0$) Let $(Q, \preceq)$ be a quasi-order and $<'$ be compatible with $\preceq$ in the sense of Definition 4.1. If $B$ and $B'$ are barriers, $f : B \to Q$, and $f' : B' \to Q$

we write $f \leq' f'$ if $\mathrm{base}(B) \subseteq \mathrm{base}(B')$, and for every $s \in B$ there exists $s' \in B'$ such that $s' \sqsubseteq s$ and $f(s) \leq' f'(s')$. We write $f <' f'$ if $f \leq' f'$ and for some $s \in B$, $s' \in B'$ with $s' \sqsubseteq s$ we have $f(s) <' f'(s')$. $f$ is *minimal bad with respect to* $<'$ if it is bad with respect to $\preceq$ and there is no $f' <' f$ which is bad with respect to $\preceq$.

**Statement 4.4** (minimal bad array lemma) Let $(Q, \preceq)$ be a quasi-order and $<'$ a well-founded relation which is compatible with $\preceq$. If $B'$ is a barrier and $f' : B' \to Q$ is bad with respect to $\preceq$ then there exist a barrier $B$ and $f : B \to Q$ such that $f \leq' f'$ and $f$ is minimal bad with respect to $<'$.

A milder generalization of the minimal bad sequence lemma is also useful: it was actually the first version of the minimal bad array lemma proved for a specific quasi-order by Nash-Williams in [50] and was isolated in [37].

**Definition 4.5** ($\mathsf{RCA}_0$) Let $(Q, \preceq)$ be a quasi-order and $<'$ be compatible with $\preceq$ in the sense of Definition 4.1. If $B$ and $B'$ are barriers, $f : B \to Q$, and $f' : B' \to Q$ we write $f \leq'_\ell f'$ if $B \subseteq B'$ and $\forall s \in B\ f(s) \leq' f'(s)$. We write $f <'_\ell f'$ if $f \leq'_\ell f'$ and $\exists s \in B\ f(s) <' f'(s)$. $f$ is *locally minimal bad with respect to* $<'$ if it is bad with respect to $\preceq$ and there is no $f' <'_\ell f$ which is bad with respect to $\preceq$.

**Statement 4.6** (locally minimal bad array lemma) Let $(Q, \preceq)$ be a quasi-order and $<'$ a well-founded relation which is compatible with $\preceq$: if $B'$ is a barrier and $f' : B' \to Q$ is bad with respect to $\preceq$ then there exist a barrier $B$ and $f : B \to Q$ such that $f \leq'_\ell f'$ and $f$ is locally minimal bad with respect to $<'$.

The minimal bad sequence lemma and the locally minimal bad array lemma have been shown to be equivalent to the strongest of the big five by Simpson and Marcone in [38, Theorem 6.5].

**Theorem 4.7** ($\mathsf{RCA}_0$) *The following are equivalent:*

*(i)* $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$*;*
*(ii) the minimal bad sequence lemma;*
*(iii) the locally minimal bad array lemma.*

On the other hand, the proofs of the minimal bad array lemma use very strong set-existence axioms: a crude analysis shows that they can be carried out within $\mathbf{\Pi}^1_2$-$\mathsf{CA}_0$.

**Question 4.8** What is the axiomatic strength of the minimal bad array lemma?

## 5 Structural Results

In this section we consider theorems showing that wqos satisfy specific properties as partial orders.

The better known of these theorems is due to de Jongh and Parikh [28] (an exposition of essentially the original proof appears in [20, Sect. 8.4]; a proof based on the study of the partial order of the initial segments of the wqo is included in [11, Sect. 4.11]; proofs with a strong set-theoretic flavor appear as [29, Theorem 4.7] and [1, Proposition 52]).

**Statement 5.1** (maximal linear extension theorem) If $(Q, \preceq)$ is wqo, then there exist a linear extension $\preceq_L$ of $Q$ which is maximal, meaning that every linear extension of $Q$ embeds in an order-preserving way into $\preceq_L$.

A less known result is due to Wolk ([67, Theorem 9], actually Wolk's statement is slightly stronger) and also appears as [29, Theorem 4.9] and [20, Theorem 8.1.7].

**Statement 5.2** (maximal chain theorem) If $(Q, \preceq)$ is wqo, then there exist a chain $C \subseteq Q$ which is maximal, meaning that every chain contained in $Q$ embeds in an order-preserving way into $C$.

Marcone and Shore [42] studied the strength of the maximal linear extension theorem and of the maximal chain theorem.

**Theorem 5.3** ($\mathsf{RCA}_0$) *The following are equivalent:*

*(i)* $\mathsf{ATR}_0$*;*
*(ii) the maximal linear extension theorem;*
*(iii) the maximal chain theorem.*

The proofs of the two theorems within $\mathsf{ATR}_0$ differ from the proofs found in the literature: to avoid using more induction than available in $\mathsf{ATR}_0$ one fixes a wqo $Q$ and looks respectively at the *tree of finite bad sequences in $Q$*

$$\mathrm{Bad}(Q) = \left\{ s \in Q^{<\mathbb{N}} : \forall i, j < \mathrm{lh}\, s (i < j \rightarrow s(i) \npreceq s(j)) \right\}$$

and at the *tree of descending sequences in $Q$*

$$\mathrm{Desc}(Q) = \left\{ s \in Q^{<\mathbb{N}} : \forall i, j < \mathrm{lh}\, s (i < j \rightarrow s(j) \prec s(i)) \right\}.$$

(Here $Q^{<\mathbb{N}}$ is the set of finite sequences of elements of $Q$.) Since $Q$ is wqo both these trees are well-founded and $\mathsf{ATR}_0$ can compute their rank functions. Focusing on the maximal linear extension theorem (the other proof follows the same strategy), by recursion on the rank of $s \in \mathrm{Bad}(Q)$ we assign to $s$ a maximal linear extension of the restriction of $\preceq$ to $\left\{ x \in Q : s^\frown\langle x \rangle \in \mathrm{Bad}(Q) \right\}$; when $s$ is the empty sequence we have the maximal linear extension of $Q$.

The two reversals contained in Theorem 5.3 have quite different proofs. The proof that the maximal chain theorem implies $\mathsf{ATR}_0$ is very simple (using the well-known equivalence between $\mathsf{ATR}_0$ and comparability of well-orders), while the proof that the maximal linear extension theorem implies $\mathsf{ATR}_0$ is more involved. In fact there is first a bootstrapping, showing that the maximal linear extension theorem implies $\mathsf{ACA}_0$.

To this end it is useful a partial order $Q$ such that the existence of any bad sequence in $Q$ implies $\mathsf{ACA}_0$: thus if $\mathsf{ACA}_0$ fails then $Q$ is wqo, we can apply the theorem and reach a contradiction from the existence of a maximal linear extension. We can now argue within $\mathsf{ACA}_0$ and, assuming the failure of $\mathsf{ATR}_0$ and using Theorem 6.23, build a wqo $Q'$ which cannot have a maximal linear extension. The difference of the two proofs is no accident. In fact a theorem of Montalbán [45] states that every computable wqo has a computable maximal linear extension (this implies that in showing that the maximal linear extension theorem implies anything unprovable in $\mathsf{RCA}_0$ the use of partial orders that are not really wqos is unavoidable), while Marcone, Montalbán and Shore [43, Theorem 3.3] showed that for every hyperarithmetic set $X$ there is a computable wqo $Q$ with no $X$-computable maximal chain.

Another kind of structural theorems about quasi-orders concerns the decomposability of the quasi-order in finite pieces which are simple.

**Definition 5.4** ($\mathsf{RCA}_0$) Let $(Q, \preceq)$ be a quasi-order. $I \subseteq Q$ is an *ideal* if

- $\forall x, y \in Q (x \in I \land y \preceq x \to y \in I)$;
- $\forall x, y \in I\, \exists z \in I (x \preceq z \land y \preceq z)$.

Bonnet [2, Lemma 2] (see also [11, Sect. 4.7.2]) proved that a partial order has no infinite antichains if and only if every initial interval is a finite union of ideals (this result follows also from [9, Theorem 1]). In [15, Theorem 4.5] Frittaion and Marcone studied the left to right direction of Bonnet's result and proved, among other things, the following equivalence.

**Theorem 5.5** ($\mathsf{RCA}_0$) *The following are equivalent:*

*(i)* $\mathsf{ACA}_0$*;*
*(ii) every wqo is a finite union of ideals.*

## 6 Major Theorems About wqos and bqos

In this section we consider the major theorems about wqos and bqos, starting with Higman's basic result, first proved in [22] and then rediscovered many times.

**Definition 6.1** ($\mathsf{RCA}_0$) If $(Q, \preceq)$ is a quasi-order we define a quasi-order on $Q^{<\mathbb{N}}$ by setting $s \preceq^* t$ if and only if there exists an embedding of $s$ into $t$, i.e., a strictly increasing $f : \operatorname{lh} s \to \operatorname{lh} t$ such that $s(i) \preceq t(f(i))$ for every $i < \operatorname{lh} s$ (here $\operatorname{lh} s$ is the length of the sequence $s$).

**Statement 6.2** (Higman's theorem) If $Q$ is wqo then $(Q^{<\mathbb{N}}, \preceq^*)$ is wqo.

Before analyzing Higman's theorem from the reverse mathematics viewpoint, let us introduce other constructions of new quasi-orders starting from the one on $Q$.

We denote by $\mathcal{P}(X)$ and $\mathcal{P}_{\mathrm{f}}(X)$ the powerset of $X$ and the set of all finite subsets of $X$. If $X$ is infinite $\mathcal{P}(X)$ does not exists as a set in second order arithmetic, but we can define and study relations between elements of $\mathcal{P}(X)$. A *quasi-order on* $\mathcal{P}(X)$ is just a formula $\varphi$ with two distinguished set variables such that $\varphi(Y, Y)$ holds and $\varphi(Y, Z)$ and $\varphi(Z, W)$ imply $\varphi(Y, W)$ whenever $Y, Z, W \subseteq X$. We use symbols like $\preceq$ and infix notation to denote quasi-orders on $\mathcal{P}(X)$.

**Definition 6.3** ($\mathsf{RCA}_0$) If $\preceq$ is a quasi-order on $\mathcal{P}(X)$, a sequence $(X_n)_{n\in\mathbb{N}}$ of elements of $\mathcal{P}(X)$ is *good* (with respect to $\preceq$) if there exist $m <_{\mathbb{N}} n$ such that $X_m \preceq X_n$. If every such sequence is good we say that $\preceq$ is *wqo*.

Analogously, a sequence $(X_s)_{s\in B}$ of elements of $\mathcal{P}(X)$ indexed by a barrier $B$ is *good* (with respect to $\preceq$) if there exist $s, t \in B$ such that $s \lhd t$ and $X_s \preceq X_t$. If every such sequence is good we say that $\preceq$ is *bqo*.

The following two quasi-orders are called the *Hoare quasi-order* and the *Smyth quasi-order* in the computer science literature. (Here we follow the computer science notation: in [39] $\preceq^\flat$ was written as $\preceq^\exists_\forall$ and $\preceq^\sharp$ as $\preceq^\forall_\exists$.)

**Definition 6.4** ($\mathsf{RCA}_0$) Let $(Q, \preceq)$ be a quasi-order. If $X, Y \in \mathcal{P}(Q)$ let

$$X \preceq^\flat Y \iff \forall x \in X\, \exists y \in Y\; x \preceq y \quad \text{and}$$
$$X \preceq^\sharp Y \iff \forall y \in Y\, \exists x \in X\; x \preceq y.$$

**Theorem 6.5** ($\mathsf{RCA}_0$) *The following are equivalent:*

*(i)* $\mathsf{ACA}_0$*;*
*(ii) Higman's theorem;*
*(iii) if $Q$ is wqo then $(\mathcal{P}_\mathrm{f}(Q), \preceq^\flat)$ is wqo.*

Most proofs of Higman's theorem are based on the minimal bad sequence lemma. Theorem 4.7 implies that such a proof cannot be carried out in $\mathsf{ACA}_0$. In fact, the provability of Higman's theorem in $\mathsf{ACA}_0$ is based on the technique of reification of wqos by well-orders ([28, 55], see also [29]) and follows from the results in Sect. 4 of [61] (see [4, Theorem 3] for details). A *reification* of $Q$ by the linear order $(X, \leq_X)$ is a map $\rho$ from $\mathrm{Bad}(Q)$ to $X$ such that $\rho(t) <_X \rho(s)$ whenever $s \sqsubset t$. Thus, if $X$ is a well-order then $\rho$ is an approximation to the rank function on $B(Q)$, and suffices to witness that $\mathrm{Bad}(T)$ is well-founded and hence $Q$ is wqo.

$\mathsf{ACA}_0$ is used twice in this proof: first to show that every wqo admits a reification by a well-order and then to show that $\omega^{\omega^{\alpha+1}}$ is a well-order when $\alpha$ is a well-order (closure of well-orders under exponentiation is equivalent to $\mathsf{ACA}_0$ over $\mathsf{RCA}_0$ by [17], see [27]). This suffices, because $\mathsf{RCA}_0$ proves that if $Q$ has a reification of order type $\alpha$ then $Q^{<\mathbb{N}}$ has a reification of order type $\omega^{\omega^{\alpha+1}}$ ([61, Sublemma 4.8], which is Lemma 5.2 of [56]) and that if a quasi-order admits a reification by a well-order then it is wqo.

$\mathsf{RCA}_0$ clearly suffices to prove that *(ii)* implies *(iii)* of Theorem 6.5, while the implication from *(iii)* to $\mathsf{ACA}_0$ was proved in [39] using $\mathsf{RT}^2_2$ (but in fact only closure of wqo under product was necessary); the provability of this implication in $\mathsf{RCA}_0$ was shown in [16, Theorem 2.5].

If $Q$ is wqo then in general neither $\mathcal{P}(Q)$ with respect to $\preceq^\flat$ nor $\mathcal{P}_f(Q)$ with respect to $\preceq^\sharp$ are wqo. However if we strengthen the hypothesis to $Q$ bqo we obtain some true statements which have been studied from the reverse mathematics viewpoint. The following theorems summarize Theorems 5.4 and 5.6 in [39].

**Theorem 6.6** ($\mathsf{RCA}_0$) *If $Q$ is bqo then $(\mathcal{P}_f(Q), \preceq^\flat)$ and $(\mathcal{P}(Q), \preceq^\sharp)$ (and hence, a fortiori, also $(\mathcal{P}_f(Q), \preceq^\sharp)$) are bqo.*

**Theorem 6.7** ($\mathsf{ACA}_0$) *If $Q$ is bqo then $(\mathcal{P}(Q), \preceq^\flat)$ is bqo.*

**Question 6.8** Is the statements "if $Q$ is bqo then $(\mathcal{P}(Q), \preceq^\flat)$ is bqo" equivalent to $\mathsf{ACA}_0$ over $\mathsf{RCA}_0$?

Trying to answer affirmatively the previous question, one is faced with the problem of applying the statement to a quasi-order $Q$ which is proved to be bqo within $\mathsf{RCA}_0$. Such a $Q$ must be infinite (otherwise $\mathcal{P}(Q) = \mathcal{P}_f(Q)$ and Theorem 6.6 applies) and, unless the answer to Question 3.11 is $\mathsf{RCA}_0$, with antichains of size at most 2.

More results about the Hoare and Smyth quasi-orders (obtained by weakening the conclusion) will be discussed in Sect. 7.

Another important result about wqos is Kruskal's theorem [32], establishing a conjecture of Vázsonyi from the 1930's popularized by Erdős. This theorem deals with trees viewed as partial orders: for our purposes we can represent them in second-order arithmetic as subsets of $\mathbb{N}^{<\mathbb{N}}$ closed under initial segments.

**Definition 6.9** ($\mathsf{RCA}_0$) Let $\mathcal{T}$ be the set of all finite trees. If $T_0, T_1 \in \mathcal{T}$ let $T_0 \preceq_{\mathcal{T}} T_1$ if and only if there exists a homeomorphic embedding of $T_0$ in $T_1$, that is, an injective $f : T_0 \to T_1$ such that $f(s \wedge t) = f(s) \wedge f(t)$ for every $s, t \in T_0$ (where $s \wedge t$ is the greatest lower bound of $s$ and $t$, which is the longest common initial segment of the two sequences).

If $Q$ is a set let $\mathcal{T}^Q$ be the set of finite trees labelled with elements of $Q$, that is, pairs $(T, \ell)$ such that $T \in \mathcal{T}$ and $\ell$ is a function from $T$ to $Q$.

If $(Q, \preceq)$ is a quasi-order and $(T_0, \ell_0), (T_1, \ell_1) \in \mathcal{T}^Q$ let $(T_0, \ell_0) \preceq_{\mathcal{T}^Q} (T_1, \ell_1)$ if and only if there exists a homeomorphic embedding $f$ of $T_0$ in $T_1$ such that $\ell_0(s) \preceq \ell_1(f(s))$ for every $s \in T_0$.

$\mathsf{RCA}_0$ easily shows that $\preceq_{\mathcal{T}}$ and $\preceq_{\mathcal{T}^Q}$ are quasi-orders.

**Statement 6.10** (Kruskal's theorem) If $Q$ is wqo then $(\mathcal{T}^Q, \preceq_{\mathcal{T}^Q})$ is wqo.

The usual proof of Kruskal's theorem uses the minimal bad sequence lemma and can be carried out in $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$ using Theorem 4.7. On the other hand, this statement is $\mathbf{\Pi}^1_2$ and hence cannot imply $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$ (see [38, Corollary 1.10]).

Harvey Friedman proved the following striking result (see [60]).

**Theorem 6.11** $\mathsf{ATR}_0$ *does not prove that $(\mathcal{T}, \preceq_{\mathcal{T}})$ is wqo. A fortiori Kruskal's theorem is not provable in* $\mathsf{ATR}_0$.

To prove this theorem we build a map $\psi$ between $\mathcal{T}$ and a certain primitive recursive notation system for the ordinals less than $\Gamma_0$, and show that $\mathsf{ACA}_0$ proves that $\psi(T_0) \leq_o \psi(T_1)$ (where $\leq_o$ is the order on the ordinal notation system) whenever $T_0 \preceq_{\mathcal{T}} T_1$. Thus $\mathsf{ACA}_0$ proves that if $(\mathcal{T}, \preceq_{\mathcal{T}})$ is wqo then the system of ordinal notations is a well-order. Since $\Gamma_0$ is the proof-theoretic ordinal of $\mathsf{ATR}_0$, it follows that $\mathsf{ATR}_0$ does not prove that $(\mathcal{T}, \preceq_{\mathcal{T}})$ is wqo.

A lower bound for Kruskal's theorem is provided by the following theorem, that apparently has never been explicitly stated.

**Theorem 6.12** ($\mathsf{RCA}_0$) *Kruskal's theorem implies* $\mathsf{ATR}_0$.

*Sketch of proof.* We use the fact that $\mathsf{ATR}_0$ is equivalent, over $\mathsf{RCA}_0$, to the statement that if $X$ is a well-order then $\boldsymbol{\varphi}(X, 0)$ is a well-order, where $\boldsymbol{\varphi}$ is the formalization of the Veblen function on the ordinals. This theorem was originally proved by H. Friedman (unpublished) and then given a proof-theoretic proof in [54] and a computability-theoretic proof in [41]. We follow the notation of the latter paper.

To prove our theorem first notice that Kruskal's theorem generalizes Higman's theorem, so that we can argue in $\mathsf{ACA}_0$. Given a well-order $X$ we can mimic the construction of the proof of Theorem 6.11 using $X$ as the set of labels for the finite trees. In this way we define a map $\psi$ between $\mathcal{T}^X$ and the ordinals less than the first fixed point for the Veblen function strictly larger than $X$. We then show that $\mathsf{ACA}_0$ proves that $\psi(T_0) \leq_\varphi \psi(T_1)$ whenever $T_0 \preceq_{\mathcal{T}^X} T_1$. Since our hypothesis implies that $(\mathcal{T}^X, \preceq_{\mathcal{T}^X})$ is wqo we obtain that $(\boldsymbol{\varphi}(X, 0), \leq_\varphi)$ is a well-order, as needed. □

Thus Kruskal's theorem is properly stronger than $\mathsf{ATR}_0$ and provable in, but not equivalent to, $\boldsymbol{\Pi}^1_1$-$\mathsf{CA}_0$. In an attempt to classify statements of this kind, Henry Towsner [64] introduced a sequence of intermediate systems based on weakening the leftmost path principle (which is equivalent to $\boldsymbol{\Pi}^1_1$-$\mathsf{CA}_0$). Towsner tested his approach by looking at various statements and, by analyzing Nash-Williams' proof of Kruskal's theorem, obtained the following result.

**Theorem 6.13** *Kruskal's theorem is provable in Towsner's system* $\Sigma_2$-$\mathsf{LPP}_0$.

Unfortunately no reversal to Towsner's systems are known, so we do not know whether the upper bound for the strength of Kruskal's theorem provided by the previous theorem is optimal.

Rathjen and Weiermann [53] carried out a detailed proof-theoretic analysis of the statement "$(\mathcal{T}, \preceq_{\mathcal{T}})$ is wqo" (beware that Rathjen and Weiermann call Kruskal's theorem this statement) showing that it is equivalent over $\mathsf{ACA}_0$ to the uniform $\boldsymbol{\Pi}^1_1$ reflection principle of the theory obtained by adding transfinite induction for $\boldsymbol{\Pi}^1_2$ formulas to $\mathsf{ACA}_0$.

Harvey Friedman, inspired by ordinal notation systems, introduced a refinement of $\preceq_{\mathcal{T}}$ (obtained by requiring that the homeomorphic embedding satisfies a "gap condition") and proved that it still yields a wqo on $\mathcal{T}$. Friedman himself [60] showed, generalizing the technique of Theorem 6.11 to larger ordinals, that this wqo statement is not provable in $\boldsymbol{\Pi}^1_1$-$\mathsf{CA}_0$.

The most striking instance of this unprovability phenomenon is provided by the graph minor theorem, proved by Robertson and Seymour in a long series of papers (see [63, Chap. 5] or [6, Chap. 12] for overviews).

**Definition 6.14** ($\mathsf{RCA}_0$) If $\mathcal{G}$ is the set of all finite directed graphs (allowing loops and multiple edges) define a quasi-order on $\mathcal{G}$ by setting $G_0 \preceq_m G_1$ if and only if $G_0$ is isomorphic to a minor of $G_1$ (recall that a minor is obtained by deleting edges and vertices and contracting edges).

**Statement 6.15** (graph minor theorem) $\preceq_m$ is wqo on $\mathcal{G}$.

Friedman's generalization of Kruskal's theorem mentioned above plays a significant role in some steps of the proof of the graph minor theorem, which uses iterated applications of the minimal bad sequence lemma. This proof cannot be carried out in $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$ and the following theorem (proved by Friedman, Robertson and Seymour [14] well before the completion of the proof of the graph minor theorem) shows that there is no simpler proof.

**Theorem 6.16** *The graph minor theorem (and even special cases where $\preceq_m$ is restricted to some subset of $\mathcal{G}$) is not provable in $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$.*

This theorem is proved once more generalizing the technique of Theorem 6.11 to larger ordinals. Notice also that the graph minor theorem is a $\mathbf{\Pi}^1_1$ statement, and therefore does not imply any set-existence axiom (in fact it holds in every $\omega$-model). More recently Rathjen and Krombholz [30, 31] analyzed more in detail the proof by Robertson and Seymour in search of upper bounds for the proof-theoretic strength of this statement, showing that it can be carried out in the system obtained by adding transfinite induction for $\mathbf{\Pi}^1_2$ formulas to $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$.

It is well-known that Higman's theorem does not extend to infinite sequences, and the canonical counterexample is Rado's partial order. The notion of bqo was developed by Nash-Williams as a way of ruling out Rado's example and its generalizations. Indeed, one of the first theorems of the subject is a generalization of Higman's theorem [51].

**Definition 6.17** ($\mathsf{RCA}_0$) If $(Q, \preceq)$ is a quasi-order we can extend the quasi-order $\preceq^*$ of Definition 6.1 from $Q^{<\mathbb{N}}$ to $\tilde{Q}$, the set of all countable sequences of elements of $Q$ (i.e., the set of all functions from a countable well-order to $Q$).

**Statement 6.18** (Nash-Williams' theorem) If $Q$ is bqo then $(\tilde{Q}, \preceq^*)$ is bqo.

Notice that $\tilde{Q}$ is uncountable, and hence we express "$(\tilde{Q}, \preceq^*)$ is bqo" in a way similar to Definition 6.3.

The following theorem is [38, Theorem 4.5].

**Theorem 6.19** $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$ *proves Nash-Williams' theorem.*

The most natural proof of Nash-Williams' theorem uses the minimal bad array lemma, and therefore to prove Theorem 6.19 a new argument is needed. This is obtained by using the locally minimal bad array lemma (provable in $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$ by Theorem 4.7) to establish the following weak version of Nash-Williams' theorem.

**Statement 6.20** (Generalized Higman's theorem) If $Q$ is bqo then $(Q^{<\mathbb{N}}, \preceq^*)$ is bqo.

Assuming the generalized Higman's theorem, we can prove Nash-Williams' theorem in $\mathsf{ATR}_0$. Thus the proof of Theorem 6.19 yields the following result.

**Theorem 6.21** ($\mathsf{ATR}_0$) *The following are equivalent:*

*(i) Nash-Williams' theorem;*
*(ii) the generalized Higman's theorem.*

Nash-Williams' theorem cannot imply $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$, even over $\mathsf{ATR}_0$ [38, Theorem 5.7]. In fact, the proof of Theorem 6.19 actually establishes a $\mathbf{\Pi}^1_2$ statement that, over $\mathsf{ATR}_0$, implies Nash-Williams' theorem. The argument mentioned before Theorem 6.11 then establishes the assertion. (Both Nash-Williams' theorem and the generalized Higman's theorem are $\mathbf{\Pi}^1_3$ statements, so we cannot apply the argument directly.) Towsner [64] looked also at the proof of the locally minimal bad array lemma.

**Theorem 6.22** *The generalized Higman's theorem, and therefore also Nash-Williams' theorem, is provable in Towsner's system* $\mathsf{TLPP}_0$.

$\mathsf{TLPP}_0$ is much stronger than the system $\Sigma_2$-$\mathsf{LPP}_0$ appearing in Theorem 6.13. Unfortunately, as already mentioned, no reversal to Towsner's systems are known, so Theorem 6.22 provides just an upper bound for the strength of Nash-Williams' theorem. Regarding lower bounds, Shore [58] proved the following important result.

**Theorem 6.23** ($\mathsf{RCA}_0$) *The following are equivalent:*

*(i)* $\mathsf{ATR}_0$
*(ii) every infinite sequence of countable well-orders contains two distinct elements which are comparable with respect to embeddability (as defined in Definition 6.25).*

It is immediate that Nash-Williams' theorem implies *(ii)*, and hence $\mathsf{ATR}_0$, within $\mathsf{RCA}_0$.

**Question 6.24** Is Nash-Williams' theorem equivalent to $\mathsf{ATR}_0$?

A positive answer to this question was conjectured in [38, 39].

Connected to Nash-Williams' theorem is one of the most famous achievements of bqo theory, Laver's proof [33] of Fraïssé's conjecture [10]. Laver actually proved a stronger result (even stronger than the one we state below) and we keep the two statements distinct.

**Definition 6.25** ($\mathsf{RCA}_0$) If $\mathcal{L}$ is the set of countable linear orderings define the quasi-order of embeddability on $\mathcal{L}$ by setting $L_0 \preceq_{\mathcal{L}} L_1$ if and only if there exists an order-preserving embedding of $L_0$ in $L_1$, i.e., an injective $f : L_0 \to L_1$ such that $x <_{L_0} y$ implies $f(x) <_{L_1} f(y)$ for every $x, y \in L_0$.

**Statement 6.26** (Fraïssé's conjecture) $(\mathcal{L}, \preceq_{\mathcal{L}})$ is wqo.

**Statement 6.27** (Laver's theorem) $(\mathcal{L}, \preceq_{\mathcal{L}})$ is bqo.

Again, $\mathcal{L}$ is uncountable, and hence we express "$(\mathcal{L}, \preceq_{\mathcal{L}})$ is wqo (bqo)" by imitating Definition 6.3.

The strength of Fraïssé's conjecture is one of the most important open problems about the reverse mathematics of wqo and bqo theory. All known proofs of Fraïssé's conjecture actually establish Laver's theorem. Basically only one proof was known until 2016: this proof uses the minimal bad array lemma and can be carried out in $\mathbf{\Pi}^1_2$-$\mathsf{CA}_0$. Recently Montalbán [47] made a major breakthrough by finding a new proof, which avoids any form of "minimal bad" arguments. This proof is based on Montalbán's earlier analysis of Fraïssé's conjecture [44] and uses the Ramsey property for subsets of $[\mathbb{N}]^{\mathbb{N}}$ and determinacy, yielding the following result.

**Theorem 6.28** $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$ *proves Fraïssé's conjecture and Laver's theorem.*

Montalbán defines $\mathbf{\Delta}^0_2$-bqo by using $\mathbf{\Delta}^0_2$ functions in Simpson's definition of bqo given after Definition 3.5. Using the fact that $\mathbf{\Sigma}^0_2$ sets are Ramsey (which is known to be equivalent to $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$), he then shows that this notion is equivalent to bqo. Within $\mathsf{ATR}_0$, using $\mathbf{\Delta}^0_1$-determinacy (which is equivalent to $\mathsf{ATR}_0$) Montalbán proves that if 3 is $\mathbf{\Delta}^0_2$-bqo then Laver's theorem holds. Since 3 is bqo is provable in $\mathsf{ATR}_0$ by item *(2)* of Theorem 3.10 the proof of Theorem 6.28 in $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$ is then complete.

Theorem 6.23 entails that Fraïssé's conjecture (and a fortiori Laver's theorem) implies $\mathsf{ATR}_0$. Moreover Fraïssé's conjecture is a $\mathbf{\Pi}^1_2$ statement and the usual considerations yield that $\mathsf{ATR}_0$ plus Fraïssé's conjecture cannot imply $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$. Montalbán's proof shows that to prove Fraïssé's conjecture in any theory weaker than $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$ it suffices to prove that 3 is $\mathbf{\Delta}^0_2$-bqo. Thus an unexpected connection with Question 3.11 comes up. Indeed, Montalbán shows that by mimicking the proof of item *(1)* of Theorem 3.10 it is easy to see that $\mathsf{ATR}_0$ proves that 2 is $\mathbf{\Delta}^0_2$-bqo.

**Question 6.29** Is Fraïssé's conjecture equivalent to $\mathsf{ATR}_0$? Is "3 is $\mathbf{\Delta}^0_2$-bqo" provable in $\mathsf{ATR}_0$?

A couple more results about Fraïssé's conjecture are worth mentioning. First, Montalbán [44] showed that Fraïssé's conjecture is equivalent, over $\mathsf{RCA}_0$ plus $\mathbf{\Sigma}^1_1$-induction, to a result about countable linear orders known as Jullien's theorem. Therefore if the answer to Question 6.29 is negative then Fraïssé's conjecture defines a system intermediate between $\mathsf{ATR}_0$ and $\mathbf{\Pi}^1_1$-$\mathsf{CA}_0$ which is equivalent to other mathematical theorems.

On the other hand, Marcone and Montalbán [40] studied the restriction of Fraïssé's conjecture to linear orders of finite Hausdorff rank. To state the result recall that $\mathsf{ACA}_0^+$ and $\mathsf{ACA}_0'$ are obtained by adding to $\mathsf{RCA}_0$ respectively "for every $X$, $X^{(\omega)}$ (the arithmetic jump of $X$) exists" and "for every $X$ and $k$, $X^{(k)}$ exists". $\mathsf{ACA}_0^+$ is strictly weaker than $\mathsf{ATR}_0$ but strictly stronger than $\mathsf{ACA}_0'$, which in turn is strictly stronger than $\mathsf{ACA}_0$. The ordinal $\varphi_2(0)$ is the first fixed point of the $\varepsilon$ function: in $\mathsf{RCA}_0$ we can define a linear order representing this ordinal, but showing that it is a well-order requires much stronger theories, since this is the proof-theoretic ordinal of $\mathsf{ACA}_0^+$.

**Theorem 6.30** $\mathsf{ACA}_0^+$ *plus "$\varphi_2(0)$ is a well-order" proves the restriction of Fraïssé's conjecture to linear orders of finite Hausdorff rank, which in turn implies, over* $\mathsf{RCA}_0$*,* $\mathsf{ACA}_0'$ *plus "$\varphi_2(0)$ is a well-order".*

## 7 A Topological Version of wqos

Recall that $Q$ wqo does not imply that $\mathcal{P}(Q)$ with respect to $\preceq^\flat$ or $\mathcal{P}_f(Q)$ with respect to $\preceq^\sharp$ is wqo. However we can still draw some conclusions about these partial orders if we weaken the conclusion, using a topological notion.

If $(Q, \preceq)$ is quasi-order we can use $\preceq$ to define a number of different topologies on $Q$. These include the *Alexandroff topology* (whose closed sets are the initial intervals of $Q$) and the *upper topology* (whose basic closed sets are of the form $\{x \in Q : \exists y \in F\, x \preceq y\}$ for $F \subseteq Q$ finite). The topological notion that turns out to be relevant is the following: a topological space is *Noetherian* if it contains no infinite strictly descending sequences of closed sets. It turns out that $Q$ is wqo if and only if the Alexandroff topology on $Q$ is Noetherian, and that $Q$ wqo implies that the upper topology on $Q$ is Noetherian. Goubault-Larrecq [18] proved that $Q$ wqo does imply that the upper topologies of $\mathcal{P}(Q)$ with respect to both $\preceq^\flat$ and $\preceq^\sharp$ are Noetherian. Frittaion, Hendtlass, Marcone, Shafer, and Van der Meeren [16] studied these results from the viewpoint of reverse mathematics, providing along the way proofs that have a completely different flavor from the category-theoretic arguments used by Goubault-Larrecq.

Before describing the results from [16] we need to explain the set-up, which in this case is not obvious because it is necessary to formalize statements about topological spaces which do not fit in the frameworks usually considered in subsystems of second-order arithmetic. (If $Q$ is not an antichain then the Alexandroff and upper topologies are not $T_1$ and are thus very different from complete separable metric spaces.) First notice that if the quasi-order $Q$ is countable the Alexandroff and upper topology can be defined in $\mathsf{RCA}_0$ within the framework of countable second-countable spaces introduced by Dorais [7]. Expressing the fact that a countable second-countable space is Noetherian, as well as the connection mentioned above between $Q$ wqo and the fact that these topologies are Noetherian are also straightforward in $\mathsf{RCA}_0$. However this still does not suffice to tackle all of Goubault-Larrecq's results, because some of them deal with topologies defined on the uncountable space $\mathcal{P}(Q)$. To express that the upper topology of $\mathcal{P}(Q)$ with respect to either $\preceq^\flat$ or $\preceq^\sharp$ is Noetherian, the authors of [16] devise a way of representing these topological spaces. This representation shares some features with other well-established representations of topological spaces, including the familiar separable complete metric spaces and the countably based MF spaces introduced by Mummert [48]. In this set-up the main results are the following ([16, Theorem 4.7]).

**Theorem 7.1** ($\mathsf{RCA}_0$) *The following are equivalent:*

*(i)* $\mathsf{ACA}_0$;
*(ii)* *if $Q$ is wqo, then the Alexandroff topology of $\mathcal{P}_f(Q)$ with respect to $\preceq^\flat$ is Noetherian;*
*(iii)* *if $Q$ is wqo, then the upper topology of $\mathcal{P}_f(Q)$ with respect to $\preceq^\flat$ is Noetherian;*
*(iv)* *if $Q$ is wqo, then the upper topology of $\mathcal{P}_f(Q)$ with respect to $\preceq^\sharp$ is Noetherian;*
*(v)* *if $Q$ is wqo, then the upper topology of $\mathcal{P}(Q)$ with respect to $\preceq^\flat$ is Noetherian;*
*(vi)* *if $Q$ is wqo, then the upper topology of $\mathcal{P}(Q)$ with respect to $\preceq^\sharp$ is Noetherian.*

In [19, Sect. 9.7] Goubault-Larrecq supports his claim that Noetherian spaces can be thought of as topological versions of wqos, by proving the following results. Starting from a topological space $X$ he introduces topologies on $X^{<\mathbb{N}}$ and $\mathcal{T}^X$ and proves the topological versions of Higman's and Kruskal's theorems, stating that if $X$ is Noetherian then $X^{<\mathbb{N}}$ and $\mathcal{T}^X$ are Noetherian. If $X$ is a countable second-countable space then so are $X^{<\mathbb{N}}$ and $\mathcal{T}^X$, which leads to the following so far unexplored question.

**Question 7.2** What is the strength of the topological versions of Higman's and Kruskal's theorems restricted to countable second-countable spaces?

## References

1. Blass, A., & Gurevich, Y. (2008). Program termination and well partial orderings. *ACM Transactions on Computational Logic*, *9*(3), Art. 18, 26.
2. Bonnet, R. (1975). On the cardinality of the set of initial intervals of a partially ordered set. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his* 60th *birthday), Vol. I*, pp. 189–198. Colloq. Math. Soc. János Bolyai, Vol. 10. North-Holland, Amsterdam.
3. Cholak, P., Marcone, A., & Solomon, R. (2004). Reverse mathematics and the equivalence of definitions for well and better quasi-orders. *The Journal of Symbolic Logic*, *69*(3), 683–712.
4. Clote, P. (1990). The metamathematics of Fraïssé's order type conjecture. In K. Ambos-Spies, G. H. Müller, & G. E. Sacks (Eds.), *Recursion theory week* (pp. 41–56). Berlin: Springer.
5. Dean, W., & Walsh, S. (2017). The prehistory of the subsystems of second-order arithmetic. *Review of Symbolic Logic*, *10*(2), 357–396.
6. Diestel, R. (2017). *Graph theory*, volume 173 of *Graduate texts in mathematics*. Springer, Heidelberg, 5th edn.
7. Dorais, F. G. (2011). Reverse mathematics of compact countable second-countable spaces. arXiv:1110.6555.
8. Downey, R. G. (1998). Computability theory and linear orderings. In *Handbook of Recursive Mathematics, Vol. 2*, volume 139 of *Studies in Logic and the Foundations of Mathematics*, pp. 823–976. North-Holland, Amsterdam.
9. Erdös, P., & Tarski, A. (1943). On families of mutually exclusive sets. *Annals of Mathematics*, *2*(44), 315–329.
10. Fraïssé, R. (1948). Sur la comparaison des types d'ordres. *C. R. Academy of Sciences Paris*, *226*, 1330–1331.
11. Fraïssé, R. (2000). *Theory of relations*, volume 145 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, revised edition, 2000. With an appendix by Norbert Sauer.

12. Friedman, H. (1975). Some systems of second order arithmetic and their use. In *Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 1*, pp. 235–242. Canad. Math. Congress, Montreal, Que.
13. Friedman, H. (1976). Systems of second order arithmetic with restricted induction, i, ii (abstracts). *The Journal of Symbolic Logic*, *41*(2), 557–559.
14. Friedman, H., Robertson, N., & Seymour, P. (1987). The metamathematics of the graph minor theorem. In S. G. Simpson (Ed.), *Logic and combinatorics* (pp. 229–261). Providence, R.I: American Mathematical Society.
15. Frittaion, E., & Marcone, A. (2014). Reverse mathematics and initial intervals. *Annals of Pure and Applied Logic*, *165*(3), 858–879.
16. Frittaion, E., Hendtlass, M., Marcone, A., Shafer, P., & Van der Meeren, J. (2016). Reverse mathematics, well-quasi-orders, and Noetherian spaces. *Review of Symbolic Logic*, *55*(3–4), 431–459.
17. Girard, J.-Y. (1987). *Proof theory and logical complexity*. Naples: Bibliopolis.
18. Goubault-Larrecq, J. (2007). On Noetherian spaces. In *Proceedings of the 22nd Annual IEEE Symposium on Logic in Computer Science (LICS'07)*, pages 453–462, Wrocław, Poland, July 2007. IEEE Computer Society Press.
19. Goubault-Larrecq, J. (2013). *Non-Hausdorff topology and domain theory* (Vol. 22). New Mathematical Monographs Cambridge: Cambridge University Press.
20. Harzheim, E. (2005). *Ordered sets* (Vol. 7), Advances in Mathematics (Springer) New York: Springer.
21. Hatzikiriakou, K., & Simpson, S. G. (2017). Reverse mathematics, Young diagrams, and the ascending chain condition. *The Journal of Symbolic Logic*, *82*(2), 576–589.
22. Higman, G. (1952). Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society*, *3*(2), 326–336.
23. Hilbert, D., & Bernays, P. (1968). *Grundlagen der Mathematik I*. Berlin: Springer.
24. Hilbert, D., & Bernays, P. (1970). *Grundlagen der Mathematik II*. Berlin: Springer.
25. Hirschfeldt, D. R. (2015). *Slicing the truth*, volume 28 of *Lecture Notes Series. Institute for Mathematical Sciences. National University of Singapore*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2015. On the computable and reverse mathematics of combinatorial principles, Edited and with a foreword by Chitat Chong, Qi Feng, Theodore A. Slaman, W. Hugh Woodin and Yue Yang.
26. Hirschfeldt, D. R., & Shore, R. A. (2007). Combinatorial principles weaker than Ramsey's theorem for pairs. *The Journal of Symbolic Logic*, *72*(1), 171–206.
27. Hirst, J. L. (1994). Reverse mathematics and ordinal exponentiation. *Annals of Pure and Applied Logic*, *66*(1), 1–18.
28. Jongh, D. H. J. de, & Parikh, R. (1977). Well-partial orderings and hierarchies. *Nederl. Akad. Wetensch. Proc. Ser. A* **80**=*Indag. Math.*, *39*(3), 195–207.
29. Kříž, I., & Thomas, R. (1990). Ordinal types in Ramsey theory and well-partial-ordering theory. In *Mathematics of Ramsey theory*, vol. 5 of *Algorithms Combin.*, pp. 57–95. Springer, Berlin.
30. Krombholz, M. (2018). *Proof theory of graph minors and tree embeddings*. Ph.d. thesis, University of Leeds, 2018.
31. Krombholz, M., & Rathjen, M., *Upper bounds on the graph minor theorem*. arXiv:1907.00412.
32. Kruskal, J. B. (1960). Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture. *Transactions of the American Mathematical Society*, *95*, 210–225.
33. Laver, R. (1971). On Fraïssé's order type conjecture. *Annals of Mathematics*, *2*(93), 89–111.
34. Laver, R. (1978). Better-quasi-orderings and a class of trees. In G.-C. Rota (Ed.), *Studies in Foundations and Combinatorics* (pp. 31–48). New York: Academic Press.
35. Lerman, M., Solomon, R., & Towsner, H. (2013). Separating principles below Ramsey's theorem for pairs. *Journal of Mathematical Logic*, *13*(2), 1350007, 44.
36. Liu, J. (2012). $RT^2_2$ does not imply $WKL_0$. *The Journal of Symbolic Logic*, *77*(2), 609–620.
37. Marcone, A. (1994). Foundations of BQO theory. *Transactions of the American Mathematical Society*, *345*(2), 641–660.

38. Marcone, A. (1996). On the logical strength of Nash-Williams' theorem on transfinite sequences. In W. Hodges, M. Hyland, C. Steinhorn, & J. Truss (Eds.), *Logic: from foundations to applications* (pp. 327–351). New York: The Clarendon Press.
39. Marcone, A. (2005). Wqo and bqo theory in subsystems of second order arithmetic. *Reverse Mathematics 2001* (Vol. 21, pp. 303–330)., Lecture Notes in Mathematics La Jolla, CA: Assoc. Symbol. Logic.
40. Marcone, A., & Montalbán, A. (2009). On Fraïssé's conjecture for linear orders of finite Hausdorff rank. *Annals of Pure and Applied Logic*, *160*(3), 355–367.
41. Marcone, A., & Montalbán, A. (2011). The Veblen functions for computability theorists. *The Journal of Symbolic Logic*, *76*(2), 575–602.
42. Marcone, A., & Shore, R. A. (2011). The maximal linear extension theorem in second order arithmetic. *Archive for Mathematical Logic*, *50*(5–6), 543–564.
43. Marcone, A., Montalbán, A., & Shore, R. A. (2012). Computing maximal chains. *Archive for Mathematical Logic*, *51*(5–6), 651–660.
44. Montalbán, A. (2006). Equivalence between Fraïssé's conjecture and Jullien's theorem. *Annals of Pure and Applied Logic*, *139*(1–3), 1–42.
45. Montalbán, A. (2007). Computable linearizations of well-partial-orderings. *Order*, *24*(1), 39–48.
46. Montalbán, A. (2011). Open questions in reverse mathematics. *Bulletin of Symbolic Logic*, *17*(3), 431–454.
47. Montalbán, A. (2017). Fraïssé's conjecture in $\Pi^1_1$-comprehension. *International Journal of Mathematics*, *17*(2), 1750006, 12.
48. Mummert, C. (2006). Reverse mathematics of MF spaces. *Journal of Mathematical Logic*, *6*(2), 203–232.
49. Nash-Williams, C. St, & J. A., (1963). On well-quasi-ordering finite trees. *Proceedings of the Cambridge Philosophical Society*, *59*(833–835), 1963.
50. Nash-Williams, C. St, & J. A.,. (1965). On well-quasi-ordering infinite trees. *Proceedings of the Cambridge Philosophical Society*, *61*, 697–720.
51. Nash-Williams, C. St, & J. A.,. (1968). On better-quasi-ordering transfinite sequences. *Proceedings of the Cambridge Philosophical Society*, *64*, 273–290.
52. Patey, L. (2016). Partial orders and immunity in reverse mathematics. In *Pursuit of the Universal*, volume 9709 of Lecture Notes in Computer Science, pp. 353–363. Springer, [Cham].
53. Rathjen, M., & Weiermann, A. (1993). Proof-theoretic investigations on Kruskal's theorem. *Annals of Pure and Applied Logic*, *60*(1), 49–88.
54. Rathjen, M., & Weiermann, A. (2011). Reverse mathematics and well-ordering principles. *Computability in Context* (pp. 351–370). London: Imperial College Press.
55. Schmidt, D. (1979). *Well-partial Orderings and their maximal order types*. Habilitationschrift: Heidelberg University.
56. Schütte, K., & Simpson, S. G. (1985). Ein in der reinen Zahlentheorie unbeweisbarer Satz über endliche Folgen von natürlichen Zahlen. *Archive for Mathematical Logik Grundlag*, *25*(1–2), 75–89.
57. Seetapun, D., & Slaman, T. A. (1995). On the strength of Ramsey's theorem. *Notre Dame Journal of Formal Logic*, *36*(4), 570–582, 1995. Special Issue: Models of arithmetic.
58. Shore, R. A. (1993). On the strength of Fraïssé's conjecture. In J. N. Crossley, J. B. Remmel, R. A. Shore, & M. E. Sweedler (Eds.), *Logical methods* (pp. 782–813). Boston, MA: Birkhäuser.
59. Simpson, S. G., & (1985). Bqo-theory and Fraïssé's conjecture. Chapter 9, Mansfield, R., & Weitkamp, G.,. (1985). *Recursive aspects of descriptive set theory*. New York: Oxford University Press.
60. Simpson, S. G. (1985). Nonprovability of certain combinatorial properties of finite trees. In L. A. Harrington, M. D. Morley, A. Ščedrov, & S. G. Simpson (Eds.), *Harvey Friedman's research on the foundations of mathematics* (pp. 87–117). Amsterdam: North-Holland.
61. Simpson, S. G. (1988). Ordinal numbers and the Hilbert basis theorem. *The Journal of Symbolic Logic*, *53*(3), 961–974.

62. Simpson, S. G. (2009). *Subsystems of second order arithmetic*. Perspectives in Logic. Cambridge University Press, Cambridge; Association for Symbolic Logic, Poughkeepsie, NY, second edition.
63. Thomassen, C. (1995). Embeddings and minors. In R. L. Graham, M. Grötschel, & L. Lovász (Eds.), *Handbook of combinatorics* (Vol. 1, pp. 301–349). Amsterdam: Elsevier Science.
64. Towsner, H. (2013). Partial impredicativity in reverse mathematics. *The Journal of Symbolic Logic*, *78*(2), 459–488.
65. Towsner, H. (2016). Constructing sequences one step at a time. arXiv:1609.05509.
66. Weyl, H. (1918). *Das Kontinuum: Kritische Untersuchungen über die Grundlagen der Analysis*. Leipzig: Veit.
67. Wolk, E. S. (1967). Partially well ordered sets and partial ordinals. *Fundamenta Mathematicae*, *60*, 175–186.

# Well Quasi-orders and the Functional Interpretation

**Thomas Powell**

**Abstract** The purpose of this article is to study the role of Gödel's functional interpretation in the extraction of programs from proofs in well quasi-order theory. The main focus is on the interpretation of Nash–Williams' famous minimal bad sequence construction, and the exploration of a number of much broader problems which are related to this, particularly the question of the constructive meaning of Zorn's lemma and the notion of recursion over the non-wellfounded lexicographic ordering on infinite sequences.

## 1 Introduction

In this article, my goal is to write something that reflects the extraordinary richness of the theory of well quasi-orders. The reader familiar with the well quasi-order will have observed first hand how this rather innocent looking mathematical object plays a crucial role in many seemingly disparate areas, ranging from proof theory, computability theory and reverse mathematics on the one hand to to term rewriting, program verification and the world of automata and formal languages on the other. While I could never do justice to such diversity in one paper, my aim is to at least explore a variety of interesting problems in my own field which arise from the study of well quasi-orders, and to present some new research in this direction.

The basis of this article is Gödel's functional interpretation, and the role it plays in making constructive sense of well quasi-orders. I have chosen to organise what follows around a somewhat superficial challenge, namely the development of a program which realizes Higman's lemma for boolean alphabets:

T. Powell (✉)
Fachbereich Mathematik, Technische Universität Darmstadt, Schlossgartenstraße 7,
64289 Darmstadt, Germany
e-mail: powell@mathematik.tu-darmstadt.de



P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_9

PROBLEM. Write a program $\Phi$ which takes as input an infinite sequence $u$ of words over a two letter alphabet, and returns a pair of indices $i < j \in \mathbb{N}$ such that $u_i$ is embedded in $u_j$.

Of course, as long as one has proven that two such indices must exist one could simply write a program which carries out a blind search until they are found! However, I am interested in the question of how one can formally construct a subrecursive program which constitutes a computational analogue of Nash–Williams' famous minimal bad sequence construction—an elegant combinatorial idea which appears throughout well quasi-order theory.

It is important to stress that this relatively simple problem provides merely a narrative framework: My ulterior motive is to explore a number of much more elusive problems which lurk underneath. So while on the surface we will work towards the construction of our program $\Phi$, the real contribution of the work is to address several deeper questions, chief among them being:

1. What is the computational meaning of Zorn's lemma?
2. Is it possible to sensibly define recursive functionals on chain-complete partial orders?
3. How can one describe formally extracted programs so that they can be easily understood by a human?

Each of these questions has significance far beyond Higman's lemma, and yet the fact that they are all naturally prompted by our elementary problem is, I believe, testament to the richness inherent to the theory of well quasi-orders.

## 1.1 Proof Interpretations and Well Quasi-Orders: A Brief History

In 1958, Gödel published a landmark paper [10] which introduced his functional, or 'Dialectica' interpretation, which he had already conceived in the 1930s as a response to Hilbert's program and his own incompleteness theorems. Initially, the functional interpretation translated Peano arithmetic to a calculus of primitive recursive functionals in all finite types known as System $\mathsf{T}$, thereby reducing the consistency of the former theory to the latter. In modern day parlance, System $\mathsf{T}$ is nothing more than a simple functional programming language which permits the construction of higher-type primitive recursive functionals. The soundness of the functional interpretation guarantees that, whenever some statement $A$ is provable in Peano arithmetic, we can extract a total functional program in $\mathsf{T}$ which witnesses its translation $A^I$.

The functional interpretation was just one of a number of techniques designed during the mid 20th century to establish relative consistency proofs. Kreisel soon observed that these techniques could be flipped on their head and viewed from a different perspective: namely as tools for extracting computational information from non-constructive proofs [14, 15]. While the significance of this idea was not fully appreciated at the time, in recent decades the application of proof theoretic methods to extract programs from proofs has flourished, and now proof interpretations are primarily used for this purpose. Variants of Gödel's functional interpretation in particular are central to the highly successful 'proof mining' program pioneered by Kohlenbach [13], which has led to new quantitative results in several areas of mathematics. At the same time, the arrival of the computer has meant that the extraction of programs from proofs can be automated, and there are now proof assistants such as MINLOG [1] which are dedicated to this, and which implement sophisticated refinements of the traditional proof theoretic techniques.

So where do well quasi-orders feature in all of this?

The vast majority of proofs in 'normal' mathematics use only a very small amount of set theory. Often, proofs of existential theorems in mathematics analysis which officially require choice or comprehension, use it in such a limited way that it doesn't really contribute to the complexity of extracted programs. However, the theory of well quasi-orders contains a number of key theorems which *do* use choice in a crucial way, the most notorious being those such Kruskal's theorem which historically rely on variant of Nash–Williams' minimal bad sequence construction.

As a result, these theorems have become something of a focal point for research in program extraction, as canonical existential statements which come with concise, elegant, but proof theoretically non-trivial classical proofs. The question of the computational meaning of such proofs is so deep that entire doctoral theses have been dedicated to it (such as [18, 26]). By now, even comparatively simple results like Higman's lemma have an extensive body of research devoted to them. Thus the theory of well quasi-orders has firmly established a foothold in the world of proof theory, and it is from this perspective that I study them here.

### *1.2 The Origins and Purpose of This Chapter*

Given the popularity of Higman's lemma among researchers in proof theory, it's perhaps important to outline my own motivation in adding yet another paper to this menagerie.

My interest in well quasi-orders began when I was a doctoral student studying Gödel's functional interpretation. Paulo Oliva suggested to me that Higman's lemma might prove a useful exercise in program extraction via the functional interpretation, as up to that point this had never been done: The majority of attempts at giving a constructive proof of the lemma had utilised some form of realizability instead. So I undertook this challenge and published my work as [21].

While this indeed turned out to be a valuable exercise, improving my own understanding of the functional interpretation and providing me with a welcome excuse to learn about well quasi-orders, in most other respects I found my work rather unsatisfactory. In order to give a computational interpretation of the instance of dependent choice used in the proof of the theorem, I resorted to the standard technique at one's disposal—a higher-type form of bar recursion. But due to the subtlety of Nash–Williams' construction, the resulting instance of bar recursion is extremely complex, leading to an extracted term whose *operational* behaviour as a program is somewhat obscure, to say the least! While after a certain amount of effort I began to see what the underlying program did, this was still very difficult to describe, and I doubt that anyone who has read [21] will have gained any meaningful insight into the computational meaning of Higman's lemma!

I believe that the shortcomings of this paper were partly due to my own inexperience at the time, and partly due to the fact that the basic technology for extracting programs from proofs is largely unchanged since its introduction over half a century ago. While admittedly a range of refinements have been developed, and proof mining in particular has produced a number of extremely powerful metatheorems which guarantee the extractability of low-complexity programs from proofs in specific areas of analysis, these do not really help us when comes to non-constructive proofs in well quasi-order theory which rely in an essential way on dependent choice.

In the years that followed I ended up thinking about much more general problems which were prompted from my analysis of Higman's lemma. In particular, I studied forms of higher-order recursion closely related to Nash–Williams' construction [22], and tried to develop notation systems which allow one to describe extracted programs in a more intuitive way [23]. In this article, I take the opportunity to expand my original work in light of these developments, and present some new results in this direction.

As I have already emphasised, the 'official' goal of building a program $\Phi$ which witnesses Higman's lemma over boolean alphabets is nothing more than an organisational device. Indeed, this is by no means the first place in which such a program has been presented, and I reiterate that real content of this essay lies in the methods which we use to obtain it, and the series of theoretical results presented in Sects. 7 and 8, particularly Theorems 7.8 and 8.3, which I publish here for the first time.

Much of the technical groundwork I will present here has been done elsewhere, and this allows me to adopt a lighter style of presentation, in which my priority will be to stress the key points and skim over the heavier details. I have nevertheless tried to keep everything as self-contained as possible. So, for example, the reader not familiar with Gödel's functional interpretation will be given the main definition and plenty of intuition on what it means, and should be able to follow later sections without too much confusion, although whenever a key concept is introduced I take care to include references to introductory material in which a more extensive presentation is given.

On the other hand, the expert reader may wish to skip straight ahead to later sections in which the main technical contributions are presented, and so I have indicated whenever a section is comprised mostly of background material.

In the area of program extraction, it is not uncommon to see technical achievements presented with few examples to illustrate them, and concrete case studies which give little insight into the underlying techniques on which they are based (and I have certainly been guilty of both of these at one point or another!). But my aim here is to endeavour to strike a balance between both theory and practice, and as a result I hope that this article will form a pleasant read for both specialists in proof theory as well as those with a more general interest in well quasi-orders.

## 2 Well Quasi-orders and Zorn's Lemma

We begin with the definition of a well quasi-order. There are numerous equivalent formulations of this concept—one of the simplest and most widely seen being the following:

**Definition 2.1** A quasi-order $(X, \preceq)$ is a set $X$ equipped with a binary relation $\preceq$ which is reflexive and transitive. It is a well quasi-order (or WQO) if it satisfies the additional property that for any infinite sequence of elements $x_0, x_1, x_2, \ldots$ there exists some $i < j$ such that $x_i \preceq x_j$.

It is not difficult to see that a quasi-order is a WQO iff it contains no infinite strictly decreasing chains and no infinite sequences of pairwise incomparable elements. Therefore being a WQO is a strictly stronger property than being well-founded. For example, the quasi-order $(\mathbb{N}, \mid)$ of natural numbers ordered by divisibility is well-founded, but not a WQO. The following is perhaps slightly less obvious:

**Lemma 2.2** *A quasi-order $(X, \preceq)$ is a WQO iff any infinite sequence $x_0, x_1, x_2, \ldots$ contains an infinite increasing subsequence $x_{g(0)} \preceq x_{g(1)} \preceq x_{g(2)} \preceq \ldots$ (where $g(0) < g(1) < \ldots$).*

*Proof* For the non-trivial direction, let $(X, \preceq)$ be a WQO, and take some infinite sequence $x_0, x_1, \ldots$. Define $T \subseteq \mathbb{N}$ by $T :\equiv \{i \in \mathbb{N} \mid (\forall j > i)\neg(x_i \preceq x_j)\}$. Then $T$ must be finite, otherwise we would be able to construct a sequence contradicting the assumption that $X$ is a WQO. Therefore there is some $N \in \mathbb{N}$ such that for all $i \geq N$ there exists some $j > i$ with $x_i \preceq x_j$, which allows us to construct our infinite increasing subsequence. □

Given some mathematical property, such as being well quasi-ordered, we are often interested in identifying constructions which preserve that property. WQO theory is particularly rich in such results. A simple example is the following:

**Proposition 2.3** *If $(X, \preceq_X)$ and $(Y, \preceq_Y)$ are WQOs, then so is their cartesian product $(X \times Y, \preceq_{X\times Y})$ under the pointwise ordering.*

*Proof* Given an infinite sequence $\langle x_0, y_0\rangle, \langle x_1, y_1\rangle, \ldots$, consider the first component $x_0, x_1, \ldots$. Since $X$ is a WQO, by Lemma 2.2 there exists an infinite increasing sequence $x_{g(0)} \preceq_X x_{g(1)} \preceq_X \ldots$. Now consider the sequence $y_{g(0)}, y_{g(1)}, \ldots$. Since $Y$ is a WQO, there exists some $i < j$ with $y_{g(i)} \preceq_Y y_{g(j)}$. But by transitivity we also have $x_{g(i)} \preceq x_{g(j)}$, and therefore $\langle x_{g(i)}, y_{g(i)}\rangle \preceq_{X\times Y} \langle x_{g(j)}, y_{g(j)}\rangle$. □

A far more subtle result, which forms the basis of this article, is the following theorem, widely known as *Higman's lemma*:

**Theorem 2.4** (Higman's lemma [11]) *If* $(X, \preceq)$ *is a WQO, then so is* $(X^*, \preceq_*)$*, the set of finite sequences over* $X$ *ordered under the embeddability relation, where* $[x_0, \ldots, x_{m-1}] \preceq_* [y_0, \ldots, y_{n-1}]$ *whenever there is a strictly increasing map* $f$ *with* $x_i \preceq y_{f(i)}$ *for all* $i < m$.

A short and extremely elegant proof of Higman's lemma was given by Nash–Williams, using the so-called minimal bad sequence construction, which is a central topic of our paper.

*Proof of Higman's lemma* [19]. Suppose for contradiction that $X$ is a WQO but that there exists an infinite sequence of words $u_0, u_1, \ldots$ such that $\neg(u_i \preceq_* u_j)$ for all $i < j$. We call such a sequence a 'bad sequence'. Now, using the axiom of dependent choice, pick a *minimal* bad sequence $v_0, v_1, \ldots$ as follows:

> Given that we have already constructed $[v_0, \ldots, v_{k-1}]$, define $v_k$ to be such that $[v_0, \ldots, v_{k-1}, v_k]$ extends to some infinite bad sequence, but $[v_0, \ldots, v_{k-1}, a]$ does not for any $a \lhd v_k$, by which mean any strict prefix $a$ of $v_k$.

Note that such a $v_k$ exists by the minimum principle over the wellfounded prefix relation $\lhd$, together with the fact that $[v_0, \ldots, v_{k-1}]$ must extend to *some* bad sequence: For $k = 0$ this follows from our assumption that least one bad sequence exists, while for $k > 0$ it is true by construction.

Now, the crucial point is that this minimal sequence must itself be bad: If instead there were some $i < j$ with $v_i \preceq_* v_j$, then $[v_0, \ldots, v_j]$ could not extend to a bad sequence, contradicting our construction. Therefore in particular each $v_n$ must be non-empty, otherwise we would trivially have $v_n = [] \preceq_* v_{n+1}$. This means that each $v_n$ must be a concatenation of the form $\tilde{v}_n * \bar{v}_n$ where $\tilde{v}_n \in X^*$ and $\bar{v}_n \in X$. By Lemma 2.2 the sequence $\bar{v}_0, \bar{v}_1, \ldots$ contains some increasing subsequence $\bar{v}_{g0} \preceq \bar{v}_{g1} \preceq \ldots$, so let's now consider the sequence

$$w := v_0, \ldots, v_{g0-1}, \tilde{v}_{g0}, \tilde{v}_{g0+1}, \tilde{v}_{g0+2}, \ldots$$

Since $\tilde{v}_{g0} \lhd v_{g0}$, by minimality of $v$ the sequence $w$ must be good, which means that $w_i \preceq_* w_j$ for some $i < j$. There are three possibilities: First $j < g0$ and so $v_i = w_i \preceq_* w_j = v_j$, second $i < g0$ and $j = gj'$ and so $v_i \preceq_* \tilde{v}_{gj'}$ which implies that $v_i \preceq_* v_{gj'}$ since $\tilde{v}_{gj'} \lhd v_{gj'}$, and finally $i, j = gi', gj'$ and so $\tilde{v}_{gi'} \preceq_* \tilde{v}_{gj'}$ which implies that $v_{gi'} \preceq_* v_{gj'}$ since $\bar{v}_{gi'} \preceq \bar{v}_{gj'}$. In all cases we have $v_i \preceq_* v_j$, contradicting the fact that $v$ is bad. Hence our original assumption was false, and we can conclude that there are no bad sequence, or equivalently that $X^*$ is a WQO. □

As an immediate consequence of Higman's lemma, we see that our main problem can, in theory, be solved:

**Corollary 2.5** *Given an infinite sequence* $u$ *of words over a two letter alphabet* $\{0, 1\}$*, there exists a pair of indices* $i < j$ *such that* $u_i$ *is embedded in* $u_j$.

*Proof* The set $(\{0, 1\}, =)$ trivially a WQO, therefore by Higman's lemma so is $(\{0, 1\}^*, =_*)$. □

## 2.1 The Minimal Bad Sequence Construction and Zorn's Lemma

The existence of a minimal bad sequence in Nash–Williams' proof of Higman's lemma can be viewed in a much broader context as a particular instance of Zorn's lemma, or equivalently, as an inductive principle over chain-complete partial orders. This was first observed by Raoult [24], and since it informs our approach to program extraction, we will explain in a little more detail what is meant by this.

Suppose that $(Y, \sqsupseteq)$ is a chain-complete partial order, where for each non-empty chain $\gamma$ in $Y$ we fix some lower bound $\bigwedge \gamma$, which is usually taken to be the greatest lower bound if it exists (note that the fact that we talk about lower rather than upper bounds is purely cosmetic, as it sounds slightly more natural when generalising the notion of a *minimal* bad sequence). The following result is essentially just the contrapositive of the principle of open induction discussed in [24]:

**Proposition 2.6** *Let $B$ be a predicate on $Y$ which satisfies the property that for any* non-empty *chain $\gamma$,*

$$(\forall x \in \gamma) B(x) \to B\left(\bigwedge \gamma\right). \tag{1}$$

*Then whenever $B(x)$ holds for some $x \in Y$, there is some minimal $y$ such that $B(y)$ holds, but $y \sqsupset z \to \neg B(z)$.*

*Proof* Define $S :\equiv \{x \in Y \mid B(x)\}$. Then whenever $B(x)$ holds for some $x$, the set $S$ is chain complete: For the empty chain we just take $x$ as a lower bound, while any non-empty chain $\gamma$ in $S$ we have that $\bigwedge \gamma \in S$ by (1). Therefore by Zorn's lemma $S$ has some minimal element $y$. □

Now, consider some set $X$ which comes equipped with a given strict partial order $\lhd$ on $X$ which is wellfounded. Define the lexicographic extension $\lhd_{\text{lex}}$ of $\lhd$ by

$$u \lhd_{\text{lex}} v \text{ iff } (\exists n)([u](n) = [v](n) \wedge u_n \lhd v_n)$$

where $u$ and $v$ are infinite sequences of type $X^{\mathbb{N}}$ and $[u](n) := [u_0, \ldots, u_{n-1}]$ denotes the initial segment of $u$ of length $n$. It is easy to show that $\lhd_{\text{lex}}$ is also (strict) a partial order. Note that $\lhd_{\text{lex}}$ is not wellfounded - for example, setting $X := 0, 1$ and defining $0 \lhd 1$ we would have

$$1, 1, 1, \ldots \rhd_{\text{lex}} 0, 1, 1, \ldots \rhd_{\text{lex}} 0, 0, 1, \ldots \rhd_{\text{lex}} \ldots$$

However, $\lhd_{\text{lex}}$ is *chain-complete*. In fact, given a chain $\gamma$ in $(X^{\mathbb{N}}, \rhd_{\text{lex}})$, we can construct its greatest lower bound by defining $v_0 \in X$ to be the minimum with respect to $\rhd$ of the first components of the elements of $\gamma$, then $v_1 \in X$ to be the minimum of the second components of all elements $x \in \gamma$ with $x_0 = v_0$, then $v_2$ to be the minimum of the third components of all elements $x \in \gamma$ with $x_0, x_1 = v_0, v_1$ and so on, and it is not difficult to show that $\bigwedge \gamma := v$ is a greatest lower bound of $\gamma$.

Moreover, this greatest lower bound $v$ has the property that for any $n \in \mathbb{N}$, there is some $x_n \in \gamma$ which agrees with $v$ on the first $n$ elements i.e. $[v](n) = [x_n](n)$. This motivates the following definition:

**Definition 2.7** A formula $B(u)$ on infinite sequences $u \in X^{\mathbb{N}}$ is piecewise definable, or just piecewise, if it can be expressed in the form $(\forall n)P([u](n))$ for some formula $P(s)$ on finite sequences $s \in X^*$.

**Theorem 2.8** *Let $B(u) \equiv (\forall n)P([u](n))$ be a piecewise formula, and suppose that $B(u)$ holds for some $u$. Then there exists some minimal 'bad' sequence $v$ such that $B(v)$ holds, but $\neg B(w)$ for any $w \lhd_{lex} v$.*

*Proof* Take any non-empty chain $\gamma$ such that $(\forall x \in \gamma)B(x)$, and let $v := \bigwedge \gamma$. We want to show that $B(v)$ holds i.e. $P([v](n))$ holds for all $n \in \mathbb{N}$. But as observed above, for any $n$ there exists some $x_n \in \gamma$ with $[v](n) = [x_n](n)$, and $P([x_n](n))$ follows from $B(x_n)$. Therefore the existence of a minimal bad sequence $v$ follows directly from Proposition 2.6. □

Theorem 2.8 is nothing more than a generalisation of the minimal-bad-sequence construction in Nash–Williams' proof of Higman's lemma: The predicate '$u$ is bad' can be expressed as $B(u) :\equiv (\forall n)(\forall i < j < n)\neg(u_i \preceq_* u_j)$ which is clearly a piecewise formula, and so the existence of a minimal bad sequence follows as a special case of the instance of Zorn's lemma given in Proposition 2.6, where $Y := (X^*)^{\mathbb{N}}$ and $\sqsupset$ is taken to be $\rhd_{\text{lex}}$ over the lexicographic extension of the prefix order.

## 2.2 Zorn's Lemma as an Axiom

The reason for the short digression above is to encourage the reader to think of the minimal bad sequence construction, not as a derived result which follows from dependent choice, but as an axiomatic minimum principle over the chain-complete partial order $((X^*)^{\mathbb{N}}, \rhd_{\text{lex}})$ which can be considered a weak form of Zorn's lemma, namely

$$(\exists u)B(u) \to (\exists v)(B(v) \wedge (\forall w \lhd_{\text{lex}} v)\neg B(w)), \tag{2}$$

where $B(u)$ ranges over *piecewise* formulas. Note that in this case, the premise of Zorn's lemma, namely chain-completeness of $S = \{u \in (X^*)^{\mathbb{N}} \mid B(u)\}$, is encoded by both the premise $(\exists u)B(u)$ and the assumption that $B$ is piecewise.

While this slight shift of emphasis from dependent choice to Zorn's lemma might not seem significant from an ordinary mathematical perspective, it completely alters the way in which we give a computational interpretation to Nash–Williams' proof of Higman's lemma, as this depends entirely on the manner in which we choose to formalise that proof. We will discuss proof interpretations and program extraction in much more detail in Sect. 4, but since the difference between dependent choice and above formulation of Zorn's lemma motivates our formal proof in the next section, it is important to at least roughly explain why this distinction matters to us here.

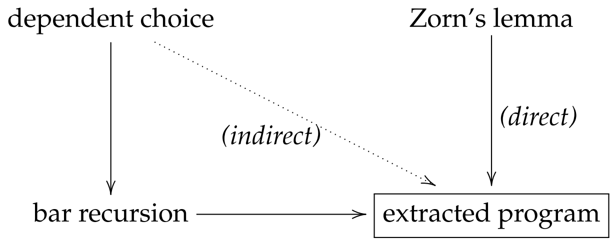
The extraction of programs from proofs typically works by assigning basic programs to the axioms and rules of some mathematical theory, and then constructing general programs recursively over the structure of formal proofs. Thus the programs which interpret the axioms of our theory form our basic building blocks, and the overall size and complexity of an extracted program in terms of these blocks reflects the size and complexity of the formal proof from which it was obtained.

In the early days of proof theory, when proof interpretations were primarily used to obtain relative consistency proofs, ‘extracted programs’ were nothing more than hypothetical objects which gave a computational interpretation to falsity, whose existence within some formal calculus was necessary but whose structure as programs was uninteresting and irrelevant. As such, it was sensible to work in a minimal axiomatic theory which was easy to reason about on a meta-level, but not necessarily convenient for extracting programs in practice. This was the approach taken by Spector [27], who extended Gödel’s consistency proof to full mathematical analysis by showing that countable dependent choice could be interpreted by the scheme of *bar recursion* in all finite types, a form of recursion which, while elegant, can be rather abstruse when it comes to understanding its operational semantics as part of a real program.

For us, on the other hand, a proof interpretation is a *tool* for extracting an *actual* program from the proof of Higman’s lemma whose algorithmic behaviour can be understood to some extent, as opposed to some obscure syntactical object which essentially acts as a black box. As a result, we want to work in a axiomatic system in which Nash–Williams’ minimal bad sequence construction can be cleanly and concisely formalised. So it is natural to ask whether, to this end, we can give a more *direct* proof of Nash–Williams’ construction which circumvents the use of bar recursion, and leads to a more intuitive extracted program.

Our idea will be the following: Instead of taking dependent choice as a basic axiom and using this to prove the existence of a minimal-bad-sequence, we will instead take (2) as a basic axiom, from which the existence of minimal-bad-sequences follows trivially. As a result, though, we will no longer be able to rely on Spector’s computational interpretation of dependent choice, and will have to instead construct a new realizer for (2).

It is now perhaps becoming clear to the reader why the three deeper questions outlined in the introduction emerge naturally from Higman’s lemma! The construction of our direct realizer for the principle (2) carried out in Sects. 6–8 offers a partial solution to Question 1 for the particular instance of Zorn’s lemma used here. Section 7 will focus specifically on the special variant of recursion over $\rhd_{\text{lex}}$ that will be required in order to do all this, and will therefore in turn address Question 2. Question 3 is something that will be on our minds throughout the paper, and in particular influences our description of the realizing term in terms of *learning procedures*.




Before we go on, it is important to observe that the idea of replacing dependent choice with some variant of Zorn's lemma has already been considered by Berger in the setting of modified realizability, in which a variant of Raoult's principle of open induction was given a direct realizability interpretation by a form of open recursion [4]. Here we will give an analogous interpretation for the *functional interpretation* of a principle classically equivalent open induction, and our work here differs considerably from the realizability setting in a number of crucial respects, all of which we make clear later.

## 3 A Formal Proof of Higman's Lemma

As I highlighted above, in order to apply a proof interpretation to a proof, we first need to have some kind of formal representation of this proof in mind. The route from 'textbook' to formal proof is no mere preprocessing step—the structure and hence usefulness of our extracted program is entirely dependent on the way in which we make precise the logical steps encoded by our textbook proof. The power of applied proof theory is due to the fact that the careful analysis of logical subtleties in formal proofs can reveal quantitative information that is not apparent from an ordinary mathematical perspective, hence the frequent characterisation of this information as being 'hidden' in the proof. In our case, as emphasised already, the fact that we will formalise Nash–Williams' proof of Higman's lemma using an axiomatic form of Zorn's lemma is absolutely crucial to our approach.

When it comes to the application of proof interpretations, one encounters two rather distinct styles in the literature. In proof mining as conceived by Kohlenbach [13], the formal analysis of a proof is typically done 'by hand'. Here, a proof interpretation is simply a means to an end (typically a numerical bound on e.g. a rate of convergence), and as such plays the role of a tool to be wielded by a mathematician. In contrast, for automated program extraction in proof assistants such as MINLOG [1], a proof interpretation forms a high level description of a procedure which has to be implemented, and programs are then extracted synthetically at the push of a button. In this case it goes without saying that the user must provide as input a full machine checkable formal proof, written within the confines of some predetermined logical system.

In this article, we take a somewhat mixed approach. On the one hand, this is not a paper on the formalisation of mathematics: We are interested in a range of rather broad theoretical issues which we intend to present through focusing on the key features of Nash–Williams' proof, and in this sense our construction of a realizing term is based on the pen-and-paper style familiar in proof mining. On the other hand, the novelties of our approach are useful partially because they could be automated within a proof assistant, and so throughout we present our construction in a semi-formal manner in the hope that the reader clearly sees that the main ideas could be implemented at some point.

## 3.1 The Logical System

The main logical theory in which we work will be the theory $\mathrm{PA}^\omega$ of Peano arithmetic in all finite types. Here, we define the finite types to include the base types $\mathbb{B}$ and $\mathbb{N}$ for booleans and natural numbers respectively, and allow the construction of product $X \times Y$, finite sequence $X^*$ and function types $X \to Y$. The theory $\mathrm{PA}^\omega$ is just the usual theory of Peano arithmetic, but with variables and quantifiers for objects of any type. We write $x : X$ or $x^X$ to denote that $x$ has type $X$. Equality symbols $=_\mathbb{B}$ and $=_\mathbb{N}$ for base types are taken as primitive, whereas equality for other types is defined inductively terms of these, so for example $s =_{X \to Y} t :\equiv (\forall x)(sx =_Y tx)$. There are various ways of treating extensionality: for reasons which we will not go into here, the functional interpretation does not interpret the axiom of extensionality—only a weak rule form—and so if the reader prefers they can take $\mathrm{PA}^\omega$ to be the weakly-extensional variant WE-$\mathrm{PA}^\omega$ defined in e.g. [13, 29], although it should be stressed that extensionality is only an issue for the interpreted theory, and when it comes to *verifying* our extracted programs in later sections we freely make use of full extensionality. In any case, the exact details of the logical system are not important in this paper, as our extraction of a program is not fully formal.

What *is* important is the way in which we extend our base theory in order to deal with the minimal bad sequence argument. First note that when reasoning in higher types it is essential to be able to add to our base theory the very weak axiom of choice for quantifier-free formulas:

$$\text{QF-AC} : (\forall x^X)(\exists y^Y)A_0(x, y) \to (\exists f^{X \to Y})(\forall x)A_0(x, f(x)).$$

which as we will see is completely harmless from a computational point of view. In contrast, the crucial additional axiom we choose in order to formalise Nash–Williams' argument is the following syntactic variant of Theorem 2.8 already stated as (2), which we interpret as an axiom schema labelled $\mathrm{ZL}_{\mathrm{lex}}$,

$$\mathrm{ZL}_{\mathrm{lex}} : (\exists u^{\mathbb{N} \to X})B(u) \to (\exists v)(B(v) \wedge (\forall w \lhd_{\mathrm{lex}} v)\neg B(w))$$

where $B(u)$ is understood to range over all *piecewise* formulas of the form $(\forall n)P([u](n))$, and $\lhd$ is some primitive recursive relation on $X$, transfinite induction over which is provable in $\mathrm{PA}^\omega$. Of course, technically we should include this additional wellfounded assumption as a premise so that $\mathrm{ZL}_{\mathrm{lex}}$ becomes a proper axiom schema, but we will omit it for simplicity and if the reader prefers they can just imagine that $\mathrm{ZL}_{\mathrm{lex}}$ is defined relative to some arbitrary but fixed $(X, \rhd)$. Both here and in this chapter, $X$ will actually be of the form $X^*$ and $\rhd$ will be nothing more than the prefix relation on finite sequences which is trivially wellfounded.

Note that the contrapositive of $\mathrm{ZL}_{\mathrm{lex}}$ can be identified with open induction over the lexicographic ordering as treated in [4]:

$$\mathrm{OI}_{\mathrm{lex}} \,:\; (\forall v)((\forall w \lhd_{\mathrm{lex}} v)U(w) \to U(v)) \to (\forall u)U(u)$$

where now $U(u)$ must be an open formula of the form $(\exists n)P([u](n))$ (and so in our terminology, being piecewise is the negation of being open). In the realizability setting of [4], there is a genuine difference between $\mathrm{ZL}_{\mathrm{lex}}$ and $\mathrm{OI}_{\mathrm{lex}}$: the latter is an intuitionistic principle which can be given a direct computational interpretation via open recursion, whereas the former is a non-constructive principle which cannot be realized without the use of e.g. the $A$-translation. On the other hand, for the functional interpretation combined with the negative translation, both of $\mathrm{ZL}_{\mathrm{lex}}$ and $\mathrm{OI}_{\mathrm{lex}}$ are interpreted by exactly the same term (informally, this is due to the fact that the functional interpretation of implication is much more intricate than that of realizability), so they are essentially interchangeable here. We choose $\mathrm{ZL}_{\mathrm{lex}}$ as primitive, because in our opinion the realizing term is a little more intuitive when viewed as an approximation to a minimal element $v$. In any case we will discuss these nuances later. For now, we proceed straight to the formal proof.

## 3.2 *The Formal Proof*

Suppose that $X$ is some arbitrary finite type which comes equipped with some quasi-order $\preceq$, which we take to mean some primitive recursive function $t : X \times X \to \mathbb{B}$ for which reflexivity and transitivity are provable in $\mathrm{PA}^\omega$. We now introduce two predicates which represent the two equivalent definitions of a WQO which we required in Sect. 2:

$$\mathrm{WQO}(\preceq) :\equiv (\forall x^{\mathbb{N}\to X})(\exists i < j)(x_i \preceq x_j)$$
$$\mathrm{WQO}_{\mathrm{seq}}(\preceq) :\equiv (\forall x^{\mathbb{N}\to X})(\exists g^{\mathbb{N}\to\mathbb{N}})(\forall i < j)(g(i) < g(j) \wedge x_{g(i)} \preceq x_{g(j)}).$$

Now, given $\preceq$ we can formally define the embeddability relation $\preceq_*$ as a primitive recursive function in $\preceq$, as in order to check that $a \preceq_* b$ we simply need to carry out a bounded search over all increasing functions $\{0, \ldots, |a| - 1\} \to \{0, \ldots, |b| - 1\}$, where $|a|$ denotes the length of $a$. Note that reflexivity and transitivity of $\preceq_*$ is easily provable from that of $\preceq$ in $\mathrm{PA}^\omega$. The main result of this section is the following:

**Theorem 3.1** *For any fixed quasi-order $\preceq$ on $X$ we have* $\mathrm{PA}^\omega + \text{QF-AC} + \mathrm{ZL}_{lex} \vdash \mathrm{WQO}_{seq}(\preceq) \to \mathrm{WQO}(\preceq_*)$.

The basic strategy of Nash–Williams' proof is first to construct a hypothetical minimal bad sequence, then to deal with a number of simple but quite fiddly cases in order to derive a contradiction. It will be greatly helpful to us in later sections if we separate these two parts here, and prove the following numerically explicit form of the latter:

**Lemma 3.2** *Given a sequence $v : (X^*)^{\mathbb{N}}$, define two sequences $\tilde{v} : (X^*)^{\mathbb{N}}$ and $\bar{v} : X^{\mathbb{N}}$ from $v$ as follows:*

$$\tilde{v}_n, \bar{v}_n := \begin{cases} [], 0_X & \text{if } v_n = [] \\ [x_1, \ldots, x_{k-1}], x_k & \text{if } v_n = [x_1, \ldots, x_k] \end{cases}$$

*where $0_X$ denotes some canonical element of type $X$. Given in addition some function $g : \mathbb{N} \to \mathbb{N}$, define the sequence $w : (X^*)^{\mathbb{N}}$ by*

$$w_n := \begin{cases} v_n & \text{if } n < g(0) \\ \tilde{v}_{g(i)} & \text{if } n = g(0) + i. \end{cases}$$

*Suppose that there exists some $k : \mathbb{N}$ such that*

$$w_{g(0)} \lhd v_{g(0)} \to (\exists i < j < k)(w_i \preceq_* w_j) \tag{3}$$

*where $\lhd$ denotes the strict prefix relation on words, and that $g$ satisfies*

$$(\forall i < j \leq k)(g(i) < g(j) \wedge \bar{v}_{g(i)} \preceq \bar{v}_{g(j)}). \tag{4}$$

*Then there exists a pair of indices $i < j < g(k) + 2$ such that $v_i \preceq_* v_j$.*

*Proof* This is a simple case distinction that we prove in excruciating detail. First of all, we note that by induction on (4) it follows that $i \leq g(i)$ for all $i < k$, which we use below. There are two main cases: A degenerate one where $v_{g(0)} = []$, in which case $v_{g(0)} \preceq_* v_{g(0)+1}$ and so we can set $i, j = g(0), g(0) + 1 < g(0) + 2 \leq g(k) + 2$. For the non-degenerate case where $v_{g(0)} \neq []$ then we have $w_{g(0)} = \tilde{v}_{g(0)} \lhd v_{g(0)}$ and hence $w_i \preceq_* w_j$ for some $i < j < k$ by (3). There are three further possibilities:

(i) $i < j < g(0)$: Then $v_i = w_i \preceq_* w_j = v_j$ and $j < k \leq g(k) < g(k) + 2$.
(ii) $i < g(0) \leq j$: Then $v_i = w_i \preceq_* w_j = \tilde{v}_{g(j')}$ where $j = g(0) + j'$. Either $v_{g(j')} = []$ and so trivially $v_{g(j')} \preceq_* v_{g(j')+1}$, with $j' \leq j$ and hence $g(j') + 1 \leq g(j) + 1 < g(k) + 2$, or $\tilde{v}_{g(j')} \lhd v_{g(j')}$ and therefore $v_i \preceq_* v_{g(j')}$ with $i < g(0) \leq g(j') < g(k)$.

(iii) $g(0) \leq i < j$: Then $\tilde{v}_{g(i')} = w_i \preceq_* w_j = \tilde{v}_{g(j')}$ with $i = g(0) + i'$, $j = g(0) + j'$. If either $\tilde{v}_{g(i')} = []$ or $\tilde{v}_{g(j')} = []$ then the result follows exactly as in part (ii), and otherwise we have $v_{g(i')} = \tilde{v}_{g(i')} * \bar{v}_{g(i')}$ and $v_{g(j')} = \tilde{v}_{g(j')} * \bar{v}_{g(j')}$ and since $i' < j' \leq j < k$ it follows by (4) that $g(i') < g(j')$ and $\bar{v}_{g(i')} \preceq \bar{v}_{g(j')}$ and hence $v_{g(i')} \preceq_* v_{g(j')}$, and since $g(j') \leq g(j) < g(k)$ we're done.

In all cases we have found some $i'' < j'' < g(k) + 2$ with $v_{i''} \preceq_* v_{j''}$. □

*Proof of Theorem* 3.1 First of all, let $\lhd$ denote the strict prefix relation on words, so that $a \lhd b$ iff $|a| < |b|$ and $(\forall i < |a|)(a_i = b_i)$. This is clearly wellfounded, and we can assume for argument's sake that it is decidable, which is automatically the case when $X$ is a base type. Now, define the piecewise predicate $B(u)$ on infinite sequences of words by $B(u) :\equiv (\forall n) P([u](n))$, where

$$P(s) :\equiv (\forall i < j < |s|)(s_i \not\preceq_* s_j).$$

Suppose that $(\exists u) B(u)$. Then by $\mathrm{ZL_{lex}}$ there exists some $v : (X^*)^{\mathbb{N}}$ such that

$$(*) \quad B(v) \wedge (\forall w \lhd_{\mathrm{lex}} v) \neg B(w).$$

Now let $\tilde{v}$, $\bar{v}$ and $w$ be defined as in Lemma 3.2, where $g$ is the function satisfying

$$(\forall i < j)(g(i) < g(j) \wedge \bar{v}_{g(i)} \preceq \bar{v}_{g(j)})$$

which exists by $\mathrm{WQO_{seq}}(\preceq)$. Then (3) holds for some $k$ by minimality of $v$, since if $w_{g(0)} \lhd v_{g(0)}$ then $w \lhd_{\mathrm{lex}} v$, and (4) holds for *any* $k$, therefore by Lemma 3.2 there exists $i < j$ such that $v_i \preceq_* v_j$, contradicting $B(v)$. Therefore $(\exists u) B(u)$ is false, which implies that for all $u$ there exists some $i < j$ such that $u_i \preceq_* u_j$, which is $\mathrm{WQO}(\preceq_*)$. Hence we have shown that $\mathrm{WQO_{seq}}(\preceq) \to \mathrm{WQO}(\preceq_*)$. □

While the derivation above is not fully formal in the sense that would be expected were we to formally extract a program using a proof assistant, in contrast to the textbook proof given in Sect. 1 it makes explicit important quantitative information which will guide us in constructing a realizing term, as we will see later.

Now to our main problem, which is to prove that $(\{0, 1\}, =_*)$ is a WQO. From now on we will equate the two letter alphabet $\{0, 1\}$ with our type $\mathbb{B}$. Suppose that $\preceq$ is now just $=_{\mathbb{B}}$, which is clearly a WQO. In order to be able to apply Higman's lemma, we need to establish $\mathrm{WQO_{seq}}(=_{\mathbb{B}})$, either by formalising Lemma 2.2 or by a direct argument. We choose the latter, since for the simple case of $=_{\mathbb{B}}$ a single instance of the law of excluded-middle suffices, whereas the general case given as Lemma 2.2 would require some form of comprehension.

**Theorem 3.3** $\mathrm{PA}^{\omega} + \text{QF-AC} \vdash \mathrm{WQO}_{seq}(=_{\mathbb{B}})$.

*Proof* We will first show that

$$(\forall x^{\mathbb{N} \to X})(\exists b^{\mathbb{B}})(\forall n)(\exists k \geq n)(x_k = b). \tag{5}$$

Fix some sequence $x : \mathbb{B}^{\mathbb{N}}$. By the law of excluded middle we have

$$(\exists n)(\forall k \geq n)(x_k = 0) \vee (\forall n)(\exists k \geq n)(x_k = 1).$$

If the left hand side of the disjunction holds we set $b := 0$. We have that there is some $N$ such that $x_k = 0$ for all $k \geq N$, and so for an arbitrary number $n$, setting $k := \max\{N, n\}$ yields $k \geq n$ and $x_k = 0$. If the right hand side holds, we set $b := 1$ and we are done by definition.

So we have proved (5). To establish $\mathrm{WQO}_{\mathrm{seq}}(=_{\mathbb{B}})$, take some $x$ and let $b$ be such that $(\forall n)(\exists k \geq n)(x_k = b)$. By QF-AC there exists some $f : \mathbb{N} \to \mathbb{N}$ satisfying

$$(\forall n)(f(n) \geq n \wedge x_{f(n)} = b).$$

Now, define $g : \mathbb{N} \to \mathbb{N}$ via primitive recursion as

$$g(0) := f(0) \text{ and } g(n+1) := f(g(n)+1).$$

Then it is clear that $g(n) < g(n+1)$ and $x_{g(n)} = b$, and therefore

$$(\forall i < j)(g(i) < g(j) \wedge x_{g(i)} = b = x_{g(j)})$$

and we're done. □

Now, putting together Theorems 3.1 and 3.3, we have:

**Corollary 3.4** $\mathrm{PA}^{\omega} + \text{QF-AC} + \mathrm{ZL}_{lex} \vdash \mathrm{WQO}(=_{\mathbb{B},*})$.

We summarise the main structure of our proof of $\mathrm{WQO}(=_{\mathbb{B},*})$ in Figure 1. There are three main parts to the proof, each of which will be treated somewhat separately in what follows, namely:

(1) A proof of $\mathrm{WQO}_{\mathrm{seq}}(=_{\mathbb{B}})$ given as Theorem 3.3, which uses an instance of the law of excluded middle for $\Pi^0_2$ formulas.
(2) A single instance of $\mathrm{ZL}_{\mathrm{lex}}$ applied to the piecewise formula $(\forall i < j)(u_i \not\preceq_* u_j)$, set out in the main proof of Theorem 3.1.
(3) The derivation of a contradiction from $\mathrm{WQO}_{\mathrm{seq}}(\preceq)$ combined with the existence of a minimal bad sequence, which is Lemma 3.2.

Having now introduced the theory of WQOs and given a formal proof of our main result, the remainder of the paper will be dedicated to constructing a program which finds an embedded pair of words in an arbitrary input sequence. We will introduce our main tool - Gödel's functional interpretation—in the next section, then each of the three main components will be analysed in turn in Sects. 5, 6–8 and 9, respectively.

**Fig. 1** A map of the formal proof

$$\dfrac{\dfrac{\text{Theorem 3.3}}{\vdots \atop \mathrm{WQO}_{\mathrm{seq}}(=_{\mathbb{B}})} \qquad \dfrac{(\exists u)B(u)}{(\exists v)(B(v) \wedge (\forall w \lhd_{\mathrm{lex}} v)\neg B(w))}\,\mathrm{ZL}_{\mathrm{lex}}}{\dfrac{\bot}{\mathrm{WQO}(=_{\mathbb{B},*})}}\,\text{Lemma 3.2}$$

# 4 Gödel's Functional Interpretation

We now put well quasi-orders aside for a moment, and introduce the second main topic of this paper: Gödel's functional (or 'Dialectica') interpretation. The reader already familiar with this may wish to skip ahead to Sect. 6, which continues with the computational interpretation of $\mathrm{ZL}_{\mathrm{lex}}$.

Presenting the functional interpretation is something of a challenge for an author: The interpretation is one of those syntactical objects—particularly common in proof theory—whose basic definition can be given in a few lines and whose characterising theorem (in this case soundness) can be set up in a couple of pages, and yet none of this is necessarily remotely helpful in giving the unacquainted reader any real insight into what it actually does! In reality, the functional interpretation is an extraordinarily subtle idea which continues to be studied from a range of perspectives, and the fact that it forms one of the central techniques of the highly successful proof mining program is testament to its power. For a comprehensive treatment of the functional interpretation and its role in program extraction, the reader is encouraged to consult the standard textbook [13], or alternatively the shorter chapter [3].

Nevertheless, in an effort to make this essay as accessible as possible it is important that I say something about the interpretation here. So my plan is as follows: in Sects. 4.1–4.3 I will begin by defining the interpretation, and will state without proof the main results on program extraction. This will all be standard material. Then in Sect. 4.4 I will employ the slightly unconventional tactic of explaining on a high level how the functional interpretation treats a series of formulas of a specific logical shape, which appear several times in the remainder of this work. Finally, in Sect. 5, I will present in quite some detail the extraction of a simple program from the proof of Theorem 3.3, which will conveniently serve simultaneously as a illustration of the functional interpretation in action and the first step in our main challenge!

## *4.1 The Basics*

In one sentence, Gödel's functional interpretation is a syntactic translation which takes as input a formula $A$ in some logical theory $\mathcal{L}$ and returns a new formula $A' := (\exists x)(\forall y)|A|^x_y$ where $x$ and $y$ are sequences of potentially higher type variables, and $|A|^x_y$ is in some sense computationally neutral, which in this article will just mean quantifier-free and hence decidable (since characteristic functions for all quantifier-free formulas can be constructed in $\mathrm{PA}^\omega$). The idea behind the translation is that $A \leftrightarrow A'$ over some reasonable higher-type theory, but the latter can be witnessed by some term in a calculus $T$. We say that the functional interpretation soundly interprets $\mathcal{L}$ in $T$, if for any formula in the language of $\mathcal{L}$ we have

$$\mathcal{L} \vdash A \Rightarrow \text{ there exists some closed term } t \text{ of } T \text{ such that } \mathcal{T} \vdash |A|^t_y$$

where $\mathcal{T}$ represents some verifying theory which allows us to reason about terms in our calculus $T$. Crucially, the soundness proof comes equipped with a method of constructing such a realizer $t$ from the proof of $A$. The direct approach above typically works for *intuitionistic* theories $\mathcal{L}$ extended with some weak semi-classical axioms (for example Markov's principle), but for theories $\mathcal{L}_c$ based on full classical logic, we need to precompose the functional interpretation with a negative translation $A \mapsto A^{\neg\neg}$. Therefore from now on, soundness of the functional interpretation for classical theories refers to the following:

$$\mathcal{L}_c \vdash A \Rightarrow \text{ there exists some closed term } t \text{ of } T \text{ such that } \mathcal{T} \vdash |A^{\neg\neg}|^t_y.$$

In this article, our $\mathcal{L}_c$ will be $\mathrm{PA}^\omega + \text{QF-AC}$, later extended with $\mathrm{ZL}_{\mathrm{lex}}$. But before we go further, we need to introduce our functional calculus $T$.

## *4.2 The Programming Language*

Our interpreting calculus will be a standard variant of Gödel's system T, extended with product and finite sequence types as with our variant of $\mathrm{PA}^\omega$. System $T$ is well-known enough that we feel no need to give a proper definition here: in any case full details can be found in many places, including the aforementioned sources [3, 13, 29]. In a sentence: System $T$ is a simply typed lambda calculus which allows the definition of functionals via primitive recursion in all higher types. We summarise the basic constructions of the calculus below, if only to allow the reader to become familiar with our notational conventions. We take the types of $T$ to be the same as those in our logical system $\mathrm{PA}^\omega$, namely those build from $\mathbb{B}$ and $\mathbb{N}$ via product, sequence and arrow types. Terms of the calculus include the following:

- **Functions**. We allow the construction of terms via lambda abstraction and application: if $x : X$ and $t : Y$ then $\lambda x.t : X \to Y$, while if $t : X \to Y$ and $s : X$ then $ts : Y$, and these satisfy the usual axioms, e.g. $(\lambda x.t[x])(s) = t[s\backslash x]$.
- **Canonical objects**. For each type $X$ we define a canonical 'zero object' $0_X : X$ in the standard manner, with $0_{\mathbb{N}} = 0$, $0_{\mathbb{B}} = 0$, $0_{\mathbf{1}} = ()$, $0_{X\times Y} = \langle 0_X, 0_Y\rangle$, $0_{X^*} = []$ and $0_{X\to Y} = \lambda x.0_Y$.
- **Products**. Given $z : X \times Y$ we often write just $z_0, z_1$ for the projections $\pi_0 z : X$, $\pi_1 z \in Y$. This will also be the case for sequences, where for $z : (X \times Y)^{\mathbb{N}}$, $z_0 : X^{\mathbb{N}}$ is defined by $(z_0)_n := \pi_0 z_n$ and so on. For $x : X$ and $y : Y$ we have a pairing operator $\langle x, y\rangle : X \times Y$.

- **Sequences**. As before, given $s : X^*$ we denote by $|s|$ the length of $s$, for $x : X$ we define $s * x : X^*$ by $[s_0, \ldots, s_{k-1}, x]$ i.e. the concatenation of $s$ with $x$, and use this also for the concatenation of $s$ with another finite sequence $s * t : X^{\mathbb{N}}$ or an infinite sequence $s * \alpha : X^{\mathbb{N}}$. For $\alpha : X^{\mathbb{N}}$ we let $[\alpha](n) := [\alpha_0, \ldots, \alpha_{n-1}]$.
- **Recursors**. For each type we have a recursor $\mathrm{Rec}_X$ which has the defining equations

$$\mathrm{Rec}_X^{a,h}(0) =_X a \text{ and } \mathrm{Rec}_X^{a,h}(n+1) =_X hn(\mathrm{Rec}^{a,h}(n)).$$

for parameters $a : X$ and $h : \mathbb{N} \to X \to X$.

Note that having access to recursors of arbitrary finite type means that along with all normal primitive recursive functions we can define e.g. the Ackermann function (using $\mathrm{Rec}_{\mathbb{N}\to\mathbb{N}}$). Indeed, the closed terms of type $\mathbb{N} \to \mathbb{N}$ definable in $T$ are the provably recursive functions of Peano arithmetic, a fact which follows from the soundness of the functional interpretation.

There are a couple of further remarks to be made. First, we have presented system $T$ as a equational calculus, but of course we could have instead used a conversion rule $\to_T$, in which case system $T$ can be viewed as a fragment of PCF consisting only of total objects. More concretely, any term of system $T$ can be straightforwardly written as a functional program, and we encourage the reader to think of system $T$ in this manner, as a high level means of describing real programs.

Finally, in the previous section we wrote $\mathcal{T} \vdash |A|^t_y$, which implies that $T$ also comes equipped with a logic $\mathcal{T}$ for verifying programs. There are various ways of defining the underlying logic of system $T$—traditionally it is presented as a minimal quantifier-free calculus with an induction axiom, although alternatively we can just identify $\mathcal{T}$ with the fully extensional variant of $\mathrm{PA}^\omega$, extended with additional axioms whenever we need them. Such distinctions are more relevant for foundational issues such as relative consistency proofs, where it was the goal to make $\mathcal{T}$ as weak as possible. Here we have no such concerns, and so we reason about the correctness of our extracted programs in a fairly free manner.

### *4.3 The Interpretation*

The functional interpretation $|A|^x_y$ of a formula $A$ in the language of $\mathrm{PA}^\omega$ is defined inductively over the logical structure of $A$ as follows:

(i) $|A| :\equiv A$ if $A$ is prime
(ii) $|A \wedge B|^{x,u}_{y,v} :\equiv |A|^x_y \wedge |B|^u_v$
(iii) $|A \vee B|^{b^{\mathbb{B}},x,u}_{y,v} :\equiv |A|^x_y \vee_b |B|^u_v$
(iv) $|A \to B|^{f,g}_{x,v} :\equiv |A|^x_{gxv} \to |B|^{fx}_v$
(v) $|\exists t^X A(t)|^{z,x}_y :\equiv |A(z)|^x_y$
(vi) $|\forall t^X A(t)|^f_{z,y} :\equiv |A(z)|^{fz}_y$

where in clause (iii) we define

$$P \vee_b Q :\equiv (b = 0 \to P) \wedge (b = 1 \to Q).$$

At first glance, the functional interpretation looks very much like a standard BHK interpretation, with the exception of the treatment of implication (iv), which is in many ways the characterising feature of the interpretation. Though (iv) may appear to be a little mysterious, it should be viewed as the 'least non-constructive' Skolemization of the formula

$$(\exists x)(\forall y)|A|^x_y \to (\exists u)(\forall v)|B|^u_v$$

which goes via

$$(\forall x)(\exists u)(\forall v)(\exists y)(|A|^x_y \to |B|^u_v)$$

as an intermediate step. In the original 1958 paper [10], Gödel proved that the usual first order theory of Heyting arithmetic could be soundly interpreted in System $T$. It follows directly that Peano arithmetic can also be interpreted in $T$ when precomposed with the negative translation, and in fact it is not difficult at all to extend these results to the higher-order extensions of arithmetic with quantifier-free choice:

**Theorem 4.1** *Let $A(a)$ be a formula in the language of (weakly-extensional)* $\mathrm{PA}^\omega$ *containing only a free. Then*

$$\mathrm{PA}^\omega + \text{QF-AC} \vdash A(a) \Rightarrow \mathcal{T} \vdash |A(a)^{\neg\neg}|^{ta}_y$$

*where $t$ is a closed term of $T$ which can be formally extracted from the proof of $A(a)$.*

A modern presentation and proof of this result can be found in [13], which also discusses the various theories $\mathcal{T}$ in which soundness can be formalised. As simple as Theorem 4.1 appears, understanding how the combination of negative translation and functional interpretation treats even simple logical formulas is far from straightforward, and is often a stumbling block when one first encounters the ideas of applied proof theory. We now try to provide some insight into this.

## *4.4 The* Meaning *of the Interpretation*

In order that the reader not familiar with the functional interpretation and program extraction can follow the main part of the paper, it is important that we highlight how the functional interpretation combined with the negative translation treats a handful of key formulas. Note that we have not yet stated which negative translation we use. Unless one wants to formalise the translation this is not so important: Typically what we do is take some arbitrary choice of $A^{\neg\neg}$ and use a range of semi-intuitionistic laws which are admitted by the functional interpretation to rearrange it into a simpler formula which is easier to interpret.

### 4.4.1 $\Pi_2$ Formulas $A :\equiv (\forall x)(\exists y)B(x, y)$

We begin with one of the key properties of the functional interpretation, which makes it so useful for program extraction. The negative translation of a $\Pi_2$ formula is equivalent to $(\forall x)\neg\neg(\exists y)B(x, y)$. However, it is not difficult to see that the functional interpretation translates $\neg\neg(\exists y)B(x, y)$ to $(\exists y)B(x, y)$, and so in particular the functional interpretation admits Markov's principle. Therefore

$$(\forall x)\neg\neg(\exists y)B(x, y) \text{ is translated to } (\exists f)(\forall x)B(x, fx)$$

and therefore in theory we can extract a program directly witnessing a $\Pi_2$ formula, even when that formula is proven classically. Note that the statement WQO($\preceq$) is a $\Pi_2$ formula, and so even though we use a number of non-constructive principles in its proof, we can still hope to extract a program $\Phi$ witnessing it!

### 4.4.2 $\Sigma_2$ Formulas $A :\equiv (\exists x)(\forall y)B(x, y)$

In contrast to $\Pi_2$ formulas, $\Sigma_2$ formulas are more problematic, as there provable $\Sigma_2$ formulas which cannot in general be directly witnessed by a computable function. Here, the negative translation is equivalent to $\neg\neg(\exists x)(\forall y)B(x, y)$, and functional interpretation acts as follows

$$\begin{aligned} \neg\neg(\exists x)(\forall y)B(x, y) &\mapsto \neg(\exists f)(\forall x)\neg B(x, fx) \\ &\mapsto (\forall f)(\exists x)\neg\neg B(x, fx) \\ &\mapsto (\forall f)(\exists x)B(x, fx) \quad (*) \\ &\mapsto (\exists\Phi)(\forall f)B(\Phi f, f(\Phi f)). \end{aligned}$$

Note that we can omit double negations in front of $B(x, y)$ as this is a quantifier-free formula. Nevertheless, the interpretation of our original formula gives us something 'indirect', which in this case coincides with Kreisel's well-known 'no-counterexample' interpretation (although in general the functional interpretation is different). The intuitive idea is that $f$ is a function which attempts to witness falsity of $A$ i.e. $(\forall x)(\exists y)\neg B(x, fx)$. Then the functional $\Phi$ takes any proposed 'counterexample function' and shows that it must fail. Over classical logic, the existence of such a functional $\Phi$ is equivalent to the existence of some $x$ satisfying $(\forall y)B(x, y)$, but unlike $x$ it can be directly constructed.

Another way of understanding the meaning of $\Phi$ is as a program which constructs an *approximation* to the non-constructive object $x$. In general, we cannot compute an $x$ which satisfies $B(x, y)$ for all $y$, but given some $f$ we can find an $x$ which satisfies $B(x, fx)$. In this setting, $f$ should be seen as a function which calibrates *how good* the approximation should be. We make extensive use of this intuition later, where we explain how the functional interpretation of the minimal bad sequence construction can be viewed as the statement that arbitrarily good 'approximate' minimal bad sequences exist.

We also note that throughout this paper, we will often express the interpretation of $\Pi_2$ formulas in its penultimate form $(\forall f)(\exists x)B(x, fx)$ indicated by $(*)$ above. This is for no other reason than that it is much easier to talk about $x$ instead of $\Phi f$, and so we avoid a lot of rather messy notation!

### 4.4.3 Implication $A :\equiv (\exists x)(\forall y)B(x, y) \rightarrow (\exists u)(\forall v)C(u, v)$

Finally, it is important to sketch what happens when we want to infer the existence of a non-constructive object $u$ from the existence of another non-constructive object $x$. In this case, the negative translation of $A$ is intuitionistically equivalent to

$$(\exists x)(\forall y)B(x, y) \rightarrow \neg\neg(\exists u)(\forall v)C(u, v),$$

Now, by the previous section, the functional interpretation of the conclusion yields

$$(\exists x)(\forall y)B(x, y) \rightarrow (\forall f)(\exists u)C(u, fu).$$

and so interpreting the implication as a whole following clause (iv) we get

$$(\exists F, G)(\forall x, f)(B(x, Gfx) \rightarrow C(Fxf, f(Fxf))).$$

In terms of our discussion above, $F$ and $G$ can be read as follows: For any given $x$, $Fx$ is a functional which computes an approximation to the conclusion of the implication i.e. $(\forall f)(\exists u)C(u, fu)$, where now it uses that $(\forall y)B(x, y)$ holds. The functional $G$ tells us, in a certain sense, how good the approximation of the *premise* has to be in order to build an approximation of the conclusion: this is given by $Gf$. Note that whenever we have a functional $\Phi$ which builds an approximation to the premise in this way i.e. $B(\Phi(Gf), Gf(\Phi(Gf)))$ we can use it to construct an approximation to the conclusion.

While all this may sound extremely intricate, it will hopefully become clearer when we see some concrete examples in the sequel.

## 5 Interpreting the Proof of $\mathrm{WQO}_{\mathrm{seq}}(=_{\mathbb{B}})$

We now give our first illustration of the functional interpretation in action. In Theorem 3.3 we showed that the proof of $\mathrm{WQO}_{\mathrm{seq}}(=_{\mathbb{B}})$ could be formalised in $\mathrm{PA}^\omega + \mathrm{QF\text{-}AC}$, and hence by Theorem 4.1 we know for sure that we can construct a program in $T$ which witnesses its functional interpretation. However, actually doing so, and ending up with a program whose behaviour can be comprehended is another matter, and in what follows we outline the philosophy emphasised later in the paper of combining formal program extraction with intuition. Note that nothing in this section is new, and if the reader prefers they can simply glance at the program we obtain in Sect. 5.3 and move straight on to Sect. 6.

The proof of Theorem 3.3 has three main components. The first is an obviously non-constructive axiom, namely the law of excluded middle for $\Sigma_2$ formulas applied to $(\exists n)(\forall k \geq n)(x_k = 0)$. The second is the derivation of the auxiliary statement (5) from this instance of the law of excluded-middle, and the final is the derivation of $\mathrm{WQO}_{\mathrm{seq}}(=_{\mathbb{B}})$ from (5) by constructing the necessary primitive recursive function. We will treat each of these in turn. Before we do so, it is worth spelling out explicitly what our goal is. First note that $\mathrm{WQO}_{\mathrm{seq}}(=_{\mathbb{B}})$ can be equivalently formulated as the $\Pi_3$ formula

$$(\forall x)(\exists g)(\forall n)(\forall i < j \leq n)(g(i) < g(j) \wedge x_{g(i)} = x_{g(j)}).$$

Note that we have merged the two quantifiers $\forall i, j$ into a single 'real' quantifier $\forall n$ so that they are now bounded and hence decidable. In particular, this means they are now ignored by the functional interpretation. The negative translation of this formulas is equivalent to

$$(\forall x)\neg\neg(\exists g)(\forall n)(\forall i < j \leq n)(g(i) < g(j) \wedge x_{g(i)} = x_{g(j)}).$$

Therefore, referring back to Sect. 4.4.2 our challenge is to produce a program $\Phi$ which takes as input $x$ together with some 'counterexample functional' $\omega : (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ and witnesses $(\exists g)$ in the formula

$$(\forall x, \omega)(\exists g)(\forall i < j \leq \omega g)(g(i) < g(j) \wedge x_{g(i)} = x_{g(j)}). \tag{6}$$

In other words, while we cannot hope to effectively construct a monotone subsequence $g$ in general, we can always do the next best thing and construct an approximation to it which works for all $i < j \leq \omega g$. Then when it comes to using $\mathrm{WQO}_{\mathrm{seq}}(=_{\mathbb{B}})$ as a lemma in the proof of $\mathrm{WQO}(=_{\mathbb{B},*})$ we will need to calibrate exactly how big this approximation needs to be, in other words construct some concrete $\omega$ such that $\mathrm{WQO}(=_{\mathbb{B},*})$ follows from (6).

### 5.1 The Law of Excluded-Middle for $\Sigma_2^0$ Formulas

Our first step when interpreting a classical proof is to interpret the main non-constructive axioms which are needed. When interpreting $\mathrm{WQO}_{\mathrm{seq}}(\preceq) \to \mathrm{WQO}(\prec)$ later, our focus will be on $\mathrm{ZL}_{\mathrm{lex}}$. Here we must deal with the somewhat simpler law of excluded-middle for $\Sigma_2^0$ formulas:

$$(\exists n)(\forall k)P(n, k) \vee (\forall m)(\exists l)\neg P(m, l).$$

The functional interpretation Skolemizes this as the $\Sigma_1$ formula

$$(\exists b, n, h)(\forall k, m)[P(n,k) \vee_b \neg P(m, hm)]$$

and following Sect. 4.4.2 the interpretation of the double negation of this formula is given by

$$(\forall \phi, \psi)(\exists b, n, h)[P(n, \phi bnh) \vee_b \neg P(\psi bnh, h(\psi bnh))]. \tag{7}$$

This already looks rather complex thanks to all the function dependencies, but the way to think of $\phi$ and $\psi$ is again as counterexample functionals which represent the quantifiers $(\forall k)$ and $(\forall m)$ respectively. Now, we have two options in front of us: We can either carefully analyse the formal derivation of the negative translation of the law of excluded-middle in intuitionistic logic in order to produce a realizer for (7), or we can try to do this directly using our intuition. We choose the latter—and it is this kind of thing that characterizes our 'semi-formal' approach to program extraction.

Let's look more closely at (7). Our boolean $b$ is just a marker which tells us which side of the conjunction holds, so essentially what we must do is find a pair $n_L, n_R$ and $h_L, h_R$ which satisfy either $P(n_L, \phi 0 n_L h_L)$ or $\neg P(\psi 1 n_R h_R, h_R(\psi 1 n_R h_R))$. In the first case we can then define $b, n, h$ to be $0, n_L, h_L$, and in the second to be $1, n_R, h_R$. In order to do this, we want to define these so that

$$P(n_L, \phi 0 n_L h_L) \leftrightarrow P(\psi 1 n_R h_R, h_R(\psi 1 n_R h_R))$$

so that $P(n_L, \phi 0 n_L h_L) \vee \neg P(\psi 1 n_R h_R, h_R(\psi 1 n_R h_R))$ follows directly from the law of excluded-middle for quantifier-free formulas. Note that we can force this equivalence to hold if

$$n_L = \psi 1 n_R h_R \quad \text{and} \quad \phi 0 n_L h_L = h_R(\psi 1 n_R h_R).$$

So can we solve these equations? Well, the first thing we notice is that $n_R$ and $h_L$ do not depend on anything and so can be freely chosen, so we just set these to be canonical elements $n_R, h_L := 0_{\mathbb{N}}, 0_{\mathbb{N}\to\mathbb{N}}$ (note that this makes sense intuitively, since $n$ only plays a role in the left disjunct, and $h$ only in the right). We can then define $n_L := \psi 1 0 h_R$. It remains to find some $h_R$ which satisfies

$$h_R(\psi 1 0 h_R) = \phi 0 n_L 0 = \phi 0 (\psi 1 0 h_R) 0,$$

where the latter equality follows by substituting in our definition for $n_L$. But this is easily achieved if we set $h_R := \lambda i.\phi 0 i 0$. So we're done, and to summarise, (7) is solved by setting

$$b, n, h := \begin{cases} 0, \psi 1 0 h_R, 0 & \text{if } P(\psi 1 0 h_R, \phi 0 (\psi 1 0 h_R) 0) \\ 1, 0, h_R & \text{otherwise} \end{cases}$$

The reader can now easily check that this is indeed a solution by substituting it back into (7). Note that while we use the law of excluded-middle in a very specific way in our proof, the above would work for *any* instance of $\Sigma^0_2$ law of excluded middle (in fact we don't even need the quantifiers to be of lowest type).

## 5.2 *Interpreting* $(\forall x)(\exists b)(\forall n)(\exists k \geq n)(x_k = b)$

We now use the realizing term given above to witness the functional interpretation of our intermediate result $(\forall x)(\exists b)(\forall n)(\exists k \geq n)(x_k = b)$. In order to distinguish this $b$ from that of the previous section, we relabel it as $c$. Taking into account the negative translation, what we mean is to interpret is

$$(\forall x)\neg\neg(\exists c, f)(\forall n)(fn \geq n \wedge x_{fn} = c)$$

and hence

$$(\forall x, \xi)(\exists c, f)(f(\xi cf) \geq \xi cf \wedge x_{f(\xi cf)} = c).$$

Again, this looks somewhat intricate, but the term $\xi cf$ simply represents the quantifier $(\forall n)$. Now, in order to prove this statement we used the law of excluded-middle for the formula $P(n, k) :\equiv (k \geq n \rightarrow x_k = 0)$ given some fixed sequence $x$. What we need to do is work out exactly how this was used, and following our discussion in Sect. 4.4.3 this means realizing the implication

$$(\exists b, n, h)(\forall k, m)(P(n, k) \vee_b P(m, hm)) \rightarrow (\forall \xi)(\exists c, f)(f(\xi cf) \geq \xi cf \wedge x_{f(\xi cf)} = c) \tag{8}$$

and therefore

$$(\forall b, n, h, \xi)(\exists k, m, c, f)[P(n, k) \vee_b P(m, hm) \rightarrow f(\xi cf) \geq \xi cf \wedge x_{f(\xi cf)} = c].$$

To this end, let us fix $b, n, h, \xi$. There are two possibilities. If $b = 0$ then we must find some $k, c, f$ (we can set $m = 0$) such that the conclusion follows from $P(n, k)$. It's sensible to choose $c := 0$, then it remains to find $k, f$ satisfying

$$(k \geq n \rightarrow x_k = 0) \rightarrow f(\xi 0 f) \geq \xi 0 f \wedge x_{f(\xi 0 f)} = 0.$$

Following our formal proof, let's define $f(i) := \max\{n, i\}$ and

$$k := f(\xi 0 f) = \max\{n, \xi 0 f\} = \max\{n, \xi 0(\lambda i. \max\{n, i\})\}.$$

Then clearly $f(\xi 0 f) \geq n, \xi 0 f$ and $x_{f(\xi 0 f)} = 0$ follows from the premise.

In the second case $b = 1$, setting $c := 1$ (and this time $k = 0$) we must establish the conclusion from $\neg P(m, hm)$, i.e. find $m, f$ satisfying

$$hm \geq m \wedge x_{hm} = 1 \rightarrow f(\xi 1 f) \geq \xi 1 f \wedge x_{f(\xi 1 f)} = 1.$$

But this is more straightforward: $f := h$ and $m := \xi 1 h$ work, so we're done. In other words, defining

$$\phi_\xi 0 n h := \max\{n, \xi 0(\lambda i. \max\{n, i\})\} \quad \text{and} \quad \psi_\xi 1 n h := \xi 1 h$$

with $\phi_\xi 1 n h = \psi_\xi 0 n h = 0$, we can eliminate the quantifiers $(\forall m, k)$ in (8), and we have proven that

$$(\forall b, n, h, \xi)[P(n, \phi_\xi b n h) \vee_b P(\psi_\xi b n h, h(\psi_\xi b n h)) \rightarrow f(\xi c f) \geq \xi c f \wedge x_{f(\xi c f)} = c]$$

for

$$c, f := \begin{cases} 0, \lambda i. \max\{n, i\} & \text{if } b = 0 \\ 1, h & \text{otherwise.} \end{cases}$$

But we know how to find $b, n, h$ which solve the premise for $\phi_\xi$ and $\psi_\xi$, so substituting those solutions in the definition above we have

$$(\forall \xi)(f(\xi c f) \geq \xi c f \wedge x_{f(\xi c f)} = c) \tag{9}$$

for

$$c, f := \begin{cases} 0, \lambda i. \max\{\psi_\xi 1 0 h_R, i\} & \text{if } P(\psi_\xi 1 0 h_R, \phi_\xi 0(\psi_\xi 1 0 h_R) 0) \\ 1, h_R & \text{otherwise} \end{cases} \tag{10}$$

where $\phi_\xi$, $\phi_\xi$, $P$ and $h_R$ are defined as above.

## 5.3 Simplifying the Realizing Term

The solution given above for finding $c$ and $f$ in $\xi$ is perfectly valid, but still somewhat tricky to understand, as it is couched in terms of the abstruse functionals which arise from our formal proof. So while an automated extraction may produce something like this, for a human being it is desirable to simplify everything and see if there is an underlying pattern.

It is immediately clear by inspecting the definition (10) above, that there are three key terms which play a role, namely $h_R$, $\psi_\xi 1 0 h$ and $\phi_\xi 0 i 0$, with substitutions $h \mapsto h_R$ and $i \mapsto \psi_\xi 1 0 h_R$. So it makes sense to unwind each of these terms. First, notice that from the definitions of $\phi_\xi$, $\psi_\xi$ we have

$$h_R(i) = \phi_\xi 0 i 0 = \max\{i, \xi 0(\lambda j. \max\{i, j\})\} \quad \text{and} \quad \psi_\xi 1 0 h = \xi 1 h$$

and so in particular

$$\psi_\xi 10 h_R = \xi 1(\lambda i.\max\{i, \xi 0(\lambda j.\max\{i, j\})\}) =: a$$
$$\phi_\xi 0(\psi_\xi 10 h_R)0 = h_R(a)$$

and our solution can already be simplified to

$$c, f := \begin{cases} 0, \lambda i.\max\{a, i\} & \text{if } x_{\max\{a, \xi 0(\lambda j.\max\{a, j\})\}} = 0 \\ 1, \lambda i.\xi 0(\lambda j.\max\{i, j\}) & \text{otherwise.} \end{cases}$$

where we use the fact that $P(a, \max\{a, \xi 0(\lambda j.\max\{a, j\})\}) \leftrightarrow x_{\max\{a, \xi 0(\lambda j.\max\{a, j\})\}} = 0$.

We now see that, far from being the syntactic mess it appeared, our realizing term can be expressed in a very natural way. By looking closer, an interesting structure emerges: Given a function $q : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, and a pair of functions $\varepsilon, \delta : (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$, define the pair $(\varepsilon \otimes \delta)(q) := \langle a, b[a] \rangle$ where

$$b[i] := \delta(\lambda j.q(i, j))$$
$$a := \varepsilon(\lambda i.q(i, b[i])).$$

This is the so-called *binary product of selection functions* studied by Escardó and Oliva [7]. Intuitively it gives a optimal play in a two player sequential game, where $q$ is assigns an outcome to each pair of moves, and $\varepsilon, \delta$ dictate the strategy of the first and second players respectively. Using this new notation, our realizer becomes

$$c, f := \begin{cases} 0, \lambda i.\max\{a, i\} & \text{if } x_{\max\{a, b[a]\}} = 0 \\ 1, \lambda i.\xi 0(\lambda j.\max\{i, j\}) & \text{otherwise.} \end{cases}$$

where $\langle a, b[a] \rangle = (\xi 1 \otimes \xi 0)(\max)$. An extension of this idea for nested sequences of the law of excluded-middle were first considered in [20], and generalisations of the product of selection functions to so-called 'unbounded' games have been used to give computational interpretations to choice principles, thereby opening up a fascinating bridge between functional interpretations and game theory [7–9].

## 5.4 The Final Step

It now remains to compute our realizer $g$ in $x$ and $\omega$ satisfying (6). For this we need to define a suitable functional $\xi$ such that (6) follows from (9). The key here is the following trick. Suppose that $\tilde{\xi}$ is another functional, and define

$$\xi c f := (\mu i \leq \tilde{\xi} c f)\neg(f(i) \geq i \wedge x_{f(i)} = c).$$

where $(\mu i \leq n)P(i)$ denotes the least $i \leq n$ satisfying $P(i)$, and just 0 if none exist. Then

$$f(\xi cf) \geq \xi cf \wedge x_{f(\xi cf)} = c \to (\forall i \leq \tilde{\xi} cf)(f(i) \geq i \wedge x_{f(i)} = c).$$

So, to this end, define $g_f(0) := f(0)$ and $g_f(n+1) := f(g_f(n)+1)$, and let

$$\xi_{x,\omega} cf := (\mu i \leq \tilde{\xi}_\omega cf)\neg(f(i) \geq i \wedge x_{f(i)} = c).$$

for

$$\tilde{\xi}_\omega cf := \max_{k \leq \omega g_f} \{g_f(k) + 1\}.$$

Then we claim that

$$f(\xi_{x,\omega} cf) \geq \xi_{x,\omega} cf \wedge x_{f(\xi_{x,\omega} cf)} = c \to (\forall i < j \leq \omega g_f)(g_f(i) < g_f(j) \wedge x_{g_f(i)} = x_{g_f(j)}).$$

Take $i < \omega g_f$. Then $g_f(i) + 1 \leq \xi_\omega cf$ and therefore $g_f(i+1) = f(g_f(i)+1) \geq g_f(i) + 1 > g_f(i)$ and $x_{g_f(i+1)} = c$, from which we can infer that for any $i < j$ we have $g_f(i) < g_f(j)$ and $x_{g_f(i)} = x_{g_f(j)} = c$. Therefore, by the previous sections, we have

$$(\forall x, \omega)(\forall i < j \leq \omega g)(g(i) < g(j) \wedge x_{g(i)} = x_{g(j)})$$

where $g = g_f$ for the $f$ satisfying (9) relative to $\xi_{x,\omega}$ defined above.

## 5.5 Summary

Our aim in this section was to lead the reader through an actual example of program extraction, from a much simpler classical principle than that about to be considered below. Rather than just presenting an extracted term, our hope was to illustrate how by analysing extracted programs and applying a degree of ingenuity, one can devise descriptions of these programs which are of interest in their own right. Here, the observation by Escardó and Oliva that the functional interpretation of the law of excluded-middle concealed a natural game-theoretic construction which could be extended to encompass much stronger principles led to a large body of research in a somewhat unexpected direction. Later in this paper, a notion of a *learning procedure* will play somewhat analogous role to the product of selection functions above, in the sense that it will describe a very natural computational pattern that underlies our realizer, and helps us understand its behaviour.

## 6 The Functional Interpretation of $\mathrm{ZL}_{\mathrm{lex}}$—Part 1

The basic soundness proof of the functional interpretation (Theorem 4.1) guarantees that we are able to extract a program from any proof which can be formalised in $\mathrm{PA}^\omega + \mathrm{QF\text{-}AC}$. However, our formalisation of the minimal bad sequence construction involves something stronger, namely an instance of $\mathrm{ZL}_{\mathrm{lex}}$. The following three sections contain the chief novelty of our approach, namely the solution of the functional interpretation of $\mathrm{ZL}_{\mathrm{lex}}$ via a form of *open recursion*, which will allow us in Sect. 9 to extract a program witnessing $\mathrm{WQO}(=_{\mathbb{B},*})$.

So what exactly is the functional interpretation of $\mathrm{ZL}_{\mathrm{lex}}$? Let's begin by writing out the axiom in full, where now replace $B(u)$ with the piecewise formula $(\forall n) P([u](n))$, and for the remainder of the paper we now assume that $P(s)$ is quantifier-free, as it is in the case of Theorem 3.1. In order to avoid nested expressions such as $P([[v](m)](n))$ we will use the notation $\bar{P}(u, n) :\equiv P([u](n))$. Then $\mathrm{ZL}_{\mathrm{lex}}$ becomes

$$(\exists u)(\forall n)\bar{P}(u,n) \to (\exists v)((\forall n)\bar{P}(v,n) \wedge (\forall w \lhd_{\mathrm{lex}} v)(\exists n)\neg\bar{P}(w,n)).$$

Now, there is still an additional quantifier implicit in $(\forall w \lhd_{\mathrm{lex}} v)$, but note that

$$(\forall w \lhd_{\mathrm{lex}} v)A(w) \leftrightarrow (\forall m, w)(w_0 \lhd v_m \to A([v](m) * w))$$

and so $\mathrm{ZL}_{\mathrm{lex}}$ can be written out in a fully explicit form as

$$(\exists u)(\forall n)\bar{P}(u,n) \to (\exists v)((\forall n)\bar{P}(v,n) \wedge (\forall m, w)(w_0 \lhd v_m \to (\exists n)\neg\bar{P}([v](m) * w, n))). \tag{11}$$

Of course, we want to apply the functional interpretation to the negative translation of (11), which is equivalent to

$$(\exists u)(\forall n)\bar{P}(u,n) \to \neg\neg(\exists v)((\forall n)\bar{P}(v,n) \wedge (\forall m, w)(w_0 \lhd v_m \to (\exists n)\neg\bar{P}([v](m) * w, n))). \tag{12}$$

Since this is a rather intricate formula, let's break its interpretation up into pieces. Focusing on the conclusion first, and applying the interpretation under the double negation only, we obtain

$$\neg\neg(\exists v, \gamma^{\mathbb{N}\to X^{\mathbb{N}}\to\mathbb{N}})(\forall n, m, w)(\bar{P}(v,n) \wedge \underbrace{(w_0 \lhd v_m \to \neg\bar{P}([v](m) * w, \gamma m w))}_{C(v,\gamma,m,w)}) \tag{13}$$

where from now on we will use the abbreviation

$$C(v, \gamma, m, w) :\equiv w_0 \lhd v_m \to \neg\bar{P}([v](m) * w, \gamma m w)$$

as indicated in (13). Now, applying the functional interpretation to (13) and referring back to the discussion in Sect. 4.4.2 we arrive at

$$(\forall N, M, W)(\exists v, \gamma)(\bar{P}(v, Nv\gamma) \wedge C(v, \gamma, Mv\gamma, Wv\gamma)) \tag{14}$$

where $N, M : X^{\mathbb{N}} \to (\mathbb{N} \to X^{\mathbb{N}} \to \mathbb{N}) \to \mathbb{N}$ and $W : X^{\mathbb{N}} \to (\mathbb{N} \to X^{\mathbb{N}} \to \mathbb{N}) \to X^{\mathbb{N}}$. Substituting (14) back into (12) and referring to Sect. 4.4.3 our challenge is to witness the following expression:

$$(\forall u, N, M, W)(\exists n, v, \gamma)(\bar{P}(u, n) \to \bar{P}(v, Nv\gamma) \wedge C(v, \gamma, Mv\gamma, Wv\gamma))). \tag{15}$$

So what does the expression (15) *intuitively* mean? In Sect. 4.4 we characterised the functional interpretation as a translation which takes fundamentally non-constructive existence statements and converts them into 'approximate' existence statements, which in theory can be given a direct computational interpretation. In its original form, $\mathrm{ZL}_{\mathrm{lex}}$ simply states that

> if there exists a bad sequence $u$ then there exists a bad sequence $v$ which is minimal with respect to $\lhd_{\mathrm{lex}}$,

where we call $u$ bad whenever $(\forall n)\bar{P}(u, n)$ holds. Now, very roughly, we can read the interpreted statement (15) as saying something like

> for any sequence $u$ and counterexample functionals $N, M, W$, there exists $n, v$ and $\gamma$ such that $\bar{P}(u, n)$ implies that $v$ is approximately bad with respect to $N$, and $\gamma$ witnesses that it is approximately minimal with respect to $M$ and $W$.

When using $\mathrm{ZL}_{\mathrm{lex}}$ as a lemma in the proof of a $\Pi_2$ statement, as we do in Corollary 3.4, the task of extracting a program from this proof involves calibrating exactly what kind of approximations we need.

## 6.1 A Rough Idea

So how do we go about solving (15)—in other words computing a suitable $n$, $v$ and $\gamma$ in terms of $u$, $N$, $M$ and $W$? A natural idea might be to simply use trial and error, as follows. Given some initial sequence $u$, we could first just try $v := u$. Let's also set $\gamma := \gamma_u$, where $\gamma_u$ is some function that we will need to sensibly define later, and put $n := Nu\gamma_u$. Now suppose that $\bar{P}(u, Nu\gamma_u)$ holds. There are two possibilities: Either $u$ is approximately minimal in the sense that $C(u, \gamma_u, Mu\gamma_u, Wu\gamma_u)$ holds, and then we're done, or $\neg C(u, \gamma_u, Mu\gamma_u, Wu\gamma_u)$ i.e.

$$(Wu\gamma_u)_0 \lhd u_{Mug_u} \wedge \bar{P}([u](Mu\gamma_u) * Wu\gamma_u, \gamma(Mu\gamma_u)(Wu\gamma_u)).$$

But in this case, we have found a sequence $u_1 := [u](Mu\gamma_u) * Wu\gamma_u$ which is lexicographically less that $u$ and approximately bad, so could we just set $v := u_1$ and repeat this process, generating a sequence $u \rhd_{\mathrm{lex}} u_1 \rhd_{\mathrm{lex}} u_2 \rhd_{\mathrm{lex}} \ldots\ldots \rhd_{\mathrm{lex}} u_k$ until we reach some $u_k$ which works? Of course, there are a lot of details to be filled in here, in particular a formal definition of $\gamma$, but the aim of Sect. 8 will be to demonstrate that this informal idea *does* actually work.

However, the obvious problem we face is that we seem to be carrying out recursion over the non-wellfounded ordering $\rhd_{\text{lex}}$, and so first we must establish a set of conditions under which this kind of recursion is well-defined. This is the purpose of Sect. 7 which follows. Before we get into the technical details, though, we want to pause for a moment and explore the general pattern hinted at above, and introduce the notion of a learning procedure, which we have alluded to several times earlier.

## *6.2 Learning Procedures*

Our challenge in the next Sections is to take some initial sequence $u$ which is 'approximately bad' and produce a $v$ which is also approximately bad, but in addition approximately minimal. For simplicity, let's forget for a moment that we're working with infinite sequences and the lexicographic ordering, and just consider a set $X$ which comes equipped with two decidable predicates $P_0(x)$ and $C_0(x)$.

Of course, in our case $P_0(x)$ represents that $x$ is approximately bad while $C_0(x)$ represents that it's approximately minimal, but here everything is greatly simplified and we do not assume anything about these formulas beyond the following property, which states that if $x$ is not minimal then there must be some $y \prec x$ satisfying $P_0(y)$:

$$(\forall x)(\neg C_0(x) \to (\exists y \prec x) P_0(y)). \tag{16}$$

Therefore on an *abstract* level, the algorithmic problem we face is the following: Given some initial $x$ which satisfies $P_0(x)$, to find some $y$ which satisfies both $P_0(y)$ and $C_0(y)$, where the *failure* of $C_0(x)$ always leads to a 'better' guess $y$—this is captured by (16). In other words, using (16) we want to produce a $y$ satisfying

$$(\forall x)(P_0(x) \to (\exists y)(P_0(y) \wedge C_0(y))). \tag{17}$$

It is not too hard to come up with an algorithm which takes us from a realizer of (16) to a realizer of (17).

**Lemma 6.1** *Suppose that $\xi : X \to X$ is a function which satisfies*

$$(\forall x)(\neg C_0(x) \to x \succ \xi(x) \wedge P_0(\xi(x))). \tag{18}$$

*For any $x : X$, the* learning procedure $\mathcal{L}_{\xi,C_0}[x]$ *starting at $x$ denotes the sequence $(x_i)_{i\in\mathbb{N}}$ given by*

$$x_0 := x \quad \textit{and} \quad x_{i+1} := \begin{cases} x_i & \textit{if } C_0(x_i) \\ \xi(x_i) & \textit{otherwise.} \end{cases}$$

*Whenever $\succ$ is wellfounded, there exists some $k$ such that $C_0(x_k)$ holds, and we call the minimal such $x_k$ the limit of $\mathcal{L}_{\xi,C_0}[x]$, which we denote by*

$$\lim \mathcal{L}_{\xi,C_0}[x].$$

*Then the functional $\lambda x.\lim \mathcal{L}_{\xi,C_0}[x]$ is definable using wellfounded recursion over $\succ$, and realizes (17) in the sense that*

$$(\forall x)(P_0(x) \to P_0(\lim \mathcal{L}_{\xi,C_0}[x]) \wedge C_0(\lim \mathcal{L}_{\xi,C_0}[x])). \tag{19}$$

*Proof* To formally construct the limit, given $C_0$ and $\xi$ define the function $L_{\xi,C_0} : X \to X^*$ by

$$L_{\xi,C_0}(x) := \begin{cases} [x] & \text{if } C_0(x) \\ [x] * L_{\xi,C_0}(\xi(x)) & \text{otherwise,} \end{cases}$$

which is definable via wellfounded recursion over $\succ$ since the recursive call $L_{\xi,C_0}(\xi(x))$ is only made in the event that $\neg C_0(x)$ and so $x \succ \xi(x)$ by (18). A simple induction over the length of $L_{\xi,C_0}(x)$ then establishes that $\lim \mathcal{L}_{\xi,C_0}[x]$ is the last element of $L_{\xi,C_0}(x)$.

That the limit satisfies (19) essentially follows from the definition. If $P_0(x_i)$ but $\neg C_0(x_i)$ then we have $x_{i+1} = \xi(x_i)$ with $x_i \succ x_{i+1}$ and $P_0(x_{i+1})$. So it follows that if $P_0(x)$ then $P_0(x_i)$ for all $i \in \mathbb{N}$. Then by the existence of a limit $x_k$ satisfying $C_0(x_k)$ we're done, since we then have $P_0(x_k) \wedge C_0(x_k)$. □

Algorithms of the above kind can be characterised as 'learning procedures' because we start with some initial attempt $x_0$ for our minimal element, and either this works or it fails, in which case we replace $x_0$ with some 'improved' guess $x_0 \prec x_1$ which we have learned from the failure of $x_0$ and continue in this way until we have produced an attempt $x_k$ which works and satisfies $P_0(x_k) \wedge C_0(x_k)$.

The next two sections involve adapting this basic idea to the more complex situation of constructing a realizer for the functional interpretation of $\mathrm{ZL}_{\mathrm{lex}}$, where the predicates $P_0$ and $C_0$ will need to take into account the counterexample functionals which determine precisely what an approximation constitutes. Moreover, we will need to adapt Lemma 6.1 so that it applies to the non-wellfounded ordering $\rhd_{\mathrm{lex}}$.

Learning procedures as described above form the main subject of the author's paper [23], which in particular contains a solution to the functional interpretation of the least element principle for wellfounded $\succ$ that essentially forms a simple version of the realizer we construct here. Moreover, learning procedures even for certain non-wellfounded orderings are discussed in [23, Sect. 5], although none of this encompasses the variant of recursion over $\rhd_{\mathrm{lex}}$ which we require below. Note that these learning procedures should not be confused with the learning realizability of Aschieri and Berardi (e.g. [2]), which although based on a similar idea, takes place in a very different setting.

# 7 Recursion over $\rhd_{\text{lex}}$ in the Continuous Functionals

In order to give a functional interpretation of $\text{ZL}_{\text{lex}}$, it is necessary that we extend system $\mathsf{T}$ with some form of recursion over the relation $\rhd_{\text{lex}}$. Since $\rhd_{\text{lex}}$ is not wellfounded, it is clear that naively introducing a general recursor over $\rhd_{\text{lex}}$ will lead to problems. However, just as $\text{ZL}_{\text{lex}}$ is equivalent to an induction principle $\text{OI}_{\text{lex}}$ over $\rhd_{\text{lex}}$, which comes with the caveat that formulas must be *open* (cf. Sect. 3.1), we will show that we can define a recursor over $\rhd_{\text{lex}}$ which exists in continuous models of higher-type functionals, provided that we introduce an analogous restriction for the recursor.

The notion of recursion over $\rhd_{\text{lex}}$ is not new: In particular this forms the main topic of Berger's analysis of open induction in the framework of modified realizability [4]. However, the functional interpretation requires a non-trivial adaptation of these ideas, which is the main purpose of this section.

## 7.1 The Problem with Recursion over $\rhd_{lex}$

We begin by highlighting why a naive lexicographic recursor does not behave in the same way as Gödel's wellfounded recursors Rec, as identifying the problems provides some insight into how we can potentially circumvent them. Suppose that given some pair $(X, \rhd)$ where $X$ is a type and $\rhd$ a wellfounded decidable relation on $X$, together with output type $Y$, we add to our programming language $\mathsf{T}$ an open recursor $\text{ORec}_{(X,\rhd),Y}$ which has the defining equation

$$\text{ORec}^H_{(X,\rhd),Y}(u) =_Y Hu(\lambda n, v\ .\ \text{ORec}^H([u](n) * v) \text{ if } v_0 \lhd u_n)$$

where 'if $v_0 \lhd u_n$' is short for 'if $v_0 \lhd u_n$, else $0_Y$'. Does our recursor give rise to well-defined functionals?

Let's consider the very simple case $X = \mathbb{B}$ where $b_0 \rhd b_1$ only holds in the case $1 \rhd 0$, and set the output type $Y := \mathbb{N}$. Define the closed functional $\Phi : (\mathbb{B}^{\mathbb{N}} \to (\mathbb{N} \to \mathbb{B}^{\mathbb{N}} \to \mathbb{N}) \to \mathbb{N}) \to \mathbb{N}$ by

$$\Phi H := \text{ORec}^H_{(\mathbb{B},\rhd),\mathbb{N}}(\lambda k.1).$$

Then we can show that the type structure of all set-theoretic functionals is no longer a model of $\mathsf{T} + (\text{ORec}_{(\mathbb{B},\rhd),\mathbb{N}})$. To see this, consider the functional $H : \mathbb{B}^{\mathbb{N}} \to (\mathbb{N} \to \mathbb{B}^{\mathbb{N}} \to \mathbb{N}) \to \mathbb{N}$ defined by

$$Huf := \begin{cases} 1 + fn(0, 1, 1, \ldots) & \text{for the least } n \text{ with } u_n = 1 \\ 0 & \text{if no such } n \text{ exists.} \end{cases}$$

Suppose that $\Phi H = N$ for some natural number $N$. Then unwinding the defining equation of $\mathrm{ORec}^H_{(\mathbb{B},\rhd),\mathbb{N}}$ we get

$$N = \Phi H = 1 + \mathrm{ORec}^H(0, 1, 1, \ldots) = 2 + \mathrm{ORec}^H(0, 0, 1, 1, \ldots)$$
$$= \ldots = N + 1 + \mathrm{ORec}^H(\underbrace{0, \ldots, 0}_{N+1 \text{ times}}, 1, 1, \ldots) \geq N + 1,$$

a contradiction. Here it is not necessarily surprising that we run into problems. But suppose that we demand that $H$ be *continuous*, in the sense that we can determine the value of $Huf$ based on a finite initial segment of $u$ and $f$. Unfortunately, it turns out that if we increase the output type to $Y := \mathbb{N} \to \mathbb{N}$ then not even continuity (or indeed even computability) can save us: Let $G : \mathbb{B}^{\mathbb{N}} \to (\mathbb{N} \to \mathbb{B}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}) \to \mathbb{N}^{\mathbb{N}}$ be defined by

$$Gufn := 1 + fn(0, 1, 1, \ldots)(n + 1),$$

and let $N : \mathbb{N}$ be given by $N := \mathrm{ORec}^G(\lambda k.0)(0)$. Then similarly to before, we have

$$N = 1 + \mathrm{ORec}^G(0, 1, 1, \ldots)(1) = 2 + \mathrm{ORec}^G(0, 0, 1, 1, \ldots)(2)$$
$$= \ldots = N + 1 + \mathrm{ORec}^G(\underbrace{0, \ldots, 0}_{N+1 \text{ times}}, 1, 1, \ldots)(N + 1) \geq N + 1,$$

which is again inconsistent with the axioms of Peano arithmetic. So even though $N$ is a closed term in $\mathsf{T} + (\mathrm{ORec}_{(\mathbb{B},\rhd),\mathbb{N}^{\mathbb{N}}})$ of base type, there is no natural interpretation of $N$ in even in continuous models. So what does it take to ensure that recursion over $\rhd_{\mathrm{lex}}$ *does* have an interpretation in continuous models? To this end we will discuss two possible restrictions, namely:

- Leave the defining equation of the recursor unchanged but restrict $Y$ to being a base type.
- Allow $Y$ to be an arbitrary type but introduce an explicit 'control functional' into the defining equation.

The former is the approach taken by Berger [4] and works well in the setting of modified realizability. However, for the functional interpretation we need a recursor whose output type $Y$ can be arbitrary, and so we appeal to the second strategy which we will describe in detail in Sect. 7.3. However, to put our solution in context, first we will quickly sketch Berger's solution.

## 7.2 *The Continuous Functionals and Berger's Open Recursor*

In order to extend functional interpretations to subsystems of mathematical analysis, it is traditionally necessary to extend the usual interpreting calculus of functionals with a strong form of recursion, which is typically only satisfiable the *continuous* models. This was originally the case with Spector's bar recursion, and also here with our variants of open recursion.

In this section we assume a basic knowledge of the type structures of partial and total continuous functionals, as a full presentation here is beyond the scope of our paper. Continuous type structures of functionals were formally constructed from the 1960s onwards: The total continuous functionals being conceived simultaneously by Kleene [12] and Kreisel [16] and the partial model, developed independently by Scott [25] and Ershov [6]. Variants of the latter play an important role in domain theory, where in particular they are used to give a denotational semantics to abstract functional programming languages such as PCF. For an up-to-date presentation of these things and much more in this direction, the reader is encouraged to consult [17].

Very roughly, the continuous functionals $C^\omega_{X\to Y}$ of type $X \to Y$ consist of functionals $F$ from $X$ to $Y$ which satisfy the property that

> in order to determine a finite amount of information about $F(x)$ one only needs a finite amount of information about $x$,

where the notion of finiteness is made precise by introducing a suitable topology for each type. Note that continuity is a strictly weaker property than being *computable*: In particular *any* function $f : \mathbb{N} \to \mathbb{N}$ is continuous by definition, since $f(n)$ only depends on a natural number $n$, and natural numbers are here considered to be finite pieces of information. On the other hand, not all functionals $F : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ are continuous, in fact the continuous functionals $C^\omega_{\mathbb{N}^{\mathbb{N}}\to\mathbb{N}}$ of type 2 are precisely those such that for any $\alpha : \mathbb{N}^{\mathbb{N}}$ there exists some $N$ such that for all $\beta$, if $[\alpha](N) = [\beta](N)$ then $F(\alpha) = F(\beta)$. Both of the aforementioned properties can be generalised in the following way:

(i) The continuous functionals $C^\omega_{\mathbb{N}\to X}$ consist of *all* sequences $\mathbb{N} \to C^\omega_X$, and so in particular the type structure of continuous functionals is a model of countable dependent choice.
(ii) Any $F \in C^\omega_{X^{\mathbb{N}}\to\mathbb{N}}$ satisfies the following property:

$$\text{CONT} : (\forall\alpha)(\exists N)(\forall\beta)([\alpha](N) =_{X^*} [\beta](N) \to F(\alpha) = F(\beta)).$$

Note that for $X = \mathbb{N}$ this property is equivalent to $F$ being continuous, whereas for $X$ a higher type, it is strictly weaker (since $F$ could depend on an infinite amount of information from $\alpha(0)$ but still satisfy CONT, for example).

The *partial* continuous functionals $\hat{C}^\omega$ are similar to the total continuous functionals described above, with the crucial difference that they allow functionals which are undefined in places, and so the $\hat{C}^\omega_X$ are represented by a *domains* which in particular come equipped with a bottom element $\perp$ denoting an undefined value. The partial continuous functionals have the key property that every continuous functional $X \to X$ has a continuous fixed point, which means in particular that any recursively defined functional has a natural interpretation in $\hat{C}^\omega$ (although this need not be total). The partial continuous functionals are related to the total continuous functionals in that $C^\omega$ is the extensional collapse of the total elements of $\hat{C}^\omega$ [6]. What this means in practice is that in order to show that a recursively defined functional has an interpretation in $C^\omega$, it is enough to show that its interpretation in $\hat{C}^\omega$ as a fixpoint is total.

**Theorem 7.1** (Berger [4]) *Let $\rhd$ be a well-founded, decidable relation on $X$. Then any fixpoint of the defining equation of* $\mathrm{ORec}_{(X,\rhd),\mathbb{N}}$ *is total, and hence* $\mathrm{ORec}_{(X,\rhd),\mathbb{N}}$ *exists in the total continuous functionals $C^\omega$.*

*Proof* While in [4, Proposition 5.1] this is proven using a variant of open induction, we appeal to the classical minimal bad sequence construction, to emphasise already the deep connection with Nash–Williams' proof of Higman's lemma. Suppose for contradiction that there are total arguments $H$ and $u$ such that $\mathrm{ORec}^H(u)$ is not total. Using dependent choice, which is valid for total objects of sequence type, construct the minimal bad sequence $v$ of total elements of type $X$ as follows:

If $[v_0, \dots, v_{k-1}]$ has already been constructed, define $v_k$ to be a total element of $\hat{C}^\omega_X$ such that $\mathrm{ORec}^H([v_0, \dots, v_{k-1}, v_k] * w)$ is not total for some total extension $w$, but $\mathrm{ORec}^H([v_0, \dots, v_{k-1}, a] * w)$ is total for all total $w$ whenever $a \lhd v_k$.

Now consider $\mathrm{ORec}^H(v) = Hv(\lambda n, w \,.\, \mathrm{ORec}^H([v](n) * w) \text{ if } w_0 \lhd v_n) = H\alpha$ where

$$\alpha_n := \langle v_n, \lambda w.\mathrm{ORec}^H([v](n) * w) \text{ if } w_0 \lhd v_n \rangle,$$

and note that we use a slight abuse of types here, informally identifying the type $X^\mathbb{N} \to (\mathbb{N} \to X^\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ of $H$ with $(X \times (X^\mathbb{N} \to \mathbb{N}))^\mathbb{N} \to \mathbb{N}$. But by minimality of $v$, the sequence $\alpha_n$ is total, and hence $H\alpha$ is total and then by CONT applied to the total objects $H$ and $\alpha$ there exists some $N$ such that whenever $[\alpha](N) = [\beta](N)$ then $H\alpha = H\beta$. But now consider the sequence $[v](N)$. By construction there exists some $w$ such that $\mathrm{ORec}^H([v](N) * w)$ is not total. But $\mathrm{ORec}^H([v](N) * w) = H\beta$ for

$$\beta_n := \langle ([v](N) * w)_n, \lambda w'.\mathrm{ORec}^H([[v](N) * w](n) * w') \text{ if } w'_0 \lhd ([v](N) * w)_n \rangle$$

and we have $\alpha_n = \beta_n$ for all $n < N$ and hence $H\beta = H\alpha$ which is total, a contradiction. Hence our original assumption was wrong and $\mathrm{ORec}^H(u)$ must be total, and since $H$ and $u$ were arbitrary we have that ORec is total. □

We have given this proof in great detail as we want to compare it to the corresponding totality proof of our explicit open recursor given in the next section. We now conclude our overview of Berger's open recursion by stating the main result of [4], namely:

**Theorem 7.2** (Berger [4]) *There is a functional definable in* $\mathsf{T} + (\mathrm{ORec}_{(\mathsf{X},\rhd),\mathbb{N}})$ *such that $\Phi$ satisfies the modified realizability interpretation of the axiom of open induction* $\mathrm{OI}_{lex}$ *for $\Sigma^0_1$-piecewise formulas, provably in* $\mathrm{PA}^\omega + \mathrm{CONT} + \mathrm{OI}_{lex} + (\mathrm{ORec}_{(X,\rhd),\mathbb{N}})$.

Theorem 8.3 below forms an analogue of this for the functional interpretation.

## 7.3 The Explicitly Controlled Open Recursor

Berger's variant of open recursion uses in an essential way the fact that total continuous functions of type $Z^{\mathbb{N}} \to \mathbb{N}$ (where $Z$ is arbitrary) only consider a finite initial segment of their input. In this way we avoid the problems encountered earlier in the chapter. However, as we will see, having open recursive functionals whose output type $Y$ is arbitrary is essential for the functional interpretation of $\mathrm{ZL}_{\mathrm{lex}}$, and Berger's variant is no longer total in this case as we cannot rely on continuity to 'implicitly' control the recursion. Therefore we require some other way of ensuring that the recursor only depends on a finite initial segment of its input. We accomplish this by adding an additional parameter $F$ to the recursor which is responsible for 'explicitly' controlling the recursion, in a sense that will be made clear below. We start with some definitions.

**Definition 7.3** Suppose that $\alpha : Z^{\mathbb{N}}$ and $m : \mathbb{N}$. Then the infinite sequence $[\alpha]_m : Z^{\mathbb{N}}$ is defined by

$$[\alpha]_m := \lambda n. \begin{cases} \alpha_n & \text{if } n < m \\ 0_Z & \text{otherwise.} \end{cases}$$

Now suppose in addition that $F : Z^{\mathbb{N}} \to \mathbb{N}$. Then the infinite sequence $\{\alpha\}_F : Z^{\mathbb{N}}$ is defined by

$$\{\alpha\}_F := \lambda n. \begin{cases} 0_Z & \text{if } (\exists m \leq n)(F([\alpha]_m) < m) \\ \alpha_n & \text{otherwise.} \end{cases}$$

Note that both $[\alpha]_m$ and $\{\alpha\}_F$ are primitive recursively definable.

Note that the sequence $\{\alpha\}_F$ uses the well-known stopping condition $F([\alpha]_m) < m$ due to Spector [27], which ensures that his variant of bar recursion is well-founded.

**Lemma 7.4** *Given some $F : Z^{\mathbb{N}} \to \mathbb{N}$ and $\alpha : Z^{\mathbb{N}}$, whenever there exists some $m : \mathbb{N}$ such that $F([\alpha]_m) < m$ then*

$$\{\alpha\}_F = [\alpha]_{m_0}$$

*where $m_0$ is the least such $m$. If no such $m$ exists then $\{\alpha\}_F = \alpha$.*

*Proof* This follows directly from the definition: For the first case, by minimality of $m_0$ we have $\{\alpha\}_F(n) = \alpha_n$ for all $n < m_0$, and $\{\alpha\}_F(n) = 0_Z$ otherwise, which is exactly the definition of $[\alpha]_{m_0}$. □

**Lemma 7.5** *For any functional $F : Z^{\mathbb{N}} \to \mathbb{N}$ satisfying* CONT*, then for each $\alpha : Z^{\mathbb{N}}$ there exists some $m$ such that $F([\alpha]_m) < m$.*

*Proof* Suppose that $N$ is the point of continuity of $F$ which exists by CONT, and define $m := \max\{N, F\alpha + 1\}$. Then $[[\alpha]_m](N) = [\alpha](m)$ since $N \leq m$, and therefore $F([\alpha]_m) = F\alpha < m$. □

**Theorem 7.6** *Given* $F : Z^{\mathbb{N}} \to \mathbb{N}$ *and* $\alpha : Z^{\mathbb{N}}$, *the following facts are provable assuming* CONT*:*

*(i)* $\{\alpha\}_F = [\alpha]_{m_0}$ *where* $m_0$ *satisfies* $F([\alpha]_{m_0}) < m_0$ *and is the least such number;*
*(ii)* *for any* $\beta$ *satisfying* $[\alpha](m_0) = [\beta](m_0)$ *we have* $\{\alpha\}_F = \{\beta\}_F$*;*
*(iii)* $\{\{\alpha\}_F\}_F = \{\alpha\}_F$.

*Proof* Part (i) follows directly from Lemma 7.4 together with Lemma 7.5. For part (ii), we observe that for all $n \leq m_0$ we have $[\alpha]_n = [\beta]_n$, from which it follows that the first $m_1$ satisfying $F([\beta]_{m_1}) < m_1$ is just $m_0$. Therefore by part (i) again we have $\{\beta\}_F = [\beta]_{m_0} = [\alpha]_{m_0} = \{\alpha\}_F$. Part (iii) now follows easily, since by part (i) we have $\{\{\alpha\}_F\}_F = \{[\alpha]_{m_0}\}_F$, and since $[[\alpha]_{m_0}](m_0) = [\alpha](m_0)$ then $\{[\alpha]_{m_0}\}_F = \{\alpha\}_F$ by part (ii). □

Now we are ready to define our 'explicit' recursor. Given $H : (X \times (X^{\mathbb{N}} \to Y))^{\mathbb{N}} \to Y$ and $F : (X \times (X^{\mathbb{N}} \to Y))^{\mathbb{N}} \to \mathbb{N}$, we define

$$\mathrm{EORec}^{H,F}_{(X,\rhd),Y}(u) =_Y H(\{\alpha\}_F) \quad \text{for} \quad \alpha :=_{(X\times(X^{\mathbb{N}}\to Y))^{\mathbb{N}}} \lambda n\ .\ \langle u_n, \lambda v\ .\ \mathrm{EORec}^{H,F}([u](n) * v) \text{ if } v_0 \lhd u_n\rangle.$$

This is a form of lexicographic recursion just as before, but with the crucial difference that the recursor now comes equipped with some additional functional $F : Z^{\mathbb{N}} \to \mathbb{N}$ which determines how much of the sequence $\alpha$ is 'relevant'. As soon as we have found some $m$ satisfying the condition $F([\alpha]_m) < m$ then we declare that we are not interested in $\alpha_n$ for $n \geq m$. Our introduction of this 'control' functional $F$ allows us to provide an analogue of CONT for $H$, even though the output type of $H$ is arbitrary.

**Lemma 7.7** *Suppose that* $H : Z^{\mathbb{N}} \to Y$ *and that* $F : Z^{\mathbb{N}} \to \mathbb{N}$ *satisfies* CONT*. Then* $H$ *satisfies the following property:*

$$\mathrm{CONT}^* : (\forall\alpha)(\exists N)(\forall\beta)([\alpha](N) = [\beta](N) \to H(\{\alpha\}_F) =_Y H(\{\beta\}_F)).$$

*Proof* Let $m_0$ be the least number satisfying $F([\alpha]_{m_0}) < m_0$, which exists by CONT, and define $N := m_0$. Then if $[\alpha](m_0) = [\beta](m_0)$ then $\{\alpha\}_F = \{\beta\}_F$ by part (ii) above, and therefore $H(\{\alpha\}_F) = H(\{\beta\}_F)$. □

We can use this result to show, analogously to Theorem 7.1, that $\mathrm{EORec}_{(X,\rhd),Y}$ exists in the total continuous functionals for any $Y$.

**Theorem 7.8** *Let* $\rhd$ *be a well-founded, decidable relation on* $X$*. Then the fixpoint of the defining equation of* $\mathrm{EORec}_{(X,\rhd),Y}$ *is total, and hence* $\mathrm{EORec}_{(X,\rhd),Y}$ *exists in the total continuous functionals* $C^{\omega}$.

*Proof* This follows analogously to the proof of Theorem 7.1. Suppose for contradiction that there are total arguments $H$, $F$ and $u$ such that $\mathrm{EORec}^{H,F}(u)$ is not total. Using dependent choice, construct a minimal bad sequence $v$ as follows:

> If $[v_0, \dots, v_{k-1}]$ has already been constructed, define $v_k$ to be a total element of $\hat{C}^{\omega}_X$ such that $\mathrm{EORec}^{H,F}([v_0, \dots, v_{k-1}, v_k] * w)$ is not total for some total extension $w$, but $\mathrm{EORec}^{H,F}([v_0, \dots, v_{k-1}, a] * w)$ is total for all total $w$ whenever $a \lhd v_k$.

Now consider $\text{ERec}^{H,F}(v) = H(\{\alpha\}_F)$ for

$$\alpha := \lambda n \; . \; \langle v_n, \lambda w \; . \; \text{EORec}^{H,F}([v](n) * w) \text{ if } w_0 \lhd v_n \rangle.$$

Then by construction of $v$, $\alpha$ and hence $H(\{\alpha\}_F)$ must be total, and by CONT* applied to the total objects $\alpha$, $F$ and $H$ hence there exists some $N$ such that for any total $\beta : (X \times (X^{\mathbb{N}} \to Y))^{\mathbb{N}}$, if $[\alpha](N) = [\beta](N)$ then $H(\{\alpha\}_F) = H(\{\beta\}_F)$. Now consider the sequence $[v](N)$. By construction there exists some $w$ such that $\text{EORec}^{H,F}([v](N) * w)$ is not total. But $\text{EORec}^{H,F}([v](N) * w) = H(\{\beta\}_F)$ where

$$\beta := \lambda n . \langle ([v](N) * w)_n, \lambda w' \; . \; \text{EORec}^{H,F}([[v](N) * w](n) * w') \text{ if } w'_0 \lhd ([v](N) * w)_n \rangle$$

and we have $\beta(n) = \alpha(n)$ for all $n < N$ and hence by CONT* we have $\text{EORec}^{H,F}$ $([v](N) * w) = H(\{\beta\}_F) = H(\{\alpha\}_F)$ which is total, a contradiction. Therefore $\text{EORec}^{H,F}(u)$ must be total, and since $H$, $F$ and $u$ were arbitrary total objects then EORec is total. □

## 8 The Functional Interpretation of $\text{ZL}_{\text{lex}}$—Part 2

We will now make formal the intuitive idea presented in Sect. 6. We begin by setting up an analogue of Lemma 6.1, but this time the objects $x$ of our learning procedures are sequences $u : X^{\mathbb{N}}$ (and so $P_0(u)$ and $C_0(u)$ are decidable predicates over $X^{\mathbb{N}}$) and $u_{i+1}$ is defined as $\xi(\{u_i\}_\phi)$ for some $\phi : X^{\mathbb{N}} \to \mathbb{N}$. As a result, we end up with a sequence of the form

$$u_0 \mapsto \{u_0\}_\phi \rhd_{\text{lex}} u_1 \mapsto \{u_1\}_\phi \rhd_{\text{lex}} u_2 \mapsto \dots$$

and so in order to guarantee that $P_0(u_i)$ holds for all $i$ we will require an additional condition, namely that the property $P_0$ is preserved under the map $\{\cdot\}_\phi : X^{\mathbb{N}} \to X^{\mathbb{N}}$. We will now just state and prove the result, but the reader is strongly encouraged to simultaneously refer back to the much simpler Lemma 6.1 and its proof, not only so that it is easier to grasp what is going on here, but because the differences in the formulation of the two lemmas are extremely informative.

*Remark 8.1* For the remainder of the paper, we request that the canonical element $0_X$ of type $X$ is minimal with respect to $\rhd$. This condition is not essential and could be circumvented by other means, but it makes what follows a little easier and allows us to avoid some additional syntax. In practice this assumption is completely benign, and in particular in this chapter where our type $X$ will actually be a type $X^*$ of finite words and $\lhd$ will denote the prefix relation, then the normal choice of $0_{X^*} = []$ is also minimal.

**Lemma 8.2** *Suppose that* $\xi : X^{\mathbb{N}} \to X^{\mathbb{N}}$ *is defined by*

$$\xi(u) := [u](\xi_0(u)) * \xi_1(u)$$

*where* $\xi_0 : X^{\mathbb{N}} \to \mathbb{N}$ *and* $\xi_1 : X^{\mathbb{N}} \to X^{\mathbb{N}}$, *and that* $\xi$ *satisfies*

$$(\forall u)(\neg C_0(u) \to u_{\xi_0(u)} \rhd \xi_1(u)_0 \wedge P_0(\xi(u))). \tag{20}$$

*Moreover, suppose that* $\phi : X^{\mathbb{N}} \to \mathbb{N}$ *is an additional functional which satisfies*

$$(\forall u)(P_0(u) \to P_0(\{u\}_\phi)). \tag{21}$$

*For any* $u : X^{\mathbb{N}}$, *the controlled learning procedure* $\mathcal{LC}^{\phi}_{\xi,C_0}[u]$ *starting at* $u$ *denotes the sequence* $(u_i)_{i\in\mathbb{N}}$ *given by*

$$u_0 := u \quad \text{and} \quad u_{i+1} := \begin{cases} \{u_i\}_\phi & \text{if } C_0(\{u_i\}_\phi) \\ \xi(\{u_i\}_\phi) & \text{otherwise.} \end{cases}$$

*Then provably from* CONT, *firstly there always exists some* $k$ *such that* $C_0(\{u_k\}_\phi)$ *holds, and we call* $\{u_k\}_\phi$ *for the minimal such* $u_k$ *the limit of* $\mathcal{LC}^{\phi}_{\xi,C_0}[u]$, *which we denote by*

$$\lim \mathcal{LC}^{\phi}_{\xi,C_0}[u],$$

*secondly the functional* $\lambda u.\lim \mathcal{LC}^{\phi}_{\xi,C_0}[u]$ *is definable in* $\mathsf{T} + (\mathrm{EORec}_{(\mathsf{X},\rhd)})$, *and finally we have*

$$(\forall u)(P_0(u) \to P_0(\lim \mathcal{LC}^{\phi}_{\xi,C_0}[u]) \wedge C_0(\lim \mathcal{LC}^{\phi}_{\xi,C_0}[u])). \tag{22}$$

*Proof* We first formally construct the limit functional, which the reader can skip if they like since this is nothing more that a somewhat intricate unwinding of definitions. First, we define $L^{\phi}_{\xi,C_0} : X^{\mathbb{N}} \to (X^{\mathbb{N}})^*$ by $L^{\phi}_{\xi,C_0}(u) := \mathrm{EORec}^{H,F}_{(X,\rhd),(X^{\mathbb{N}})^*}(u)$ where

$$F\alpha := \phi(\alpha_0)$$

$$H\alpha := \begin{cases} [\alpha_0] & \text{if } C_0(\alpha_0) \\ [\alpha_0] * \alpha_1 \xi_0(\alpha_0)\xi_1(\alpha_0) & \text{otherwise.} \end{cases}$$

Here we denote by $\alpha_0$ the sequence $\lambda n.\pi_0\alpha(n)$ and similarly for $\alpha_1$. Then unwinding the definition, we have $L^{\phi}_{\xi,C_0}(u) = H(\{\alpha\}_F)$ for $\alpha = \lambda n.\langle u_n, \lambda v.L^{\phi}_{\xi,C_0}([u](n) * v) \text{ if } v_0 \lhd u_n\rangle$. But since $F\alpha$ depends only on the first component $\alpha_0 = u$, we have (by Lemma 7.6) $\{\alpha\}_F = [\alpha]_{m_0}$ where $m_0$ is the least number satisfying $F([\alpha]_{m_0}) = \phi([u]_{m_0}) < m_0$. This means that $\{u\}_\phi = [u]_{m_0}$ and so $(\{\alpha\}_F)_0 = ([\alpha]_{m_0})_0 = [u]_{m_0} = \{u\}_\phi$ and $(\{\alpha\}_F)_1 = \lambda n < m_0, v.L^\phi_{\xi,C_0}([u](n) * v)$ if $v_0 \lhd u_n$, and so

$$H(\{\alpha\}_F) = \begin{cases} [\{u\}_\phi] & \text{if } C_0(\{u\}_\phi) \\ [\{u\}_\phi] * (\{\alpha\}_F)_1 \xi_0(\{u\}_\phi)\xi_1(\{u\}_\phi) & \text{otherwise.} \end{cases}$$

But now by (20), if $\neg C_0(\{u\}_\phi)$ then $(\{u\}_\phi)_{\xi_0(\{u\}_\phi)} \rhd \xi_1(\{u\}_\phi)_0$, and since $0_X$ was chosen to be minimal with respect to $\rhd$ (cf. Remark 8.1) this can only mean that $\xi_0(\{u\}_\phi) < m_0$ (else $(\{u\}_\phi)_{\xi_0(\{u\}_\phi)} = 0_X$) and therefore $(\{u\}_\phi)_{\xi_0(\{u\}_\phi)} = u_{\xi_0(\{u\}_\phi)}$. Substituting all this information into the $(\{\alpha\}_F)_1$ we have

$$(\{\alpha\}_F)_1\xi_0(\{u\}_\phi)\xi_1(\{u\}_\phi) = L^\phi_{\xi,C_0}([\{u\}_\phi](\xi_0(\{u\}_\phi)) * \xi_1(\{u\}_\phi)) = L^\phi_{\xi,C_0}(\xi(\{u\}_\phi)).$$

So to summarise, the functional $L^\phi_{\xi,C_0}$ satisfies (repressing subscripts)

$$L^\phi_{\xi,C_0}(u) = \begin{cases} [\{u\}_\phi] & \text{if } C_0(\{u\}_\phi) \\ [\{u\}_\phi] * L^\phi_{\xi,C_0}(\xi(\{u\}_\phi)) & \text{otherwise,} \end{cases}$$

and so by induction on the length of $L^\phi_{\xi,C_0}(u)$ one establishes that $\lim \mathcal{L}C^\phi_{\xi,C_0}[u]$ exists and is the last element of $L^\phi_{\xi,C_0}(u)$.

To verify (22) is similar to the proof of Lemma 6.1. We first show by induction that if $P_0(u)$ holds then $P_0(\{u_i\}_\phi)$ holds for all $i \in \mathbb{N}$. For $i = 0$ this follows by (21) applied to $u = u_0$, and otherwise if $P_0(\{u_i\}_\phi)$ is true then either $\{u_{i+1}\}_\phi = \{\{u_i\}_\phi\}_\phi = \{u_i\}_\phi$ by Lemma 7.6 or $\neg C(\{u_i\}_\phi)$ and then by (20) we have $P_0(u_{i+1})$ and hence $P_0(\{u_{i+1}\}_\phi)$ by (21). Therefore, if $\{u_k\}_\phi$ is the limit of the learning procedure, then $P_0(\{u_k\}_\phi) \wedge C_0(\{u_k\}_\phi)$ holds, and we're done. □

Our final step is now to produce a realizer for the functional interpretation of $\mathrm{ZL}_{\mathrm{lex}}$. Let's briefly recall from Sect. 6 what this means: We are given as input a sequence $\bar{u}$ (we use this new notation as we want $u$ to denote a separate variable below), a pair of functionals $M, N : X^{\mathbb{N}} \to (\mathbb{N} \to X^{\mathbb{N}} \to \mathbb{N}) \to \mathbb{N}$, together with $W : X^{\mathbb{N}} \to (\mathbb{N} \to X^{\mathbb{N}} \to \mathbb{N}) \to X^{\mathbb{N}}$, and we must produce some $n : \mathbb{N}$, $v : X^{\mathbb{N}}$ and $\gamma : (X^{\mathbb{N}} \to \mathbb{N})^{\mathbb{N}}$ satisfying

$$\bar{P}(\bar{u}, n) \to \bar{P}(v, Nv\gamma) \wedge C(v, \gamma, Mv\gamma, Wv\gamma), \tag{23}$$

where as before $C(v, \gamma, m, w) :\equiv w_0 \lhd v_m \to \neg\bar{P}([v](m) * w, \gamma m w)$. We first need some definitions. Define the functional $\Psi^N : X^{\mathbb{N}} \to \mathbb{N}$ by

$$\Psi^N(u) := \mathrm{EORec}^{\tilde{N},\tilde{N}}_{(X,\rhd),\mathbb{N}}(u)$$

where $\tilde{N}\alpha := N\alpha_0\alpha_1$. Using $\Psi$, for each $u : X^{\mathbb{N}}$ define $\gamma_u : \mathbb{N} \to X^{\mathbb{N}} \to \mathbb{N}$ by

$$\gamma_u := \lambda n, v \,.\, \Psi^N([u](n) * v) \text{ if } v_0 \lhd u_n.$$

Finally, define parameters $\phi : X^{\mathbb{N}} \to \mathbb{N}, \xi_0 : X^{\mathbb{N}} \to \mathbb{N}$ and $\xi_1 : X^{\mathbb{N}} \to X^{\mathbb{N}}$, together with predicates $P_0$ and $C_0$, by

$$\begin{aligned}
\phi(u) &:= N u \gamma_u \\
\xi_0(u) &:= M u \gamma_u \\
\xi_1(u) &:= W u \gamma_u \\
P_0(u) &:\equiv \bar{P}(\bar{u}, \Psi^N(\bar{u})) \to \bar{P}(u, \Psi^N(u)) \\
C_0(u) &:\equiv C(u, \gamma_u, M u \gamma_u, W u \gamma_u).
\end{aligned}$$

Now it is perhaps becoming clear to the reader what will come next: We will set up a controlled learning procedure $\mathcal{LC}^{\phi}_{\xi, C_0}[\bar{u}]$ on these parameters exactly as in Lemma 8.2, and the limit $v := \{u_k\}_\phi$ of $\mathcal{LC}^{\phi}_{\xi, C_0}[\bar{u}]$ will satisfy $P_0(v) \wedge C_0(v)$, or in other words, $\bar{P}(\bar{u}, \Psi^N(\bar{u})) \to \bar{P}(v, \Psi^N(v))$ and $C(v, \gamma_v, M v \gamma_v, W v \gamma_v)$, from which we will be able to construct our realizer of $\mathrm{ZL}_{\mathrm{lex}}$. Let's make this formal.

**Theorem 8.3** *Let $n$, $v$ and $\gamma$ be defined in terms of $\bar{u}$, $N$, $M$ and $W$ by*

$$\begin{aligned}
n &:= \Psi^N(\bar{u}) \\
v &:= \lim \mathcal{LC}^{\phi}_{\xi, C_0}[\bar{u}] \\
\gamma &:= \gamma_v.
\end{aligned}$$

*Then (provably in* CONT*) these satisfy (23) and therefore solve the functional interpretation of* $\mathrm{ZL}_{lex}$.

*Proof* We use Lemma 8.2, which means that we must check that each of (20) and (21) hold for our choice of $P_0$ and $C_0$, i.e. we must prove

$$(\forall u)[\neg C(u, \gamma_u, M u \gamma_u, W u \gamma_u) \to u_{\xi_0(u)} \rhd \xi_1(u)_0 \wedge (\bar{P}(\bar{u}, \Psi^N(\bar{u})) \to \bar{P}(\xi(u), \Psi^N(\xi(u))))] \tag{24}$$

and

$$(\forall u)[(\bar{P}(\bar{u}, \Psi^N(\bar{u})) \to \bar{P}(u, \Psi^N(u))) \to (\bar{P}(\bar{u}, \Psi^N(\bar{u})) \to \bar{P}(\{u\}_\phi, \Psi^N(\{u\}_\phi)))]. \tag{25}$$

For the first condition, note that $\neg C(u, \gamma_u, M u \gamma_u, W u \gamma_u)$ is just $\neg C(u, \gamma_u, \xi_0(u), \xi_1(u))$, which implies both $\xi_1(u)_0 \lhd u_{\xi_0(u)}$ and $\bar{P}([u](\xi_0(u)) * \xi_1(u), \gamma_u \xi_0(u) \xi_1(u))$. But since

$$[u](\xi_0(u)) * \xi_1(u) = \xi(u) \text{ and } \gamma_u \xi_0(u) \xi_1(u) = \Psi^N(\xi(u))$$

we have established $\bar{P}(\xi(u), \Psi^N(\xi(u)))$, and hence the conclusion of (24).

The second condition is more subtle: Either $\neg\bar{P}(\bar{u}, \Psi^N(\bar{u}))$ and we're done, or it suffices to prove $\bar{P}(u, \Psi^N(u)) \to \bar{P}(\{u\}_\phi, \Psi^N(\{u\}_\phi))$. We now need to unwind the definition of $\Psi^N(u)$: First note that we have

$$\Psi^N(u) = \tilde{N}(\{\alpha\}_{\tilde{N}})$$

where (using the definition of $\gamma_u$)

$$\alpha := \lambda n.\langle u_n, \lambda v \ . \ \Psi^N([u](n) * v) \text{ if } v_0 \lhd u_n\rangle = \lambda n.\langle u_n, \gamma_{u,n}\rangle.$$

and so in particular we have (by the definitions of $\tilde{N}$ and $\phi$)

$$\tilde{N}([\alpha]_m) = N([u]_m)([\gamma_u]_m) = N([u]_m)(\gamma_{[u]_m}) = \phi([u]_m) \tag{26}$$

where for the central equality we use the assumption that $0_X$ is minimal with respect to $\rhd$ and so

$$\begin{aligned}
\gamma_{[u]_m} &= \lambda n, v.\Psi^N([[u]_m](n) * v) \text{ if } v_0 \lhd ([u]_m)_n \\
&= \lambda n \ . \begin{cases} \lambda v.\Psi^N([u](n) * v) \text{ if } v_0 \lhd u_n & \text{if } n < m \\ \lambda v.\Psi^N([u](n) * v) \text{ if } v_0 \lhd 0_X & \text{otherwise} \end{cases} \\
&= \lambda n \ . \begin{cases} \lambda v.\gamma_u n v & \text{if } n < m \\ \lambda v.0 & \text{otherwise} \end{cases} \\
&= [\gamma_u]_m.
\end{aligned}$$

Let $m_0$ be the least number such that $\tilde{N}([\alpha]_{m_0}) < m_0$ (which exists since we are assuming CONT), and therefore by (26) the also the least number such that $\phi([u]_{m_0}) < m_0$. Then by Lemma 7.6 we have that for all $u$:

$$\Psi^N(u) = \tilde{N}(\{\alpha\}_{\tilde{N}}) \overset{L.\,7.6(i)}{=} \tilde{N}([\alpha]_{m_0}) \overset{(26)}{=} \phi([u]_{m_0}) \overset{L.\,7.6(i)}{=} \phi(\{u\}_\phi) = N\{u\}_\phi\gamma_{\{u\}_\phi}. \tag{27}$$

In particular, by Lemma 7.6 (iii) we have

$$\Psi^N(\{u\}_\phi) \overset{(27)}{=} \phi(\{\{u\}_\phi\}_\phi) \overset{L.\,7.6(iii)}{=} \phi(\{u\}_\phi) \overset{(27)}{=} \Psi^N(u) \tag{28}$$

and

$$\Psi^N(u) \overset{(27)}{=} \tilde{N}([\alpha]_{m_0}) < m_0 \tag{29}$$

and therefore

$$[\{u\}_\phi](\Psi^N(\{u\}_\phi)) \overset{(28)}{=} [\{u\}_\phi](\Psi^N(u)) \overset{L.\,7.6(i)}{=} [[u]_{m_0}](\Psi^N(u)) \overset{(29)}{=} [u](\Psi^N(u))$$

and thus $P([u](\Psi^N(u)))$ implies $P([\{u\}_\phi](\Psi^N(\{u\}_\phi)))$, which establishes (25).

It now follows from Lemma 8.2 that

$$P_0(\bar{u}) \to P_0(v) \wedge C_0(v)$$

and since $P_0(\bar{u})$ is trivially true we have established

$$\bar{P}(\bar{u}, \Psi^N(\bar{u})) \to \bar{P}(v, \Psi^N(v)) \quad \text{and} \quad C(v, \gamma_v, Mv\gamma_v, Wv\gamma_v).$$

We can now prove (23). Suppose that $\bar{P}(u, n)$ holds. Then since $n = \Psi^N(\bar{u})$ from the left hand side we have $\bar{P}(v, \Psi^N(v))$. Now, since $v = \{u_k\}_\phi$ for some element in the learning procedure $\mathcal{LC}^\phi_{\xi,C_0}[\bar{u}]$, by (28) we have

$$\Psi^N(v) = \Psi^N(\{u_k\}_\phi) = \phi(\{u_k\}_\phi) = N\{u_k\}_\phi \gamma_{\{u_k\}_\phi} = Nv\gamma_v = Nv\gamma$$

and so we have established $\bar{P}(v, Nv\gamma)$. Then, since $C(v, \gamma, Mv\gamma, Wv\gamma)$ is given to us automatically, we have proven

$$\bar{P}(u, n) \to \bar{P}(v, Nv\gamma) \wedge C(v, \gamma, Mv\gamma, Wv\gamma)$$

which is exactly (23), and so we're done. □

The results of this section mark the technical climax of the paper, and in particular form our broadest and most widely applicable contribution. While the proofs above are perhaps somewhat difficult to navigate, it is important to emphasise that most of the technical details are bureaucratic in nature, in the unwinding of all the definitions and the careful use of Lemma 7.6. The *intuition* behind our realizer, on the other hand, should hopefully be clear from the somewhat more informal discussion in Sect. 6. In any case, now that the hard work is done, a computational interpretation of Nash–William's proof of Higman's lemma follows relatively easily.

## 9 Interpreting the Proof of WQO($=_{\mathbb{B},*}$)

We are now finally ready to produce our realizer for the statement that $=_{\mathbb{B}^*}$ is a WQO. In fact we do something more general, namely give a computational interpretation to the proof that $\text{WQO}_{\text{seq}}(\preceq) \to \text{WQO}(\preceq_*)$, which is valid for *any* well quasi-order $\preceq$. Recall that the functional interpretation of $\text{WQO}_{\text{seq}}(\preceq)$ is given by

$$(\forall x^{X^{\mathbb{N}}}, \omega^{(\mathbb{N}\to\mathbb{N})\to\mathbb{N}})(\exists g)(\forall i < j \le \omega g)(g(i) < g(j) \wedge x_{g(i)} \preceq x_{g(j)}). \tag{30}$$

Lemma 3.2 makes precise exactly how we use the assumption $\text{WQO}_{\text{seq}}(\preceq)$ to prove $\text{WQO}(\preceq_*)$: Namely given a hypothetical minimal bad sequence of words $v$, we take the sequence $\bar{v}$ and require that our monotone subsequence $g$ be valid up to the point $k$, where $k$ is such that $P(w, k)$ holds for $w = [v](g(0)) * \tilde{v}_g$ as defined in Lemma 3.2. If minimality of $v$ is witnessed by some functional $\gamma$ then such a $k$ would be given by $\gamma(g0)(\tilde{v}_g)$. This motivates the following:

**Lemma 9.1** *Suppose that G is a realizer for (30):*

$$(\forall x, \omega)(\forall i < j < \omega G_{x,\omega})(G_{x,\omega}(i) < G_{x,\omega}(j) \wedge x_{G_{x,\omega}(i)} \preceq x_{G_{x,\omega}(j)}). \qquad (31)$$

*Then from this we can construct a functional $H : X^{\mathbb{N}} \to (\mathbb{N} \to X^{\mathbb{N}} \to \mathbb{N}) \to \mathbb{N}^{\mathbb{N}}$ satisfying*

$$(\forall v, \gamma)(\forall i < j < \gamma(H_{v,\gamma}(0))(\tilde{v}_{H_{v,\gamma}}))(H_{v,\gamma}(i) < H_{v,\gamma}(j) \wedge \bar{v}_{H_{v,\gamma}(i)} \preceq \bar{v}_{H_{v,\gamma}(j)}) \qquad (32)$$

*where $v_{H_{v,\gamma}}$ is shorthand for $\lambda i.v_{v_{H_{v,\gamma}}(i)}$, and $\tilde{v}$, $\bar{v}$ are defined as in Lemma 3.2.*

*Proof* Define $\omega_{v,\gamma} g := \gamma(g(0))(\tilde{v}_g)$ and then $H_{v,\gamma} := G_{\bar{v},\omega_{v,\gamma}}$. Then (32) follows directly from (31). □

We will now give a computational version of Lemma 3.2 as a whole:

**Lemma 9.2** *Suppose that H satisfies (32) and that $v$ and $\gamma$ satisfy $C(v, \gamma, H_{v,\gamma}(0), \tilde{v}_{H_{v,\gamma}})$, which analogously to before abbreviates*

$$\tilde{v}_{H_{v,\gamma}(0)} \lhd v_{H_{v,\gamma}(0)} \to \underbrace{(\exists i < j < \gamma(H_{v,\gamma}(0))(\tilde{v}_{H_{v,\gamma}}))(([v](H_{v,\gamma}(0)) * \tilde{v}_{H_{v,\gamma}})_i \preceq_* ([v](H_{v,\gamma}(0)) * \tilde{v}_{H_{v,\gamma}})_j)}_{\neg \bar{P}([v](H_{v,\gamma}(0)) * \tilde{v}_{H_{v,\gamma}}, \gamma(H_{v,\gamma}(0))(\tilde{v}_{H_{v,\gamma}}))}. \qquad (33)$$

*Then we have $\neg \bar{P}(v, H_{v,\gamma}(\gamma(H_{v,\gamma}(0))(\tilde{v}_{H_{v,\gamma}})) + 2)$ i.e.*

$$(\exists i < j < H_{v,\gamma}(\gamma(H_{v,\gamma}(0))(\tilde{v}_{H_{v,\gamma}})) + 2)(v_i \preceq_* v_j).$$

*Proof* This follows directly from Lemma 3.2. First of all, we define $g := H_{v,\gamma}$, then the sequence $w$ in Lemma 3.2 becomes identified with $[v](H_{v,\gamma}(0)) * \tilde{v}_{H_{v,\gamma}}$, and so setting $k := \gamma(H_{v,\gamma}(0))(\tilde{v}_{H_{v,\gamma}})$, the Eq. (33) is just (3), while (32) is just (4), and so by the lemma there exists some $i < j < g(k) + 2$ such that $v_i \preceq_* v_j$. But since $g(k) + 2 = H_{v,\gamma}(\gamma(H_{v,\gamma}(0))(\tilde{v}_{H_{v,\gamma}})) + 2$ we're done. □

What we have shown above is that if $C(v, \gamma, H_{v,\gamma}(0), \tilde{v}_{H_{v,\gamma}})$ then $\neg \bar{P}(v, H_{v,\gamma}(\gamma(H_{v,\gamma}(0))(\tilde{v}_{H_{v,\gamma}})) + 2)$, or in other words,

$$\bar{P}(v, H_{v,\gamma}(\gamma(H_{v,\gamma}(0))(\tilde{v}_{H_{v,\gamma}})) + 2) \wedge C(v, \gamma, H_{v,\gamma}(0), \tilde{v}_{H_{v,\gamma}})$$

must be false. But this is just the conclusion of the function interpretation of $\mathrm{ZL}_{\mathrm{lex}}$ for $Nv\gamma = H_{v,\gamma}(\gamma(H_{v,\gamma}(0))(\tilde{v}_{H_{v,\gamma}})) + 2$, $Mv\gamma = H_{v,\gamma}(0)$ and $Wv\gamma = \tilde{v}_{H_{v,\gamma}}$, and so for any $v, \gamma$ and $n$ satisfying (23) we must have $\neg \bar{P}(u, n)$, which is exactly what we want! Let's make this formal.

**Theorem 9.3** *Define $N, M : (X^*)^{\mathbb{N}} \to (\mathbb{N} \to (X^*)^{\mathbb{N}} \to \mathbb{N}) \to \mathbb{N}$ and $W : (X^*)^{\mathbb{N}} \to (\mathbb{N} \to (X^*)^{\mathbb{N}} \to \mathbb{N}) \to (X^*)^{\mathbb{N}}$ by*

$$N v\gamma := H_{v,\gamma}(\gamma(H_{v,\gamma}(0))(\tilde{v}_{H_{v,\gamma}})) + 2$$
$$M v\gamma := H_{v,\gamma}(0)$$
$$W v\gamma := \tilde{v}_{H_{v,\gamma}}$$

*where $H$ is some functional which satisfies (32). Define $P(s) :\equiv (\forall i < j < |s|)(s_i \npreceq_* s_j)$ so that $\bar{P}(u,k) \equiv (\forall i < j < k)(u_i \npreceq_* u_j)$, and let $n$, $v$ and $\gamma$ be such that they satisfy the functional interpretation of* $\mathrm{ZL}_{lex}$ *relative to $N$, $M$, $W$ defined above:*

$$\bar{P}(u,n) \to \bar{P}(v, N v\gamma) \wedge C(v, \gamma, M v\gamma, W v\gamma) \tag{34}$$

*Then we have $\neg\bar{P}(u,n)$ and hence*

$$(\forall u)(\exists i < j < n)(u_i \preceq_* u_j).$$

*Proof* As shown above, if $H$ satisfies (32) and $C(v, \gamma, M v\gamma, W v\gamma)$, then by Lemma 9.2 we have $\neg\bar{P}(v, N v\gamma)$, which implies $\neg(\bar{P}(v, N v\gamma) \wedge C(v, \gamma, M v\gamma, W v\gamma))$, and so by the contrapositive of (34) we have $\neg\bar{P}(u,n)$. □

Therefore any program which computes $v$, $\gamma$ and $n$ on any $u$, $N$, $M$ and $W$, in particular that of Theorem 8.3, can be converted to a program which realizes $\mathrm{WQO}(\preceq_*)$:

**Corollary 9.4** *Suppose that $H$ satisfies (32), and define $N v\gamma := H_{v,\gamma}(\gamma(H_{v,\gamma}(0))(\tilde{v}_{H_{v,\gamma}})) + 2$. Then provably in* CONT *the functional $\Phi : (X^*)^{\mathbb{N}} \to \mathbb{N}$ defined by*

$$\Phi(u) := \Psi^N(u)$$

*where $\Psi^N(u)$ is defined as in Theorem 8.3 witnesses* $\mathrm{WQO}(\preceq_*)$ *i.e.*

$$(\exists i < j < \Phi(u))(u_i \preceq_* u_j).$$

**Corollary 9.5** *Let $H$ be defined as in Lemma 9.1 for $G$ as defined in Sect. 5. Then $\Phi : (\mathbb{B}^*)^{\mathbb{N}} \to \mathbb{N}$ as defined in Corollary 9.4 witnesses* $\mathrm{WQO}(=_{\mathbb{B},*})$ *i.e.*

$$(\exists i < j < \Phi(u))(u_i =_{\mathbb{B},*} u_j).$$

*Remark 9.6* To construct our realizer for $\mathrm{WQO}(\preceq_*)$ we have only used the first component $n$ of the full functional interpretation $n, v, \gamma$ of $\mathrm{ZL}_{\mathrm{lex}}$. This makes sense: We actually prove via Lemma 9.2 that $(\exists v, \gamma)Q(v,\gamma) \to \bot$ where $Q(v,\gamma)$ abbreviates the conclusion of (34), and so to realize Higman's lemma we in fact only need to produce some functional $\Phi$ such that $\bar{P}(u, \Phi(u)) \to (\exists v, \gamma)Q(v,\gamma)$, and so the full computational interpretation of $\mathrm{ZL}_{\mathrm{lex}}$ via learning procedures was not strictly necessary. However, this simply emphasises the fact that we have achieved much more that a realizer for Higman's lemma—Theorem 8.3 allows us to extract a program from *any* proof which uses $\mathrm{ZL}_{\mathrm{lex}}$, and in general this program may well need a specific $v$ and $\gamma$ satisfying $Q(v,\gamma)$, even though here that was not the case. There is a further point to be made in this direction—namely that the concrete witnesses for $v$ and $\gamma$ enables us to *verify* our realizer $\Phi(u)$ in a quantifier-free theory, a fact that is relevant to those inclined towards foundational issues.

Before we conclude, it is worth pausing for a moment and trying to explain from an algorithmic point of view what the realizer we get in Corollary 9.4 actually does. Note that all of the following is essentially just an informal recapitulation of ideas contained in the preceding results. Roughly speaking, $\Phi(u)$ encodes a program which works by recursion on the lexicographic ordering $\rhd_{\text{lex}}$. First, it finds the point $m_0$ such that

$$N([u]_{m_0})([\gamma_u]_{m_0}) < m_0$$

where $\gamma_u := \lambda n, w.\ \Phi([u](n) * w)$ if $w_0 \lhd u_n$, and so in particularly it only looks at the sequence $u$ at points $n < m_0$. For simplicity let's define $u', \gamma' := [u]_{m_0}, [\gamma_u]_{m_0}$. Now, using any program $H$ which realizes $\mathrm{WQO}_{\mathrm{seq}}(\preceq)$ on the sequence $\bar{u}'$ we find a sufficiently large approximation $H_{u',\gamma'}$ to a constant subsequence, which works up to the point $\gamma'(H_{u',\gamma'}(0))(\tilde{u}'_{H_{u',\gamma'}})$.

Now if $\tilde{u}'_{H_{u',\gamma'}(0)} \lhd u'_{H_{u',\gamma'}(0)}$ then we must have $H_{u',\gamma'}(0) < m_0$ (using our assumption that $0_X$ is chosen to be minimal with respect to $\lhd$) and so $\gamma'(H_{u',\gamma'}(0))(\tilde{u}'_{H_{u',\gamma'}}) = \Phi([u'](H_{u',\gamma'}(0)) * \tilde{u}'_{H_{u',\gamma'}})$. Assuming inductively that this returns a bound for $[u'](H_{u',\gamma'}(0)) * \tilde{u}'_{H_{u',\gamma'}}$ being a good sequence, then using reasoning as in Lemma 3.2 this means that $u'$ becomes a good sequence before point $H_{u',\gamma'}(\Phi([u'](H_{u',\gamma'}(0)) * \tilde{u}'_{H_{u',\gamma'}})) + 2 = Nu'\gamma'$. But since $\Phi(u) = Nu'\gamma' < m_0$ and $u' = [u]_{m_0}$ then this means also that $u$ is good before $\Phi(u)$.

To verify that $\Phi([u'](H_{u',\gamma'}(0)) * \tilde{u}'_{H_{u',\gamma'}})$ returns a bound, we can repeat this argument for $u_1 := [u'](H_{u',\gamma'}(0)) * \tilde{u}'_{H_{u',\gamma'}}$, and we end up with a learning procedure as in Sect. 8. Eventually, this learning procedure will terminate with a minimal sequence $v$ such that $\Phi(v)$ is guaranteed to witnesses that $v$ is good.

## 10 Conclusion

I will conclude by tying up everything that we've done and outlining some directions for future work. On the route to Corollary 9.5 we took what I hope was a pleasant and instructive detour through many different areas which connect proof theory and well quasi-order theory, the most important of which I will now summarise.

Right at the start, in Chaps. 2 and 3, we discussed various nuances that arise when giving an axiomatic formalisation of results in WQO theory, in particular how the distinction between dependent choice or Zorn's lemma plays an important role in the context of program extraction. The full formalisation of Nash–Williams' minimal bad sequence argument has already been studied in e.g. [26] (in MINLOG) and [28] (in Isabelle/HOL), and we hope that our formal proof sketched in Chap. 3 may prove informative to those working in a more hands-on manner on the formalisation of WQO theory in proof assistants.

In Chaps. 4 and 5 we took the opportunity to present Gödel's functional interpretation in a way that would appeal to readers not already familiar with it. In Sect. 4.4 we placed particular emphasis on explaining how the interpretation behaves in *practice*, and in this vein we gave a carefully worked out case study in Sect. 5, which also formed a key Lemma in our proof of $\mathrm{WQO}(=_{\mathbb{B},*})$. It is my sincere hope that these chapters will be a general help to those interested in how proof interpretations work, independently of the rest of the paper.

Chapters 6 and 8 contain our main technical contribution, namely the solution of the functional interpretation of $\mathrm{ZL}_{\mathrm{lex}}$. While as a direct consequence this enables us to extract a program witnessing $\mathrm{WQO}(=_{\mathbb{B},*})$, our work in these chapters is much broader, and provides us with a method of giving a computational interpretation to *any* proof that can be formalised in $\mathrm{PA}^{\omega} + \text{QF-AC} + \mathrm{ZL}_{\mathrm{lex}}$, where moreover $\mathrm{ZL}_{\mathrm{lex}}$ can involve any relation $(X, \rhd)$ which is provably wellfounded. In particular, this paves the way for the extraction of programs from much more complex proofs in WQO theory, such as Kruskal's theorem, and we intend to address this in future work.

Contained in Chaps. 6 and 8 is also an extension of my work on *learning procedures* [23]. While in this paper they play the role of making our computational interpretation of $\mathrm{ZL}_{\mathrm{lex}}$ more intuitive, Lemma 8.2 is of interest in its own right, as it demonstrates that we can extend the notion of a learning procedure as introduced in [23] to the non-wellfounded ordering $\rhd_{\mathrm{lex}}$. We anticipate that a number of variants of open induction or Zorn's lemma over $\rhd_{\mathrm{lex}}$ could be given computational interpretations by appealing directly to Lemma 8.2 as an intermediate result, and this was part of our motivation for stating it explicitly here.

Sandwiched between these sections is Chap. 7, which itself forms a small essay on higher-type computability theory, and the various ways of carrying out recursion over $\rhd_{\mathrm{lex}}$ in the continuous functionals. There are a number of interesting questions to be answered in this direction. Firstly, what is the relationship between EORec, Berger's open recursion and the many variants of bar recursion which have been devised in the context of proof theory? I have already shown that Berger's open recursion is primitive recursively equivalent to *modified bar recursion* [5] and thus strictly stronger than Spector's original bar recursion, but I conjecture that in contrast, EORec is equivalent to Spector's bar recursion and thus weaker than Berger's open recursion. This would also imply that EORec exists in the type structure of $\mathcal{M}^{\omega}$ strongly majorizable functionals, and so does not necessarily rely on continuity to be a wellfounded form of recursion. It would be interesting to explore some of these issues in the future.

Finally, we should not forget that in this chapter we gave a new program which witnesses Higman's lemma, that works not just for the $=_{\mathbb{B}}$, but for *any* WQO for which a realizer of $\mathrm{WQO}_{\mathrm{seq}}(\preceq)$ can be given. Moreover, in contrast to [21], this realizer encodes a recursive algorithm which seems to do something fundamentally intuitive. The precise relationship between this algorithm and the many others which have been offered over the years is a question we leave open for now, although the presence of the control functional in the explicit open recursor leads me to conjecture that it is genuinely different to most of them. But for now we simply hope that, among other things, we have provided a little more insight into the computational meaning of Nash–Williams' elegant classical proof.

**Acknowledgements** In developing the ideas of this article I have benefited greatly from numerous illuminating conversations with Ulrich Berger, Paulo Oliva and Monika Seisenberger. Moreover, I am grateful to the anonymous referee, whose extremely detailed review led to a much better version of the paper.

## References

1. http://www.mathematik.uni-muenchen.de/~logik/minlog/. Official homepage of Minlog, as of January 2018.
2. Aschieri, F., & Berardi, S. (2010). Interactive learning-based realizability for Heyting arithmetic with EM1. *Logical Methods in Computer Science*, *6*(3),
3. Avigad, J., & Feferman, S. (1998). Gödel's functional ("Dialectica") interpretation. In S. R. Buss (Ed.), *Handbook of proof theory* (Vol. 137, pp. 337–405)., Studies in logic and the foundations of mathematics Amsterdam: North Holland.
4. Berger, U. (2004). A computational interpretation of open induction. In *Proceedings of LICS 2004*, pp. 326–334. IEEE Computer Society.
5. Berger, U., & Oliva, P. (2005). Modified bar recursion and classical dependent choice. *Lecture Notes in Logic*, *20*, 89–107.
6. Ershov, Y. L. (1977). Model *C* of partial continuous functionals. *Logic colloquium* (pp. 455–467). Amsterdam: North Holland.
7. Escardó, M., & Oliva, P. (2010). Selection functions, bar recursion and backward induction. *Mathematical Structures in Computer Science*, *20*(2), 127–168.
8. Escardó, M., & Oliva, P. (2011). Sequential games and optimal strategies. *Royal Society Proceedings A*, *467*, 1519–1545.
9. Escardó, M., & Oliva, P. (2012). Computing Nash equilibria of unbounded games. In *Proceedings of the Turing Centenary Conference, Manchester*, vol. 10 of EPiC Series, pp. 53–65.
10. Gödel, K. (1958). Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *dialectica*, *12*, 280–287.
11. Higman, G. (1952). Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society*, *2*, 326–336.
12. Kleene, S. C. (1959). Countable functionals. In A. Heyting (Ed.), *Constructivity in mathematics* (pp. 81–100). Amsterdam: North-Holland.
13. Kohlenbach, U. (2008). *Applied proof theory: Proof interpretations and their use in mathematics*. Monographs in Mathematics. Springer.
14. Kreisel, G. (1951). On the interpretation of non-finitist proofs. *Part I Journal of Symbolic Logic*, *16*, 241–267.
15. Kreisel, G. (1952). On the interpretation of non-finitist proofs, Part II: Interpretation of number theory. *Journal of Symbolic Logic*, *17*, 43–58.
16. Kreisel, G. (1959). Interpretation of analysis by means of functionals of finite type. In A. Heyting (Ed.), *Constructivity in mathematics* (pp. 101–128). Amsterdam: North-Holland.
17. Longley, J., & Normann, D. (2015). *Higher-order computability*. Theory and Applications of Computability: Springer.
18. Murthy, C. R. (1990). *Extracting constructive content from classical proofs*. Ph.D. thesis, Ithaca, New York.

19. Nash-Williams, C. St, & J. A., (1963). On well-quasi-ordering finite trees. *Proceedings of the Cambridge Philosophical Society*, *59*, 833–835.
20. Oliva, P. (2006). Understanding and using Spector's bar recursive interpretation of classical analysis. In A. Beckmann, U. Berger, B. Löwe, and J. V. Tucker, editors, *Proceedings of CiE'2006*, volume 3988 of *LNCS*, pp. 423–234.
21. Powell, T. (2012). Applying Gödel's Dialectica interpretation to obtain a constructive proof of Higman's lemma. In *Proceedings of Classical Logic and Computation '12*, volume 97 of *EPTCS*, pp. 49–62, 2012.
22. Powell, T. (2014). The equivalence of bar recursion and open recursion. *Annals of Pure and Applied Logic*, *165*(11), 1727–1754.
23. Powell, T. (2016). Gödel's functional interpretation and the concept of learning. In *Proceedings of Logic in Computer Science (LICS 2016)*, pp. 136–145. IEEE Computer Society.
24. Raoult, J.-C. (1988). Proving open properties by induction. *Information Processing Letters*, *29*, 19–23.
25. Scott, D. S. (1970). Outline of a mathematical theory of computation. In *4th Annual Princeton Conference on Information Sciences and Systems*, pp. 169–176.
26. Seisenberger, M. (2003). *On the constructive content of proofs*. Ph.D. thesis, Ludwig Maximilians Universität München.
27. Spector, C. (1962). Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles in current intuitionistic mathematics. In F. D. E. Dekker (Ed.), *Recursive Function Theory: Proc* (pp. 1–27), Symposia in Pure Mathematics, volume 5 Providence, Rhode Island: American Mathematical Society.
28. Sternagel, C. (2013). Certified Kruskal's tree theorem. In *Proceedings of the 3rd International Conference on Certified Programs and Proofs (CPP '13)*, volume 8307 of *LNCS*, pp. 178–193.
29. Troelstra, A. S. (1973). *Metamathematical investigation of intuitionistic arithmetic and analysis* (Vol. 344), Lecture Notes in Mathematics Berlin: Springer.

# Well-Quasi Orders and Hierarchy Theory

**Victor Selivanov**

**Abstract** We discuss some applications of WQOs to several fields were hierarchies and reducibilities are the principal classification tools, notably to Descriptive Set Theory, Computability theory and Automata Theory. While the classical hierarchies of sets usually degenerate to structures very close to ordinals, the extension of them to functions requires more complicated WQOs, and the same applies to reducibilities. We survey some results obtained so far and discuss open problems and possible research directions.

**Keywords** Well quasiorder · Better quasiorder · Quasi-polish space · Borel hierarchy · Hausdorff hierarchy · Wadge hierarchy · Fine hierarchy · Reducibility · $k$-Partition · Labeled tree · $h$-Quasiorder

## 1 Introduction

WQO-theory is an important part of combinatorics with deep connections and applications to several parts of mathematics (proof theory, reverse mathematics, descriptive set theory, graph theory) and computer science (verification of infinite-state systems, combinatorics on words and formal languages, automata theory).

In this paper we discuss some applications of WQOs to several fields where hierarchies and reducibilities are the principal classification tools, notably to descriptive set theory (DST), computability theory and automata theory. The starting point of our discussion are three important parts of DST:

(1) The classical Borel, Luzin, and Hausdorff hierarchies in Polish spaces, which are defined using set-theoretic operations.

Supported by Russian Science Foundation, project 18-11-00100.

V. Selivanov (✉)
A.P. Ershov Institute of Informatics Systems SB RAS and Kazan Federal University,
Lavrentyev av. 6, Novosibirsk 630090, Russia
e-mail: vseliv@ngs.ru

 
P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_10

(2) The Wadge hierarchy which is non-classical in the sense that it is based on a notion of reducibility that was not recognized in the classical DST, and on using ingenious versions of Gale-Stewart games rather than the properties of set-theoretic operations.
(3) The classification of Borel equivalence relations by means of Borel reducibility, which uses deep analytical tools.

Addison [2–5] suggested to develop a general hierarchy theory, in order to have precise notions and tools to study analogies between the classical hierarchies and some hierarchies that appeared later in logic and computability theory. In particular, he suggested a general notion of a hierarchy of sets. The hierarchy theory was continued in a series of the author's papers (see e.g. [88, 95, 98] and references therein) where new general notions and techniques of hierarchy theory were suggested, in particular the notion of a hierarchy of $k$-partitions and of a reducibility that fits a given hierarchy.

While the classical hierarchies of sets usually degenerate to structures very close to ordinals, the attempt to extend them to $k$-partitions requires more complicated WQOs (namely, the so called $h$-quasiorder on labeled forests), and the same applies to reducibilities. This was the original author's motivation for a systematic study of relationships between WQO-theory and hierarchy theory.

In this paper we survey some results obtained in this direction so far. The general theory of hierarchies and reducibilities based on WQO-theory seems already matured and homogeneous, including the extension to $k$-partitions. In contrast, several attempts to include in the theory more general functions and Borel equivalence relations is still in the beginning, and the role of WQOs in such further generalizations is not yet clear. For this reason we mention several open questions which seem interesting for such generalizations. We decided not to include proofs (which are sometimes technical and long), instead concentrating on the formulations of basic results and discussions of the main tools.

In the next section we recall some well-known notions and facts, but we also mention some less-known facts about the Wadge hierarchy and the extension of the classical hierarchies to the so called quasi-Polish spaces which are of interest to computer science. In Sect. 3 we recall the basic notions of WQO and BQO and provide examples which are important for the sequel. In Sect. 4 we discuss several versions and extensions of Wadge reducibility which are based on the $h$-quasiorder. In Sect. 5 we discuss some reducibilities on objects more complex than $k$-partitions, notably for equivalence relations and functions on the Baire space. In Sect. 5 we recall some notions of the general hierarchy theory needed to unify terminology. In Sects. 7 and 8 we discuss some hierarchies in computability theory and automata theory respectively, trying to relate this to WQO-theory and the $h$-quasiorder. We conclude in Sect. 9 with comments on a recent preprint and some open questions.

As is well known, outside the Borel sets in DST or the hyperarithmetical sets in computability theory, some properties of hierarchies and reducibilities depend on set-theoretic axioms. Although the axiomatic issues are important and interesting, we decided to avoid the foundational discussions and to include only results provable in the widely accepted axiomatic system ZFC. As a result, we stay mainly within the Borel sets, although many facts may be extended far beyond the Borel sets under suitable set-theoretic assumptions.

## 2 Preliminaries

In this section we recall some notation, notions and results used in the subsequent sections. We use the standard set-theoretic notation like $dom(f)$ and $rng(f)$ for the domain and range of a function $f$, respectively, $X \times Y$ for the Cartesian product, $X \oplus Y$ for the disjoint union of sets $X$ and $Y$, $Y^X$ for the set of functions $f\colon X \to Y$, and $P(X)$ for the set of all subsets of $X$. For $A \subseteq X$, $\overline{A}$ denotes the complement $X \setminus A$ of $A$ in $X$. The notation $f : X \to Y$ means that $f$ is a (total) function from a set $X$ to a set $Y$.

### *2.1 Ordinals*

We assume the reader to be acquainted with the notion of an ordinal (see e.g. [62]). Ordinals are important for the hierarchy theory because levels of hierarchies of sets are (almost) well ordered by inclusion. This opens the possibility to estimate the complexity of sets (and other objects) by ordinals.

Ordinals are denoted by $\alpha$, $\beta$, $\gamma$, .... The successor $\alpha + 1$ of an ordinal $\alpha$ is defined by $\alpha + 1 = \alpha \cup \{\alpha\}$. Every ordinal $\alpha$ is the set of all smaller ordinals, in particular $k = \{0, 1, \dots, k-1\}$ for each $k < \omega$, and $\omega = \{0, 1, 2, \dots\}$. Ordinals may be considered as the order types of well orders (see the next subsection).

We use some well-known facts about the ordinal arithmetic. As usual, $\alpha + \beta$, $\alpha \cdot \beta$ and $\alpha^\beta$ denote the ordinal addition, multiplication and exponentiation of $\alpha$ and $\beta$, respectively. The context will help to distinguish the ordinal exponentiation from the set exponentiation denoted in the same way but having a quite different meaning.

Below we will mention the ordinals $\omega, \omega^2, \omega^3, \dots$ and $\omega^\omega$. The last ordinal is the order type of finite sequences $(k_1, \dots, k_n)$ of natural numbers $k_1 \geq \dots \geq k_n$, ordered lexicographically. Any non-zero ordinal $\alpha < \omega^\omega$ is uniquely representable in the form $\alpha = \omega^{k_1} + \dots + \omega^{k_n}$ with $\omega > k_1 \geq \dots \geq k_n$. We will also use the bigger ordinal $\varepsilon_0 = sup\{\omega, \omega^\omega, \omega^{(\omega^\omega)}, \dots\}$. Any non-zero ordinal $\alpha < \varepsilon_0$ is uniquely representable in the form $\alpha = \omega^{\gamma_0} + \dots + \omega^{\gamma_k}$ for a finite sequence $\gamma_0 \geq \dots \geq \gamma_k$ of ordinals $< \alpha$. The ordinal $\varepsilon_0$ is the smallest solution of the ordinal equation $\omega^\varkappa = \varkappa$.

All concrete ordinals mentioned above are computable, i.e. they are order types of computable well orders on computable subsets of $\omega$. The first non-computable ordinal $\omega_1^{CK}$, known as the Church-Kleene ordinal, is important in computability theory. The first non-countable ordinal $\omega_1$ is important for the hierarchy theory. From this ordinal one can construct many other interesting ordinals, in particular $\omega_1^{\omega_1}, \omega_1^{(\omega_1^{\omega_1})}, \dots$. Even much bigger ordinals (like the Wadge ordinal discussed below) are of interest for the hierarchy theory.

### 2.2 Partial Orders and Quasiorders

We use some standard notation and terminology on partially ordered sets (posets), which may be found e.g. in [21]. Recall that a *quasiorder* (QO) is a structure $(P; \leq)$ satisfying the axioms of reflexivity $\forall x(x \leq x)$ and transitivity $\forall x \forall y \forall z(x \leq y \wedge y \leq z \rightarrow x \leq z)$. *Poset* is a QO satisfying the antisymmetry axiom $\forall x \forall y(x \leq y \wedge y \leq x \rightarrow x = y)$. *Linear order* is a partial order satisfying the connectivity axiom $\forall x \forall y(x \leq y \vee y \leq x)$. A linearly ordered subset of a poset is sometimes called a *chain*.

Any partial order $\leq$ on $P$ induces the relation of strict order $<$ on $P$ defined by $a < b \leftrightarrow a \leq b \wedge a \neq b$ and called the strict order related to $\leq$. The relation $\leq$ can be restored from $<$, so we may safely apply the terminology on partial orders also to the strict orders. A poset $(P; \leq)$ will be often shorter denoted just by $P$. Any subset of a poset $P$ may be considered as a poset with the induced partial order.

It is well known that any QO $(P; \leq)$ induces the partial order $(P^*; \leq^*)$ called *the quotient* of $P$. The set $P^*$ is the quotient set of $P$ under the equivalence relation defined by $a \equiv b \leftrightarrow a \leq b \wedge b \leq a$; the set $P$ consists of all equivalence classes $[a] = \{x \mid x \equiv a\}$, $a \in P$. The partial order $\leq^*$ is defined by $[a] \leq^* [b] \leftrightarrow a \equiv b$. We will not be cautious when applying notions about posets also to QOs; in such cases we mean the corresponding quotient-poset of the QO.

A partial order $(P; \leq)$ is *well-founded* if it has no infinite descending chains. In this case there are a unique ordinal $rk(P)$ and a unique rank function $rk_P$ from $P$ onto $rk(P)$ satisfying $a < b \rightarrow rk_P(a) < rk_P(b)$. It is defined by induction $rk_P(x) = sup\{rk_P(y) + 1 \mid y < x\}$. The ordinal $rk(P)$ is called the *rank* (or *height*) of $P$, and the ordinal $rk_P(x)$ is called the *rank of the element* $x \in P$ *in* $P$.

In the sequel we will often deal with semilattices expanded by some additional operations. In particular the following notions introduced in [84, 93] will often be mentioned. The abbreviation "dc-semilattice" refers to "semilattice with discrete closures".

**Definition 2.1** By *dc-semilattice* we mean a structure $(S; \leq, \cup, p_0, \ldots, p_{k-1})$ such that:

(1) $(S; \cup)$ is an upper semilattice, i.e. it satisfies $(x \cup y) \cup z = x \cup (y \cup z)$, $x \cup y = y \cup x$ and $x \cup x = x$, for all $x, y, z \in S$.
(2) $\leq$ is the partial order on $S$ induced by $\cup$, i.e. $x \leq y$ iff $x \cup y = y$, for all $x, y \in S$.
(3) Every $p_i$, $i < k$, is a closure operation on $(S; \leq)$, i.e. it satisfies $x \leq p_i(x)$, $x \leq y \rightarrow p_i(x) \leq p_i(y)$ and $p_i(p_i(x)) \leq p_i(x)$, for all $x, y \in S$.

(4) The operations $p_i$ have the following discreteness property: for all distinct $i, j < k$, $p_i(x) \leq p_j(y) \to p_i(x) \leq y$, for all $x, y \in S$.
(5) Every $p_i(x)$ is join-irreducible, i.e. $p_i(x) \leq y \cup z \to (p_i(x) \leq y \vee p_i(x) \leq z)$, for all $x, y, z \in S$.

By *dcσ-semilattice* we mean a *dc*-semilattice which is also a $\sigma$-semilattice (i.e., the supremums of countably many elements exist), and the axiom (5) holds also for the supremums of countable subsets of $S$, (i.e., $p_i(x) \leq \bigcup_{j<\omega} y_j$ implies that $p_i(x) \leq y_j$ for some $j < \omega$; we express this by saying that $p_i(x)$ is $\sigma$-join-irreducible).

## 2.3 Topological Spaces

Here we recall some topological notions and facts relevant to this paper. We assume the reader to be familiar with the basic notions of topology [25]. For the underlying set of a topological space $X$ we will write $X$, in abuse of notation. We will often abbreviate "topological space" to "space". A space is *zero-dimensional* if it has a basis of clopen sets. Recall that a *basis* for the topology on $X$ is a set $\mathcal{B}$ of open subsets of $X$ such that for every $x \in X$ and open $U$ containing $x$ there is $B \in \mathcal{B}$ satisfying $x \in B \subseteq U$.

Let $\omega$ be the space of non-negative integers with the discrete topology. Of course, the spaces $\omega \times \omega = \omega^2$, and $\omega \sqcup \omega$ are homeomorphic to $\omega$, the first homeomorphism is realized by the Cantor pairing function $\langle \cdot, \cdot \rangle$. Let $\mathcal{N} = \omega^\omega$ be the set of all infinite sequences of natural numbers (i.e., of all functions $\xi \colon \omega \to \omega$). Let $\omega^*$ be the set of finite sequences of elements of $\omega$, including the empty sequence. For $\sigma \in \omega^*$ and $\xi \in \mathcal{N}$, we write $\sigma \sqsubseteq \xi$ to denote that $\sigma$ is an initial segment of the sequence $\xi$. By $\sigma\xi = \sigma \cdot \xi$ we denote the concatenation of $\sigma$ and $\xi$, and by $\sigma \cdot \mathcal{N}$ the set of all extensions of $\sigma$ in $\mathcal{N}$. For $x \in \mathcal{N}$, we can write $x = x(0)x(1)\ldots$ where $x(i) \in \omega$ for each $i < \omega$. For $x \in \mathcal{N}$ and $n < \omega$, let $x \restriction n = x(0)\ldots x(n-1)$ denote the initial segment of $x$ of length $n$. Notations in the style of regular expressions like $0^\omega$, $0^*1$ or $0^m1^n$ have the obvious standard meaning.

By endowing $\mathcal{N}$ with the product of the discrete topologies on $\omega$, we obtain the so-called *Baire space*. The product topology coincides with the topology generated by the collection of sets of the form $\sigma \cdot \mathcal{N}$ for $\sigma \in \omega^*$. The Baire space is of primary importance for Descriptive Set Theory and Computable Analysis. The importance stems from the fact that many countable objects are coded straightforwardly by elements of $\mathcal{N}$, and it has very specific topological properties. In particular, it is a perfect zero-dimensional space such that any countably based zero-dimensional $T_0$-space topologically embeds into it. The subspace $\mathcal{C} := 2^\omega$ of $\mathcal{N}$ formed by the infinite binary strings (endowed with the relative topology inherited from $\mathcal{N}$) is known as the *Cantor space*.

We recall the well-known (see e.g. [65]) relation of closed subsets of $\mathcal{N}$ to trees. A *tree* is a non-empty set $T \subseteq \omega^*$ which is closed downwards under $\sqsubseteq$. A *leaf* of $T$ is a maximal element of $(T; \sqsubseteq)$. A *pruned tree* is a tree without leafs. A *path* *through* a tree $T$ is an element $x \in \mathcal{N}$ such that $x \upharpoonright n \in T$ for each $n \in \omega$. For any tree $T$, the set $[T]$ of paths through $T$ is closed in $\mathcal{N}$. For any non-empty closed set $A \subseteq \mathcal{N}$ there is a unique pruned tree $T$ with $A = [T]$ and, moreover, there is a continuous surjection $t : \mathcal{N} \to A$ which is constant on $A$ (such a surjection is called a retraction onto $A$). Therefore, there is a bijection between the pruned trees and the non-empty closed sets. Note that the well founded trees $T$ (i.e., trees with $[T] = \emptyset$) and non-empty well founded forests of the form $F := T \setminus \{\varepsilon\}$ will be used below, in particular in defining the $h$-quasiorders in Sects. 3.1 and 3.2.

The *Sierpinski space* $\mathbb{S}$ is the two-point set $\{\bot, \top\}$ where the set $\{\top\}$ is open but not closed. The space $P\omega$ is formed by the set of subsets of $\omega$ equipped with the Scott topology [6]. A countable base of the Scott topology is formed by the sets $\{A \subseteq \omega \mid F \subseteq A\}$, where $F$ ranges over the finite subsets of $\omega$.

Recall that a space $X$ is *Polish* if it is countably based and metrizable with a metric $d$ such that $(X, d)$ is a complete metric space. Important examples of Polish spaces are $\omega$, $\mathcal{N}$, $\mathcal{C}$, the space of reals $\mathbb{R}$ and its Cartesian powers $\mathbb{R}^n$ ($n < \omega$), the closed unit interval $[0, 1]$, the Hilbert cube $[0, 1]^\omega$ and the space $\mathbb{R}^\omega$. Simple examples of non-Polish spaces are $\mathbb{S}$, $P\omega$ and the space $\mathbb{Q}$ of rationals.

A *quasi-metric* on $X$ is a function from $X \times X$ to the nonnegative reals such that $d(x, y) = d(y, x) = 0$ iff $x = y$, and $d(x, y) \leq d(x, z) + d(z, y)$. Every quasi-metric on $X$ induces the topology $\tau_d$ on $X$ generated by the open balls $\{y \in X \mid d(x, y) < \varepsilon\}$ for $x \in X$ and $0 < \varepsilon$. A space $X$ is *quasi-metrizable* if there is a quasi-metric on $X$ which generates its topology. If $d$ is a quasi-metric on $X$, let $\hat{d}$ be the metric on $X$ defined by $\hat{d}(x, y) = max\{d(x, y), d(y, x)\}$. A sequence $\{x_n\}$ is *d-Cauchy* if for every $\varepsilon > 0$ there is $p \in \omega$ such that $d(x_n, x_m) < \varepsilon$ for all $p \leq n \leq m$. We say that the quasi-metric $d$ on $X$ is *complete* if every $d$-Cauchy sequence converges with respect to $\hat{d}$. A $T_0$ space $X$ is called *quasi-Polish* [17] if it is countably based and there is a complete quasi-metric which generates its topology.

Note that the spaces $\mathbb{S}$, $P\omega$ are quasi-Polish while the space $\mathbb{Q}$ is not. A complete quasi-metric which is compatible with the topology of $P\omega$ is given by $d(x, y) = 0$ if $x \subseteq y$ and $d(x, y) = 2^{-(n+1)}$ otherwise, where $n$ is the smallest element in $x \setminus y$ (for every, $x, y \subseteq \omega$). As shown in [17], a space is quasi-Polish iff it is homeomorphic to a $\Pi^0_2$-subset of [17] (the well known definition of $\Pi^0_2$-subset is recalled in the next subsection). There are some other interesting characterizations of quasi-Polish spaces. For this paper the following characterization in terms of total admissible representations is relevant.

A *representation* of a space $X$ is a surjection of a subspace of the Baire space $\mathcal{N}$ onto $X$. A basic notion of Computable Analysis [119] is the notion of admissible representation. A representation $\delta$ of $X$ is *admissible*, if it is continuous and any continuous function $\nu : Z \to X$ from a subset $Z \subseteq \mathcal{N}$ to $X$ is continuously reducible to $\delta$, i.e. $\nu = \delta \circ g$ for some continuous function $g : Z \to \mathcal{N}$. In [17] the following characterization of quasi-Polish spaces was obtained: A space $X$ is quasi-Polish iff it has a total admissible representation $\delta : \mathcal{N} \to X$.

### 2.4 Classical Hierarchies in Quasi-Polish Spaces

Here we recall some notions and facts on the classical hierarchies in quasi-Polish spaces [17, 65]. Note that the definitions of Borel and Luzin hierarchies look slightly different from the well-known definitions for Polish spaces [65], in order to behave correctly also on non-Hausdorff spaces (for Polish spaces the definitions are equivalent to the usual ones).

A *pointclass* on $X$ is simply a collection $\mathbf{\Gamma}(X)$ of subsets of $X$. A *family of pointclasses* [99] is a family $\mathbf{\Gamma} = \{\mathbf{\Gamma}(X)\}$ indexed by arbitrary topological spaces $X$ (or by spaces in a reasonable class) such that each $\mathbf{\Gamma}(X)$ is a pointclass on $X$ and $\mathbf{\Gamma}$ is closed under continuous preimages, i.e. $f^{-1}(A) \in \mathbf{\Gamma}(X)$ for every $A \in \mathbf{\Gamma}(Y)$ and every continuous function $f\colon X \to Y$. A basic example of a family of pointclasses is given by the family $\mathcal{O} = \{\tau_X\}$ of the topologies of all the spaces $X$.

We will use some operations on families of pointclasses. First, the usual set-theoretic operations will be applied to the families of pointclasses pointwise: for example, the union $\bigcup_i \mathbf{\Gamma}_i$ of the families of pointclasses $\mathbf{\Gamma}_0, \mathbf{\Gamma}_1, \dots$ is defined by $(\bigcup_i \mathbf{\Gamma}_i)(X) = \bigcup_i \mathbf{\Gamma}_i(X)$.

Second, a large class of such operations is induced by the set-theoretic operations of Kantorovich and Livenson (see e.g. [99] for the general definition). Among them are the operation $\mathbf{\Gamma} \mapsto \mathbf{\Gamma}_\sigma$, where $\mathbf{\Gamma}(X)_\sigma$ is the set of all countable unions of sets in $\mathbf{\Gamma}(X)$, the operation $\mathbf{\Gamma} \mapsto \mathbf{\Gamma}_\delta$, where $\mathbf{\Gamma}(X)_\delta$ is the set of all countable intersections of sets in $\mathbf{\Gamma}(X)$, the operation $\mathbf{\Gamma} \mapsto \mathbf{\Gamma}_c$, where $\mathbf{\Gamma}(X)_c$ is the set of all complements of sets in $\mathbf{\Gamma}(X)$, the operation $\mathbf{\Gamma} \mapsto \mathbf{\Gamma}_d$, where $\mathbf{\Gamma}(X)_d$ is the set of all differences of sets in $\mathbf{\Gamma}(X)$, the operation $\mathbf{\Gamma} \mapsto \mathbf{\Gamma}_\exists$ defined by $\mathbf{\Gamma}_\exists(X) := \{\exists^{\mathcal{N}}(A) \mid A \in \mathbf{\Gamma}(\mathcal{N} \times X)\}$, where $\exists^{\mathcal{N}}(A) := \{x \in X \mid \exists p \in \mathcal{N}.(p, x) \in A\}$ is the projection of $A \subseteq \mathcal{N} \times X$ along the axis $\mathcal{N}$, and finally the operation $\mathbf{\Gamma} \mapsto \mathbf{\Gamma}_\forall$ defined by $\mathbf{\Gamma}_\forall(X) := \{\forall^{\mathcal{N}}(A) \mid A \in \mathbf{\Gamma}(\mathcal{N} \times X)\}$, where $\forall^{\mathcal{N}}(A) := \{x \in X \mid \forall p \in \mathcal{N}.(p, x) \in A\}$.

The operations on families of pointclasses enable to provide short uniform descriptions of the classical hierarchies in quasi-Polish spaces. E.g., the *Borel hierarchy* is the sequence of families of pointclasses $\{\mathbf{\Sigma}^0_\alpha\}_{\alpha<\omega_1}$ defined by induction on $\alpha$ as follows [17, 92]: $\mathbf{\Sigma}^0_0(X) := \{\emptyset\}$, $\mathbf{\Sigma}^0_1 := \mathcal{O}$ (the family of open sets), $\mathbf{\Sigma}^0_2 := (\mathbf{\Sigma}^0_1)_{d\sigma}$, and $\mathbf{\Sigma}^0_\alpha(X) := (\bigcup_{\beta<\alpha} \mathbf{\Sigma}^0_\beta(X))_{c\sigma}$ for $\alpha > 2$. The sequence $\{\mathbf{\Sigma}^0_\alpha(X)\}_{\alpha<\omega_1}$ is called *the Borel hierarchy* in $X$. We also let $\mathbf{\Pi}^0_\beta(X) := (\mathbf{\Sigma}^0_\beta(X))_c$ and $\mathbf{\Delta}^0_\alpha(X) := \mathbf{\Sigma}^0_\alpha(X) \cap \mathbf{\Pi}^0_\alpha(X)$. The classes $\mathbf{\Sigma}^0_\alpha(X)$, $\mathbf{\Pi}^0_\alpha(X)$, $\mathbf{\Delta}^0_\alpha(X)$ are called the *levels* of the Borel hierarchy in $X$. The class $\mathbf{B}(X)$ of *Borel sets* in $X$ is defined as the union of all levels of the Borel hierarchy in $X$; it coincides with the smallest $\sigma$-algebra of subsets of $X$ containing the open sets. We have $\mathbf{\Sigma}^0_\alpha(X) \cup \mathbf{\Pi}^0_\alpha(X) \subseteq \mathbf{\Delta}^0_\beta(X)$ for all $\alpha < \beta < \omega_1$. For any uncountable quasi-Polish space $X$ and any $\alpha < \omega_1$, $\mathbf{\Sigma}^0_\alpha(X) \nsubseteq \mathbf{\Pi}^0_\alpha(X)$.

The *hyperprojective hierarchy* is the sequence of families of pointclasses $\{\mathbf{\Sigma}^1_\alpha\}_{\alpha<\omega_1}$ defined by induction on $\alpha$ as follows: $\mathbf{\Sigma}^1_0 = \mathbf{\Sigma}^0_2$, $\mathbf{\Sigma}^1_{\alpha+1} = (\mathbf{\Sigma}^1_\alpha)_{c\exists}$, $\mathbf{\Sigma}^1_\lambda = (\mathbf{\Sigma}^1_{<\lambda})_{\delta\exists}$, where $\alpha, \lambda < \omega_1$, $\lambda$ is a limit ordinal, and $\mathbf{\Sigma}^1_{<\lambda}(X) := \bigcup_{\alpha<\lambda} \mathbf{\Sigma}^1_\alpha(X)$. In this way, we obtain for any quasi-Polish space $X$ the sequence $\{\mathbf{\Sigma}^1_\alpha(X)\}_{\alpha<\omega_1}$, which we call here *the hyperprojective hierarchy in* $X$. The pointclasses $\mathbf{\Sigma}^1_\alpha(X)$, $\Pi^1_\alpha(X) := (\mathbf{\Sigma}^1_\alpha(X))_c$ and $\mathbf{\Delta}^1_\alpha(X) := \mathbf{\Sigma}^1_\alpha(X) \cap \Pi^1_\alpha(X)$ are called *levels of the hyperprojective hierarchy in* $X$. The finite non-zero levels of the hyperprojective hierarchy coincide with the corresponding levels of the Luzin projective hierarchy. The class of *hyperprojective sets* in $X$ is defined as the union of all levels of the hyperprojective hierarchy in $X$. We have $\mathbf{\Sigma}^1_\alpha(X) \cup \Pi^1_\alpha(X) \subseteq \mathbf{\Delta}^1_\beta(X)$ for all $\alpha < \beta < \omega_1$. For any uncountable Polish space $X$ and any $\alpha < \omega_1$, $\mathbf{\Sigma}^1_\alpha(X) \nsubseteq \Pi^1_\alpha(X)$. For any quasi-Polish space $X$, $\mathbf{B}(X) = \mathbf{\Delta}^1_1(X)$ (the Suslin theorem [17, 65]). As mentioned in the Introduction, in this paper we mostly stay within the Borel sets, hence the very important Luzin hierarchy will not be considered. We recalled its definition mainly to illustrate the general notions of hierarchy theory.

For any non-zero ordinal $\theta < \omega_1$, let $\{\mathbf{\Sigma}^{-1,\theta}_\alpha\}_{\alpha<\omega_1}$ be the Hausdorff difference hierarchy over $\mathbf{\Sigma}^0_\theta$. We recall the definition. An ordinal $\alpha$ is *even* (resp. *odd*) if $\alpha = \lambda + n$ where $\lambda$ is either zero or a limit ordinal and $n < \omega$, and the number $n$ is even (resp., odd). For an ordinal $\alpha$, let $r(\alpha) = 0$ if $\alpha$ is even and $r(\alpha) = 1$, otherwise. For any ordinal $\alpha$, define the operation $D_\alpha$ sending sequences of sets $\{A_\beta\}_{\beta<\alpha}$ to sets by

$$D_\alpha(\{A_\beta\}_{\beta<\alpha}) = \bigcup\{A_\beta \setminus \bigcup_{\gamma<\beta} A_\gamma \mid \beta < \alpha,\ r(\beta) \neq r(\alpha)\}.$$

For any ordinal $\alpha < \omega_1$ and any pointclass $\mathcal{E}$ in $X$, let $D_\alpha(\mathcal{E})$ be the class of all sets $D_\alpha(\{A_\beta\}_{\beta<\alpha})$, where $A_\beta \in \mathcal{E}$ for all $\beta < \alpha$. Finally, let $\mathbf{\Sigma}^{-1,\theta}_\alpha(X) = D_\alpha(\mathbf{\Sigma}^0_\theta(X))$ for any space $X$ and for all $\alpha, \theta < \omega$, $\theta > 0$. It is well known that $\mathbf{\Sigma}^{-1,\theta}_\alpha(X) \cup \Pi^{-1,\theta}_\alpha(X) \subseteq \mathbf{\Delta}^{-1,\theta}_\beta(X)$ and $\bigcup_{\alpha<\omega_1} \mathbf{\Sigma}^{-1,\theta}_\alpha(X) \subseteq \mathbf{\Delta}^0_{\theta+1}(X)$ for all $\alpha < \beta < \omega_1$. For any quasi-Polish space $X$ and any $0 < \theta < \omega_1$, $\bigcup_{\alpha<\omega_1} \mathbf{\Sigma}^{-1,\theta}_\alpha(X) = \mathbf{\Delta}^0_{\theta+1}(X)$ (the Hausdorff–Kuratowski theorem [17, 65]).

## 2.5 Wadge Hierarchy

Here we briefly discuss the Wadge reducibility in the Baire space. For subsets $A, B$ of the Baire space $\mathcal{N}$, $A$ is *Wadge reducible* to $B$ ($A \leq_W B$), if $A = f^{-1}(B)$ for some continuous function $f$ on $\mathcal{N}$. The quotient-poset of the QO $(P(\mathcal{N}); \leq_W)$ under the induced equivalence relation $\equiv_W$ on the power-set of $\mathcal{N}$ is called *the structure of Wadge degrees* in $\mathcal{N}$.

In [116] Wadge (using the Martin determinacy theorem) proved the following result: The structure $(\mathbf{B}(\mathcal{N}); \leq_W)$ of the Borel sets in the Baire space is semi-well-ordered (i.e., it is well-founded and for all $A, B \in \mathbf{B}(\mathcal{N})$ we have $A \leq_W B$ or $\overline{B} \leq_W A$). In particular, there is no antichain of size 3 in $(\mathbf{B}(\mathcal{N}); \leq_W)$. He has also computed the rank $\nu$ of $(\mathbf{B}(\mathcal{N}); \leq_W)$ which we call the Wadge ordinal. Recall that a set $A$ is *self-dual* if $A \leq_W \overline{A}$. Wadge has shown that if a Borel set is self-dual (resp. non-self-dual) then any Borel set of the next Wadge rank is non-self-dual (resp. self-dual), a Borel set of Wadge rank of countable cofinality is self-dual, and a Borel set of Wadge rank of uncountable cofinality is non-self-dual. This characterizes the structure of Wadge degrees of Borel sets up to isomorphism.

Recall that a pointclass $\mathbf{\Gamma} \subseteq P(\mathcal{N})$ has the *separation property* if for all disjoint sets $A, B \in \mathbf{\Gamma}$ there is $S \in \mathbf{\Gamma} \cap \mathbf{\Gamma}_c$ with $A \subseteq S \subseteq \overline{B}$. In [109, 115] the following deep relation of the Wadge reducibility to the separation property was established: For any Borel set $A$ which is non-self-dual exactly one of the principal ideals $\{X \mid X \leq_W A\}$, $\{X \mid X \leq_W \overline{A}\}$ has the separation property. The mentioned results give rise to the *Wadge hierarchy* which is, by definition, the sequence $\{\mathbf{\Sigma}_\alpha(\mathcal{N})\}_{\alpha<\nu}$ (where $\nu$ is the Wadge ordinal) of all non-self-dual principal ideals of $(\mathbf{B}(\mathcal{N}); \leq_W)$ that do not have the separation property and satisfy for all $\alpha < \beta < \nu$ the strict inclusion $\mathbf{\Sigma}_\alpha(\mathcal{N}) \subset \mathbf{\Delta}_\beta(\mathcal{N})$ where, as usual, $\mathbf{\Delta}_\beta(\mathcal{N}) = \mathbf{\Sigma}_\alpha(\mathcal{N}) \cap \Pi_\alpha(\mathcal{N})$.

The Wadge hierarchy subsumes the classical hierarchies in the Baire space, in particular $\mathbf{\Sigma}_\alpha(\mathcal{N}) = \mathbf{\Sigma}^{-1}_\alpha(\mathcal{N})$ for each $\alpha < \omega_1$, $\mathbf{\Sigma}_1(\mathcal{N}) = \mathbf{\Sigma}^0_1(\mathcal{N})$, $\mathbf{\Sigma}_{\omega_1}(\mathcal{N}) = \mathbf{\Sigma}^0_2(\mathcal{N})$, $\mathbf{\Sigma}_{\omega_1^{\omega_1}}(\mathcal{N}) = \mathbf{\Sigma}^0_3(\mathcal{N})$ and so on. Thus, the sets of finite Borel rank coincide with the sets of Wadge rank less than $\lambda = sup\{\omega_1, \omega_1^{\omega_1}, \omega_1^{(\omega_1^{\omega_1})}, \ldots\}$. Note that $\lambda$ is the smallest solution of the ordinal equation $\omega_1^\varkappa = \varkappa$. Hence, we warn the reader not to mistake $\mathbf{\Sigma}_\alpha(\mathcal{N})$ with $\mathbf{\Sigma}^0_\alpha(\mathcal{N})$. To give the reader a first impression about the Wadge ordinal we note that the rank of the QO $(\mathbf{\Delta}^0_\omega(\mathcal{N}); \leq_W)$ is the $\omega_1$-st solution of the ordinal equation $\omega_1^\varkappa = \varkappa$ [116]. As mentioned in the Introduction, Wadge hierarchy is non-standard in the sense that it is based on a highly original tool of infinite games, in contrast to the set-theoretic and topological methods used to investigate the classical hierarchies. As a result, Wadge hierarchy was originally defined only for the Baire space and its close relatives, and it is not straightforward to extend it to non-zero-dimensional spaces.

The Wadge hierarchy was an important development in classical DST not only as a unifying concept but also as a useful tool to investigate countably based zero-dimension spaces. We illustrate this with two examples. In [26] a complete classification (up to homeomorphism) of homogeneous zero-dimensional absolute Borel sets was achieved, completing a series of earlier results in this direction. In [27] it was shown that any Borel subspace of the Baire space with more than one point has a non-trivial auto-homeomorphism.

## 3 Well and Better Quasiorders

In this section we briefly discuss some notions and facts of WQO-theory relevant to our main theme. We do not try to be comprehensive in this survey or with references and from the numerous important concrete WQOs in the literature we choose mainly those directly relevant.

### *3.1 Well Quasiorders*

A *well quasiorder* (WQO) is a QO $Q = (Q; \leq)$ that has neither infinite descending chains nor infinite antichains. Note that for this paper WQOs are equivalent to the associated well partial orders (WPOs) and are only used to simplify some notation.

With any WQO $Q$ we associate its rank and also its *width* $w(P)$ defined as follows: if $P$ has antichains with any finite number of elements, then $w(Q) = \omega$, otherwise $w(Q)$ is the greatest natural number $n$ for which $Q$ has an antichain with $n$ elements. For instance, the structure of Wadge degrees of Borel sets is of width 2. Note that the notion of width maybe naturally refined in order to stratify the WQOs of infinite width (in the above "naive" sense) using ordinals.

There are several useful characterizations of WQOs. Some of them are collected in the following proposition. An infinite sequence $\{x_n\}$ in $Q$ is *good* if $x_i \leq x_j$ for some $i < j < \omega$, and *bad* otherwise. Let $\mathcal{F}(Q)$ be the class of all upward closed subsets of a QO $Q$.

**Proposition 3.1** *For a quasiorder $Q = (Q; \leq)$ the following are equivalent:*

*(1) $Q$ is WQO;*
*(2) every infinite sequence in $Q$ is good;*
*(3) every infinite sequence in $Q$ contains an increasing subsequence;*
*(4) any non-empty upward closed set in $Q$ has a finite number of minimal elements;*
*(5) the poset $(\mathcal{F}(Q); \supseteq)$ is well-founded;*
*(6) every linear order on $Q$ which extends $\leq$ is a well-order.*

It is easy to see that if $Q$ is WQO then any QO on $Q$ that extends $\leq$ is WQO, as well as any subset of $Q$ with the induced QO. Also, the cartesian product of two WQOs is WQO, and if $P$, $Q$ are WQOs which are substructures of some QO, then $P \cup Q$ is WQO. There are also many other useful closure properties of WQOs including the following two examples:

(1) If $Q$ is WQO then $(Q^*; \leq^*)$ is WQO where $Q^*$ is the set of finite sequences in $Q$ and $(x_1, \ldots, x_m) \leq^* (y_1, \ldots, y_n)$ means that for some strictly increasing $\varphi : \{1, \ldots, m\} \to \{1, \ldots, n\}$ we have $x_i \leq y_{\varphi(i)}$ for all $i$ (Higman's lemma [38]).
(2) If $Q$ is WQO then $(\mathcal{T}_Q; \leq_h)$ is WQO where $\mathcal{T}_Q$ is the set of finite $Q$-labeled trees $(T, c)$, $c : T \to Q$, and $\leq_h$ is the homomorphism QO ($h$-QO for short) defined as follows: $(T, c) \leq_h (S, d)$ if there is a monotone function $\varphi : (T, \sqsubseteq) \to (S, \sqsubseteq)$ such that $c(t) \leq d(\varphi(t))$ for all $t \in T$ (a consequence of Kruskal's theorem [49]). Recall that our trees are initial segments of $(\omega^*; \sqsubseteq)$.

We proceed with some concrete examples of WQOs. In Sect. 8.2 we will consider the important particular case of Higman's lemma $Q = \overline{k} = (k; =)$ for $2 \leq k < \omega$; in this case $(Q^*; \leq^*)$ is the subword relation on the set $k^*$ of finite words over $k$-letter alphabet $\{0, \ldots, k-1\}$. The Kruskal's theorem and Higman's lemma are close to optimal in the sense that the sets of finite structures in many natural classes (for instance, the set of finite distributive lattices of width 2 [86]) are not WQOs under the embeddability relation.

Similarly, we will be interested in the $h$-QO on the set $\mathcal{F}_k = \mathcal{F}_Q$ (where $Q = \overline{k} = (k; =)$) of finite $k$-labeled forests defined in the same manner as for trees; this QO first appeared in [34]. Also some weaker QOs $\leq_0, \leq_1, \leq_2$ on $\mathcal{F}_k$ are of interest [34, 99]. They are defined as follows: $(T, c) \leq_0 (S, d)$ (resp. $(T, c) \leq_1 (S, d)$, $(T, c) \leq_2 (S, d)$) if there is a monotone function $\varphi : (T, \sqsubseteq) \to (S, \sqsubseteq)$ such that $c = g \circ d \circ \varphi$ for some permutation $g : k \to k$ (resp. $c = g \circ d \circ \varphi$ for some $g : k \to k$, $\forall x, y \in T((x \sqsubseteq y \wedge c(x) \neq c(y)) \to d(\varphi(x)) \neq d(\varphi(y)))$. Obviously, $\leq_h \subseteq \leq_0 \subseteq \leq_1 \subseteq \leq_2$.

To obtain further interesting examples, we iterate the construction $Q \mapsto \mathcal{T}_Q$ starting with the antichain $\overline{k}$ of $k$ elements $\{0, \ldots, k-1\}$ (or with any other WQO $P$ in place of $\bar{k}$). Define the sequence $\{\mathcal{T}_k(n)\}_{n<\omega}$ of QOs by induction on $n$ as follows: $\mathcal{T}_k(0) = \overline{k}$ and $\mathcal{T}_k(n+1) = \mathcal{T}_{\mathcal{T}_k(n)}$. Note that $\mathcal{T}_k(1) = \mathcal{T}_k$. Identifying the elements $i < k$ of $\overline{k}$ with the corresponding minimal elements $s(i)$ of $\mathcal{T}_k(1)$, we may think that $\mathcal{T}_k(0)$ is an initial segment of $\mathcal{T}_k(1)$. Then $\mathcal{T}_k(n)$ is an initial segment of $\mathcal{T}_k(n+1)$ for each $n < \omega$, and hence $\mathcal{T}_k(\omega) = \bigcup_{n<\omega} \mathcal{T}_k(n)$ is WQO. Note that $\mathcal{T}_{\mathcal{T}_k(\omega)} = \mathcal{T}_k(\omega)$. The set $\mathcal{T}_k^{\sqcup}(\omega)$ of forests generated by the trees in $\mathcal{T}_k(\omega)$ is also WQO. The iterated $h$-QOs were first defined and studied in [98].

By a result in [98], in the case $k = 2$ the QO $(\mathcal{T}_k(\omega); \leq)$ is semi-well-ordered with the order type $\varepsilon_0$. This indicates a possible relation to the hierarchy theory. We will see its close relation to the so called fine hierarchy of sets in Sect. 6.1.

Finally, we mention the remarkable example of the QO of finite graphs with the graph-minor relation (we do not define this relation because do not discuss it in the sequel). Robertson–Seymour theorem [80] stating that this structure is WQO is one of the deepest known facts about finite graphs. The above-mentioned Higman's lemma and Kruskal's theorem are certainly much easier to prove than Robertson–Seymour's but their proofs are also non-trivial. Robertson–Seymour theorem is important for computer science because it implies that many graph problems are solvable in polynomial time, although such algorithms are hard to discover because it is hard to compute the minimal (under the minor relation) elements of a given upward closed sets of graphs (see e.g. [18] for details).

Along with the rank and width, there are some other important invariants of a WPO $(P; \leq)$. The most important is probably the *maximal order type* $o(P)$ which is the supremum of the order types of linearizations of $\leq$ (i.e., linear orders on $P$ which extend $\leq$). By a nice result of De Jongh and Parikh [20], every WPO $P$ has a linearization of order type $o(P)$. The computation of $o(P)$ for natural WPOs turned out an interesting and challenging task. Schmidt [82] computed the maximal order type of the Higman's WPO $(k^*; \leq^*)$ and gave upper bounds on the maximal order types of some other important WPOs including that of Kruskal's.

To our knowledge, the maximal order types of the other above-mentioned concrete WPOs are still unknown. Also the problem of relating $rk(P)$ and $o(P)$ discussed in [82] seems still to be open. In particular, there is no known characterization of pairs of ordinals $(\alpha, \beta)$ such that $\alpha = rk(P)$ and $\beta = o(P)$ for some WPO $P$ [70].

The structure $(\mathcal{T}_k^{\sqcup}(\omega); \leq_h)$ may be expanded by natural operations inducing a rich algebraic structure on the quotient-poset. These operations, introduced and studied in [97, 98, 100], are important for relating the $h$-QOs to hierarchy theory.

The binary operation $\oplus$ of disjoint union on $\mathcal{T}_k^{\sqcup}(\omega)$ is defined in the obvious way. For any $i < k$ and $F \in \mathcal{T}_k^{\sqcup}(\omega)$, let $p_i(F)$ be the tree in $\mathcal{T}_k(\omega)$ obtained from $F$ by adjoining the empty string labeled by $i$. Let $\mathbf{i}$ be the singleton tree $\{\varepsilon\}$ labeled by $i$. Define the binary operation $+$ on $\mathcal{T}_k^{\sqcup}(\omega)$ as follows: $F + G$ is obtained by adjoining a copy of $G$ below any leaf of $F$. One easily checks that $\mathbf{i} + F \equiv_h p_i(F), F \leq_h F + G, G \leq_h F + G, F \leq_h F_1 \to F + G \leq_h F_1 + G$, $G \leq_h G_1 \to F + G \leq_h F + G_1, (F + G) + H \equiv_h F + (G + H)$. Note that the set $\mathcal{T}_k^{\sqcup}(n)$ is closed under the operation $+$ for each $1 \leq n \leq \omega$. Define the function $s$ on $\mathcal{T}_k(\omega)$ as follows: $s(F)$ is the singleton tree carrying the label $F$. Note that $s(\mathbf{i}) = \mathbf{i}$ for each $i < k$, and $T \leq_h S$ iff $s(T) \leq_h s(S)$, for all $S, T \in \widetilde{\mathcal{T}}_k$. One easily checks the following properties:

**Proposition 3.2** *(1) For each* $1 \leq n \leq \omega$, $(\mathcal{T}_k^{\sqcup}(n); \oplus, \leq_h, p_0, \ldots, p_{k-1})$ *is a dc-semilattice.*

*(2) For any* $T \in \mathcal{T}_k(\omega)$, $F \mapsto s(T) + F$ *is a closure operator on* $(\mathcal{T}_k^{\sqcup}(\omega); \leq_h)$.

*(3) For all* $T, T_1 \in \mathcal{T}_k(\omega)$ *and* $F, F_1 \in \mathcal{T}_k^{\sqcup}(\omega)$, *if* $s(T) + F \leq_h s(T_1) + F_1$ *and* $T \not\leq_h T_1$ *then* $s(T) + F \leq_h F_1$.

*(4) The QO* $(\mathcal{T}_2^{\sqcup}(\omega); \leq_h)$ *is semi-well-ordered with order type* $\varepsilon_0$.

## *3.2 Better Quasiorders*

As we know, the closure properties of WQOs suffice to obtain nice WQOs using finitary constructions like forming finite labeled words or trees. But they do not suffice to establish that similar structures on, say, infinite words and trees are WQOs. A typical example is the attempt to extend the example (1) from the previous subsection to the set $Q^\omega$ of infinite $Q$-labeled sequences. As shown by Rado (see e.g. [66]), there is WQO $Q$ such that $Q^\omega$ is not WQO.

Nevertheless, it turns out possible to find a natural subclass of WQOs, called better quasiorders (BQOs) which contains most of the "natural" WQOs and has strong closure properties also for many infinitary constructions. In particular, if $Q$ is BQO then $Q^\omega$ is BQO. In this way it is possible to show that many important QOs are BQOs and hence also WQOs. The notion of BQO is due to C. Nash-Williams [73], we recall an alternative equivalent definition due to Simpson [103], see also [63].

Let $[\omega]^\omega$ be the subspace of the Baire space formed by all strictly increasing functions $p$ on $\omega$. Given $p \in [\omega]^\omega$, by $p^-$ we denote the result of dropping the first entry from $p$ (or equivalently, $p^- = p \setminus \{minX\}$, if we think of $p$ as an infinite subset $X = rng(p)$ of $\omega$). A QO Q is called BQO if, for any continuous function $f : [\omega]^\omega \to Q$ ($Q$ is assumed to carry the discrete topology), there is $p \in [\omega]^\omega$ with $f(p) \leq f(p^-)$.

It is easy to see that: any BQO is WQO; if $Q$ is BQO then any QO on $Q$ that extends $\leq$ is BQO, as well as any subset of $Q$ with the induced QO. Also, the cartesian product of two BQOs is BQO, and if $P$, $Q$ are BQO which are substructures of some

QO, then $P \cup Q$ is BQO. There are also many other useful closure properties of BQOs including the following:

(1) If $Q$ is BQO then $(Q^\omega; \leq^\omega)$ is BQO (in fact, this holds for sequences of arbitrary transfinite length [73, 74]).
(2) If $Q$ is BQO then $(\widetilde{\mathcal{T}}_Q; \leq_h)$ is BQO where $\widetilde{\mathcal{T}}_Q$ is the set of well-founded $Q$-labeled trees $(T, c)$, $c : T \to Q$, and $\leq_h$ is the homomorphism relation defined as follows: $(T, c) \leq_h (S, d)$ if there is a monotone function $\varphi : (T, \sqsubseteq) \to (S, \sqsubseteq)$ such that $c(t) \leq d(\varphi(t))$ for all $t \in T$ (a consequence of the extension of Kruskal's theorem to infinite trees [73, 74]).

We proceed with some concrete examples of BQOs. In Sect. 8.2 we will consider the particular case of (1) for $Q = \overline{k} = (k; =)$, $2 \leq k < \omega$; in this case $(Q^\omega; \leq^\omega)$ is the the subword relation on the set $k^\omega$ of infinite words over $k$-letter alphabet $\{1, \ldots, k-1\}$.

Similarly, we will be interested in the $h$-QO on the set $\widetilde{\mathcal{F}}_k = \widetilde{\mathcal{F}}_Q$, $Q = \overline{k} = (k; =)$, of well-founded $k$-labeled forests. We can also iterate the construction $Q \mapsto \widetilde{\mathcal{T}}_Q$ starting with the antichain $\overline{k}$ of $k$ elements $\{0, \ldots, k-1\}$. Define the sequence $\{\widetilde{\mathcal{T}}_k(\alpha)\}_{\alpha<\omega_1}$ of QOs by induction on $\alpha$ as follows: $\widetilde{\mathcal{T}}_k(0) = \overline{k}$, $\widetilde{\mathcal{T}}_k(\alpha+1) = \widetilde{\mathcal{T}}_{\widetilde{\mathcal{T}}_k(\alpha)}$, and $\widetilde{\mathcal{T}}_k(\lambda) = \bigcup_{\alpha<\lambda} \widetilde{\mathcal{T}}_{\widetilde{\mathcal{T}}_k(\alpha)}$ for limit $\lambda < \omega_1$. Note that $\widetilde{\mathcal{T}}_k(1) = \widetilde{\mathcal{T}}_k$. Identifying the elements $i < k$ of $\overline{k}$ with the corresponding minimal elements $s(i)$ of $\widetilde{\mathcal{T}}_k(1)$, we may think that $\widetilde{\mathcal{T}}_k(0)$ is an initial segment of $\mathcal{T}_k(1)$. Then $\widetilde{\mathcal{T}}_k(\alpha)$ is an initial segment of $\widetilde{\mathcal{T}}_k(\beta)$ for all $\alpha < \beta < \omega_1$, hence $\widetilde{\mathcal{T}}_k(\omega_1) = \bigcup_{\alpha<\omega_1} \widetilde{\mathcal{T}}_k(\alpha)$ is BQO. Note that $\widetilde{\mathcal{T}}_{\widetilde{\mathcal{T}}_k(\omega_1)} = \widetilde{\mathcal{T}}_k(\omega_1)$. The set $\widetilde{\mathcal{T}}_k^\sqcup(\omega_1)$ of countable disjoint unions of trees in $\widetilde{\mathcal{T}}_k(\omega_1)$ is also BQO. Similar iterated h-QOs were first studied in [100].

By a result in [100], in the case $k = 2$ the QO $\widetilde{\mathcal{T}}_k(\omega)$ is semi-well-ordered (in fact, $\widetilde{\mathcal{T}}_2(\omega_1)$ is also semi-well-ordered). This indicates a possible relation to the Wadge hierarchy from Sect. 2.5.

We also mention the remarkable example of the QO of countable linear orders with the embeddability relation. Laver's theorem [66] (see also [103]) stating that this structure is WQO (and thus resolving the Fraïssé conjecture) is one of the deepest applications of BQO-theory.

The maximal order types of the concrete BPOs introduced above seem to be unknown.

We conclude the list of examples of BQOs by a deep fact related to Wadge reducibility. For any QO $Q$ (equipped with the discrete topology), let $(Q^*; \leq^*)$ be the QO of Borel functions $A : \mathcal{N} \to Q$ with countable range, where $A \leq^* B$ means that for some continuous function $f$ on $\mathcal{N}$ we have $A(x) \leq B(f(x))$ for all $x \in \mathcal{N}$. Note that for $Q = (2, =)$ the QO $(Q^*; \leq^*)$ coincides with $(\mathbf{B}(\mathcal{N}); \leq_W)$. Theorem 3.2 in [27] states the following:

**Theorem 3.3** *If* $(Q; \leq)$ *is BQO then* $(Q^*; \leq^*)$ *is BQO.*

Let us briefly recall from [100] some operations on the iterated labeled forests and collect some of their properties used in the sequel. The $\omega$-ary operation $\bigoplus$ of disjoint union on $\widetilde{\mathcal{T}}_k^\sqcup(\omega_1)$ is defined in the obvious way. For any $i < k$ and

$F \in \widetilde{\mathcal{T}}_k^{\sqcup}(\omega_1)$, let $p_i(F)$ be the tree in $\widetilde{\mathcal{T}}_k(\omega_1)$ obtained from $F$ by adjoining the empty string labeled by $i$. Define the binary operation $+$ on $\widetilde{\mathcal{T}}_k^{\sqcup}(\omega_1)$ as follows: $F+G$ is obtained by adjoining a copy of $G$ below any leaf of $F$. One easily checks that $\mathbf{i}+F \equiv_h p_i(F), F \leq_h F+G, G \leq_h F+G, F \leq_h F_1 \rightarrow F+G \leq_h F_1+G$, $G \leq_h G_1 \rightarrow F+G \leq_h F+G_1, (F+G)+H \equiv_h F+(G+H)$. Note that the set $\widetilde{\mathcal{T}}_k^{\sqcup}(\alpha)$ is closed under the operation $+$ for each $1 \leq \alpha \leq \omega_1$. Define the function $s$ on $\widetilde{\mathcal{T}}_k(\omega_1)$ as follows: $s(F)$ is the singleton tree carrying the label $F$. Note that $s(\mathbf{i}) = \mathbf{i}$ for each $i < k$, and $T \leq_h S$ iff $s(T) \leq_h s(S)$, for all $S, T \in \widetilde{\mathcal{T}}_k$. One easily checks the following properties:

**Proposition 3.4** *(1) For any $1 \leq \alpha \leq \omega_1$, $(\widetilde{\mathcal{T}}_k^{\sqcup}(\alpha); \bigoplus, \leq_h, p_0, \ldots, p_{k-1})$ is a $dc\sigma$-semilattice.*
*(2) For any $T \in \widetilde{\mathcal{T}}_k(\omega_1)$, $F \mapsto s(T)+F$ is a closure operator on $(\widetilde{\mathcal{T}}_k^{\sqcup}(\omega_1); \leq_h)$.*
*(3) For all $T, T_1 \in \widetilde{\mathcal{T}}_k(\omega_1)$ and $F, F_1 \in \widetilde{\mathcal{T}}_k^{\sqcup}(\omega_1)$, if $s(T)+F \leq_h s(T_1)+F_1$ and $T \not\leq_h T_1$ then $s(T)+F \leq_h F_1$.*
*(4) The QO $(\widetilde{\mathcal{T}}_2^{\sqcup}(\omega); \leq_h)$ is semi-well-ordered with order type $sup\{\omega_1, \omega_1^{\omega_1}, \omega_1^{(\omega_1^{\omega_1})}, \ldots\}$.*

### 3.3 Computable Well Partial Orders

Here we briefly discuss computability properties of WQOs. The investigation of computable structures (see e.g. [7, 29]) is an important direction of computability theory. Recall that an algebraic structure $\mathbb{A} = (A; \sigma)$ of a finite signature $\sigma$ is *computable* if $A$ is a computable subset of $\omega$ and all signature relations and functions are computable. A structure is *computably presentable* if it is isomorphic to a computable structure. The notions of polynomial-time computable and polynomial-time computably presentable structure are defined in a similar manner using, say, the set $2^*$ of binary words instead of $\omega$.

It is easy to see that all concrete WQOs from Sect. 3.1, as well as the expansions of the $h$-QOs by functions, are computably presentable (as well as many other natural countable WQOs). As observed in [70], from results in [20] it follows that if a WPO is computable then its maximal order type is a computable ordinal. Moreover, the following result was established in [70]:

**Theorem 3.5** *(1) Every computable WPO has a computable linearization of maximal order type.*
*(2) There is no computable (even hyperatithmetical) function which, given an index for a computable WPO, returns an index for a computable maximal linearization of this WPO.*

Which of the concrete WQOs in Sect. 3.1 (and of their functional expansions) are polynomial-time presentable? By a well-known general fact [16], any computably presentable structure of a relational signature is in fact polynomial-time presentable.

Therefore, any of the structures $(\mathcal{F}_k; \leq_h)$ and $(\mathcal{F}_k; \leq_i)$ for $k \leq \omega$, $i \leq 2$ (and other WQOs in Sect. 3.1) is polynomial-time presentable. Since the presentations given by the proof in [16] are often artificial, one can ask further natural questions related to feasibility of our structures. We give some examples.

The sets $\mathcal{F}_k$, $k \leq \omega$, may be encoded in a natural way by words over a finite alphabet [36]. Will the relations $\leq_i$ be polynomial-time computable w.r.t. this coding? In [36] the following results were obtained:

**Theorem 3.6** *(1) The relation $\leq_h$ on $\mathcal{F}_\omega$ is computable in polynomial time.*
*(2) The relations $\leq_1$, $\leq_2$ on $\mathcal{F}_k$ are computable in polynomial time for $k < \omega$ and are NP-complete for $k = \omega$.*

Many natural questions concerning computability properties of WPO remain open, e.g.:

(1) Characterize the maximal order types of computable WPOs of rank $\omega$.
(2) Recall that the *degree spectrum* of a countable structure $\mathbb{A}$ is the set of Turing degrees **a** such that $\mathbb{A}$ is computably presentable relative to **a**. What are the degree spectra of countable WPOs? In particular, is it true that for any given countable graph there is a countable WPO with the same degree spectrum?
(3) Associate with any WPO $P$ the function $f_P : rk(P) \to \omega$ by: $f_P(\alpha)$ is the cardinality of $\{x \in P \mid rk_P(x) = \alpha\}$. Is $f_P$ computable provided that $P$ is computably presentable? It maybe shown that for several concrete examples of WPOs in from Sect. 3.1 the answer is positive, though in general we expect the negative answer.

## *3.4 Definability and Decidability Issues*

The study of definability and (un)decidability of first order theories is a central topic in logic and model theory. Here we briefly discuss such questions for some WQOs from Sects. 3.1 and 3.2. Along with WQOs, we also mention the infix order $\leq$ on the set of words $k^*$ where $u \leq v$ means that $v = xuy$ for some $x, y \in k^*$ (this relation is well founded but has infinite antichains).

Let $\mathbb{A} = (A; \sigma)$ be a structure of a given finite signature $\sigma$. As the understanding of definability in $\mathbb{A}$ assumes understanding of its automorphism group $Aut(\mathbb{A})$, we start with citing some facts following from the results in [57–60].

**Theorem 3.7** *(1) For any $k \geq 3$, the automorphism groups of the quotient-posets of $(\mathcal{F}_k; \leq_h)$ and $(\widetilde{\mathcal{F}}_k; \leq_h)$ are isomorphic to the symmetric group $\mathbf{S}_k$ of permutation of k elements.*
*(2) For any $k \geq 2$, $Aut(k^*; \leq^*) \simeq Aut(k^*; \leq) \simeq \mathbf{S}_k \times \mathbf{S}_2$.*

Recall that a relation $R \subseteq A^k$ is *definable in* $\mathbb{A}$ if there is a first-order $\sigma$-formula $\phi(x_1, \ldots, x_k)$ with $R = \{(x_1, \ldots, x_k) \in A^k \mid \mathbb{A} \models \phi(x_1, \ldots, x_k)\}$. A function on $A$ is definable if its graph is definable. An element $a \in A$ is definable if the set $\{a\}$

is definable. A structure is definable if its universe and all signature predicates and functions are definable.

The characterization of definable relations in a structure is quite important for understanding the structure. In a series of papers by Kudinov and Selivanov [54–60] a method of characterizing the definable relations was developed which might be of use for many similar structures on words, trees, graphs and so on (currently, the method mainly applies to well founded partial orders of rank $\omega$). We cite some facts following from [57–60] which characterize the definable relations in some of the mentioned structures.

Recall that a structure $\mathbb{A}$ equipped with a numbering $\alpha$ (i.e., a surjection from $\omega$ onto $A$) is *arithmetical*, if the equality predicate and all signature predicates and functions are arithmetical modulo $\alpha$. Obviously, any definable predicate on an arithmetical structure $(\mathbb{A}; \alpha)$ is arithmetical (w.r.t. $\alpha$) and invariant under the automorphisms of $\mathbb{A}$; we say that $(\mathbb{A}; \alpha)$ has the *maximal definability property* if the converse is also true, i.e., any arithmetical predicate invariant under the automorphisms of $\mathbb{A}$ is definable. The natural numberings of $\mathcal{F}_k$ and $k^*$ (which are not mentioned explicitly in the next theorem) make the structures $(\mathcal{F}_k; \leq_h)$, $(k^*; \leq^*)$ and $(k^*; \leq)$ arithmetical (even computable).

**Theorem 3.8** *Let $\mathbb{A}$ be one of the structures $(k^*; \leq^*)$, $(k^*; \leq)$ for $k \geq 2$ or the quotient-posets of $(\mathcal{F}_k; \leq_h)$ for $k \geq 3$. Then $\mathbb{A}$ has the maximal definability property.*

Recall (cf. [40, 75]) that a structure $\mathbb{B}$ of a finite relational signature $\tau$ is *biinterpretable* with a structure $\mathbb{C}$ of a finite relational signature $\rho$ if $\mathbb{B}$ is definable in $\mathbb{C}$ (in particular, there is a bijection $f : B \to B_1$ on a definable set $B_1 \subseteq C^m$ for some $m \geq 1$ which induces an isomorphism on the $\tau$-structure $\mathbb{B}_1$ definable in $\mathbb{C}$), $\mathbb{C}$ is definable in $\mathbb{B}$ (in particular, there is a similar bijection $g : C \to C_1$ on a definable set $C_1 \subseteq B^n$ for some $n \geq 1$), the function $g^m \circ f : B \to B^{nm}$ is definable in $\mathbb{B}$ and the function $f^n \circ g : C \to C^{mn}$ is definable in $\mathbb{C}$. Though the notion of biinterpretability looks sophisticated, its role in model theory is increasing because it gives a natural and strong equivalence relations on structures.

**Theorem 3.9** *The expansions of the structures $(k^*; \leq^*)$ and $(k^*; \leq)$, $k \geq 2$, by the constants for words of lengths 1 and 2, and the expansion of the quotient-posets of $(\mathcal{F}_k; \leq_h)$, $k \geq 3$, by the constants $\mathbf{0}, \ldots, \mathbf{k-1}$ for singleton trees, are biinterpretable with $\mathbb{N} = (\omega; +, \cdot)$.*

The closely related definability issues for embeddability relations on graphs and different classes of finite structures are now actively studied (note that most of these QOs have infinite antichains), see e.g. [44–48, 81, 120, 121]. Nevertheless, there are still many interesting open questions, including those for some structures mentioned in Sects. 3.1 and 3.2.

Recall that the *first-order theory* $FO(\mathbb{A})$ of a structure $\mathbb{A}$ of signature $\sigma$ is the set of $\sigma$-sentences true in $\mathbb{A}$. The investigation of algorithmic complexity of first-order theories of natural structures is a big chapter of logic, model theory and computability theory (see e.g. [29, 30, 114]). The proof of Theorem 3.8 implies that $FO(\mathbb{A})$ is

$m$-equivalent to first order arithmetic $FO(\mathbb{N})$ for any structure mentioned in that theorem.

**Theorem 3.10** *(1) Let $k \geq 3$ and $\mathbb{A}$ be the quotient-poset of some of $(\mathcal{F}_k; \leq_h)$, $(\mathcal{T}_k^{\sqcup}(\omega); \leq_h)$, or $(\mathcal{T}_k^{\sqcup}(n); \leq_h)$, $2 \leq n < \omega$. Then $FO(\mathbb{A}) \equiv_m FO(\mathbb{N})$.*
*(2) Let $\mathbb{A}$ be the quotient-poset of some of $(\widetilde{\mathcal{F}}_k; \leq_h)$, $(\widetilde{\mathcal{T}}_k^{\sqcup}(\omega_1); \leq_h)$, or of $(\widetilde{\mathcal{T}}_k^{\sqcup}(\alpha); \leq_h)$ for some $2 \leq \alpha < \omega_1$. Then $FO(\mathbb{N}) \leq_m FO(\mathbb{A})$*

Note that in item (2) we have only the lower bound. The natural upper bound is second-order arithmetic, the precise estimation is an interesting open problem. For many other interesting WPOs we do not know so comprehensive results as above, but for many first-order theories undecidability is known. By interpreting the finite structures of two equivalence relations [30] the following result about some other structures from Sect. 3.1 was established in [60, 99]. Recall that a structure of a finite signature is *hereditarily undecidable* if any its subtheory of the same signature is undecidable.

**Theorem 3.11** *For any $k \geq 3$, the first-order theories of the quotient-posets of $(\mathcal{F}_k; \leq_0)$, $(\mathcal{F}_k; \leq_1)$, and $(\mathcal{F}_k; \leq_2)$ are hereditarily undecidable.*

It is easy to see that the first-order theories of $(k^*; \leq^*)$, $(k^*; \leq)$ for $k = 1$ and of the quotient-poset of $(\mathcal{F}_2; \leq_h)$ are decidable. For most of the non-countable structures in Sect. 3.2 the complexity of first-order theories seem to be open.

Since the first-order theories of most of the mentioned structures are undecidable, it is natural to look for their decidable fragments. Such questions are interesting because they often originate from the computer science community, have many applications and, unlike most structures originated in mathematics, were considered relatively recently and many natural questions remain open (see e.g. [61] and references therein).

We illustrate this by the subword order on words. The study of subword order is important in many areas of computer science, e.g., in pattern matching, coding theory, theorem proving, algorithmics, automatic verification of unreliable channel systems [1, 52]. The reasoning about subwords involves ad hoc techniques quite unlike the standard tools that work well with prefixes and suffixes [50].

The study of $FO(A^*; \leq^*)$ was started by Comon and Treinen who showed undecidability for an expanded signature where $A$ has at least three letters. Kuske [61] showed that the 3-quantifier fragment of $FO(A^*; \leq^*)$ is undecidable. Karandikar and Schnoebelen showed that already the 2-quantifier theory is undecidable [52] and this is tight since the 1-quantifier fragment is decidable, in fact NP-complete [51, 61]. Karandikar and Schnoebelen also showed that the two-variable fragment is decidable [51] and that it has an elementary complexity upper bound [53]. Recently, it was shown [50] that, when constants are allowed, the 1-quantifier fragment is actually undecidable. This holds as soon as $A$ contains two distinct letters and exhibits a strong dependence on the presence of constants in the signature. To our knowledge, a similar detailed study for most of the above-mentioned structures is still to be done.

# 4 Wadge-Like Reducibilities in Quasi-Polish Spaces

As we know from Sect. 2.5, the structure of Wadge degrees in the Baire space refines the structure of levels of several popular hierarchies and serves as a tool to measure the topological complexity of some problems of interest in set-theoretic topology. There are several reasons and several ways to generalize the Wadge reducibility on the Baire space. For example, one can consider

(1) more complicated topological spaces instead of $\mathcal{N}$ (the notion of Wadge reducibility makes sense for arbitrary topological spaces);
(2) other natural classes of reducing functions in place of the continuous functions;
(3) reducibilities between functions rather than reducibilities between sets (the sets may be identified with their characteristic functions).

In any of the mentioned directions a certain progress has been achieved, although in many cases the situation typically becomes more complicated than in the classical case. In this section we mention some results in this direction.

## *4.1 Wadge-Like Reducibilities in the Baire Space*

For any family of pointclasses $\mathbf{\Gamma}$ and for any spaces $X, Y$, let $\mathbf{\Gamma}(X, Y)$ be the class of functions $f : X \to Y$ such that $f^{-1}(A) \in \mathbf{\Gamma}(X)$ for each $A \in \mathbf{\Gamma}(Y)$, and let $\mathbf{\Gamma}(X) = \mathbf{\Gamma}(X, X)$. Clearly, $\mathbf{\Gamma}(X)$ is closed under composition and contains the identity function, hence it induces a reducibility $\leq_{\mathbf{\Gamma}}$ on subsets of $X$ similar to the Wadge reducibility. For any $1 \leq \alpha < \omega_1$, let $D_\alpha(X, Y)$ denote $\mathbf{\Sigma}^0_\alpha(X, Y)$ and let $\leq_\alpha$ abbreviate $\leq_{\mathbf{\Sigma}^0_\alpha}$. Then $\leq_\alpha$ is a QO on $P(X)$. In particular $\leq_1$ coincides with the Wadge reducibility.

For any $\alpha < \omega_1$ and any spaces $X, Y$, let $D^W_\alpha(X, Y)$ be the class of functions $f : X \to Y$ such that there is a partition $\{D_n\}$ of $X$ to $\mathbf{\Sigma}^0_\alpha$-sets and a sequence $f_n : D_n \to Y$ of continuous functions with $f = \bigcup_{n<\omega} f_n$. Note that $D^W_\alpha(X, Y) \subseteq D_\alpha(X, Y)$. We again set $D^W_\alpha(X) = D^W_\alpha(X, X)$. Clearly, $D^W_\alpha(X)$ is closed under composition and contains the identity function, hence it induces a reducibility $\leq^W_\alpha$ on subsets of $X$.

The study of the reducibility by Borel functions on the Baire space was initiated by Andretta and Martin in [10], the reducibility $\leq_2$ was studied by Andretta [9], the other of just defined reducibilities (as well as many other so called amenable reducibilities) were comprehensively investigated by Motto-Ross [71]. The next result, which is a very particular case of the results in [71], shows that these reducibilities behave similarly to the Wadge reducibility:

**Theorem 4.1** *For any* $1 \leq \alpha < \omega_1$*, the quotient-posets of* $(\mathbf{B}(\mathcal{N}); \leq_\alpha)$ *and* $(\mathbf{B}(\mathcal{N}); \leq^W_\alpha)$ *are isomorphic to that of* $(\mathbf{B}(\mathcal{N}); \leq_W)$.

## 4.2 Wadge Reducibility of k-Partitions in the Baire Space

Let $2 \leq k < \omega$. By a *k-partition of* $\mathcal{N}$ we mean a function $A : \mathcal{N} \to k = \{0, \ldots, k-1\}$ often identified with the sequence $(A_0, \ldots, A_{k-1})$ where $A_i = A^{-1}(i)$ are the components of $A$. Obviously, 2-partitions of $\mathcal{N}$ can be identified with the subsets of $\mathcal{N}$ using the characteristic functions. The set of all $k$-partitions of $\mathcal{N}$ is denoted $k^{\mathcal{N}}$, thus $2^{\mathcal{N}} = P(\mathcal{N})$. The Wadge reducibility on subsets of $\mathcal{N}$ is naturally extended to $k$-partitions: for $A, B \in k^{\mathcal{N}}$, $A \leq_W B$ means that $A = B \circ f$ for some continuous function $f$ on $\mathcal{N}$. In this way, we obtain the QO $(k^{\mathcal{N}}; \leq_W)$. For any pointclass $\mathbf{\Gamma} \subseteq P(\mathcal{N})$, let $\mathbf{\Gamma}(k^{\mathcal{N}})$ be the set of $k$-partitions of $\mathcal{N}$ with components in $\mathbf{\Gamma}$.

In contrast with the Wadge degrees of sets, the structure $(\mathbf{B}(k^{\mathcal{N}}); \leq_W)$ for $k > 2$ has antichains of any finite size. Nevertheless, a basic property of the Wadge degrees of sets may be lifted to $k$-partitions, as the following very particular case of Theorem 3.2 in [27] (see Theorem 3.2 in Sect. 3.2) shows:

**Theorem 4.2** *For any* $2 \leq k < \omega$*, the structure* $(\mathbf{B}(k^{\mathcal{N}}); \leq_W)$ *is WQO.*

Although this result gives an important information about the Wadge degrees of Borel $k$-partitions, it is far from a characterization. Here we briefly discuss some steps to such a characterization made in [34, 94, 100, 101]). The approach of [94, 101] is to characterize the initial segments $(\mathbf{\Delta}^0_\alpha(k^{\mathcal{N}}); \leq_W)$ for bigger and bigger ordinals $2 \leq \alpha < \omega_1$. In [94] this was done for $\alpha = 2$, in [101] for $\alpha = 3$ where also a way to the general characterization was sketched. This characterization uses the iterated $h$-QO from Sect. 3.2. The main idea is to expand the structure by suitable operations whose properties are similar to those on the labeled forests and use the similarity to prove isomorphism. Some of these operations extend the corresponding operations on sets from [116] to $k$-partitions.

Let $\bigoplus_i A_i$ be the disjoint union of a sequence of elements $A_0, A_1, \ldots$ of $k^{\mathcal{N}}$. Let $\mathcal{N}^+ := \{1, 2, \ldots\}^\omega$ and for $x \in \mathcal{N}^+$ let $x^- := \lambda i.x(i) - 1$, so $x^- \in \mathcal{N}$. Define the binary operation $+$ on $k^{\mathcal{N}}$ as follows: $(A + B)(x) := A(x^-)$ if $x \in \mathcal{N}^+$, otherwise $(A + B)(x) := B(y)$ where $y$ is the unique element of $\mathcal{N}$ such that $x = \sigma 0 y$ for a unique finite sequence $\sigma$ of positive integers. For any $i < k$, define a unary operation $p_i$ on $k^{\mathcal{N}}$ by $p_i(A) := \mathbf{i} + A$ where $\mathbf{i} := \lambda x.i$ are the constant $k$-partitions (which are precisely the distinct minimal elements of $(k^{\mathcal{N}}; \leq_W)$). For any $i < k$, define a unary operation $q_i$ on $k^{\mathcal{N}}$ (for $k = 2$, $q_0$ and $q_1$ coincide with the Wadge's operations $\sharp$ and $\flat$ from Section III.E of [116]) as follows: $q_i(A)(x) := i$ if $x$ has infinitely many zeroes, $q_i(A)(x) := A(x^-)$ if $x$ has no zeroes, and $q_i(A)(x) := A(y^-)$ otherwise where $y$ is the unique element of $\mathcal{N}^+$ such that $x = \sigma 0 y$ for a string $\sigma$ of non-negative integers. The introduced operations are correctly defined on Wadge degrees.

The result next from [101] characterizes some subalgebras of the Wadge degrees generated from the minimal degrees $\{\mathbf{0}\}, \ldots, \{\mathbf{k-1}\}$. The proof uses Theorem 4.2.

**Theorem 4.3** *The quotient-poset of* $(\mathbf{\Delta}^0_2(k^{\mathcal{N}}); \leq_W)$ *is generated from the degrees* $\{\mathbf{0}\}, \ldots, \{\mathbf{k-1}\}$ *by the operations* $\bigoplus, p_0, \ldots, p_{k-1}$*. The quotient-poset of* $(\mathbf{\Delta}^0_3(k^{\mathcal{N}}); \leq_W)$ *is generated from* $\{\mathbf{0}\}, \ldots, \{\mathbf{k-1}\}$ *by the operations* $\bigoplus, +, q_0, \ldots, q_{k-1}$.

The next result from [101] characterizes the structures above using Proposition 3.4.

**Theorem 4.4** *(1) The quotient-posets of* $(\mathbf{\Delta}^0_2(k^{\mathcal{N}}); \leq_W)$ *and of* $(\widetilde{\mathcal{F}}_k; \leq_h)$ *are isomorphic.*
*(2) The quotient-posets of* $(\mathbf{\Delta}^0_3(k^{\mathcal{N}}); \leq_2)$ *and of* $(\widetilde{\mathcal{F}}_k; \leq_h)$ *are isomorphic.*
*(3) The quotient-posets of* $(\mathbf{\Delta}^0_3(k^{\mathcal{N}}); \leq_W)$ *and of* $(\widetilde{\mathcal{T}}^{\sqcup}_k(2); \leq_h)$ *are isomorphic.*

We describe functions that induce the isomorphisms of the quotient-posets. For (1), let $(T; c) \in \widetilde{\mathcal{T}}_k$. Associate with any node $\sigma \in T$ the $k$-partition $\mu_T(\sigma)$ by induction on the rank $rk(\sigma)$ of $\sigma$ in $(T; \sqsupseteq)$ as follows: if $rk(\sigma) = 0$, i.e. $\sigma$ is a leaf of $T$ then $\mu_T(\sigma) := \mathbf{i}$ where $i = c(\sigma)$; otherwise, $\mu_T(\sigma) := p_i(\bigoplus\{\mu_T(\sigma n) \mid n < \omega, \sigma n \in T\})$. Now, define a function $\mu : \widetilde{\mathcal{T}}_k \to k^{\mathcal{N}}$ by $\mu(T) := \mu_T(\varepsilon)$. Next extend $\mu$ to $\widetilde{\mathcal{F}}_k$ by $\mu(F) := \bigoplus\{\mu_T(n) \mid n < \omega, (n) \in T\}$. Then $\mu \to \widetilde{\mathcal{F}}_k$ induces the isomorphism in (1).

The isomorphism $\nu$ in (2) is constructed just as $\mu$ but with $q_i$ instead of $p_i$.

Towards the isomorphism in (3), let $(T; c) \in \widetilde{\mathcal{T}}_k(2)$. Relate to any node $\sigma \in T$ the $k$-partition $\rho_T(\sigma)$ by induction on the rank $rk(\sigma)$ of $\sigma$ in $(T; \sqsupseteq)$ as follows: if $rk(\sigma) = 0$ then $\rho_T(\sigma) := \nu(Q)$ where $Q = c(\sigma) \in \widetilde{\mathcal{T}}_k$; otherwise, $\rho_T(\sigma) := \nu(Q) + (\bigoplus\{\rho_T(\sigma n) \mid n < \omega, \sigma n \in T\})$. Now define a function $\rho : \widetilde{\mathcal{T}}_{\widetilde{\mathcal{T}}_k} \to k^{\mathcal{N}}$ by $\rho(T) := \rho_T(\varepsilon)$. Finally, extend $\rho$ to $\widetilde{\mathcal{T}}^{\sqcup}_k(2)$ by $\rho(F) := \bigoplus\{\rho_T(n) \mid n < \omega, (n) \in T\}$ where $T := \{\varepsilon\} \cup F$. Then $\rho$ induces the isomorphism in (3).

As conjectured in [101], the results above may be extended to larger segments $(\mathbf{\Delta}^0_\alpha(k^{\mathcal{N}}); \leq_W)$, $4 \leq \alpha < \omega_1$. Using the Kuratowski relativization technique [10, 71, 116], we can define for any $1 \leq \beta < \omega_1$ the binary operation $+_\beta$ on $k^{\mathcal{N}}$ such that $+_1$ coincides with $+$ and, for any $2 \leq \alpha < \omega_1$, the quotient-poset of $(\mathbf{\Delta}^0_\alpha(k^{\mathcal{N}}); \leq_W)$ is generated from $\{\mathbf{0}\}, \ldots, \{\mathbf{k-1}\}$ by the operations $\bigoplus$ and $+_\beta$ for all $1 \leq \beta < \alpha$. The extension of Theorem 4.4 could probably be obtained by defining suitable iterated versions of the $h$-quasiorder in the spirit of [100]. Since $\mathbf{B}(k^{\mathcal{N}}) = \bigcup_{\alpha<\omega_1} \mathbf{\Delta}^0_\alpha(k^{\mathcal{N}})$, we obtain the characterization of Wadge degrees of Borel $k$-partitions. Note that item (2) suggests that the extension of Theorem 4.1 to $k$-partitions holds.

## *4.3 Wadge Reducibility in Quasi-Polish Spaces*

A straightforward way to extend the Wadge hierarchy to non-zero-dimensional spaces would be to show that Wadge reducibility in such spaces behaves similarly to its behaviour in the Baire space, e.g. it is a semi-well-order. Unfortunately, this is not the case for many natural spaces: Wadge reducibility is often far from being WQO.

Using the methods of [116] it is easy to check that the structure $(\mathbf{B}(X); \leq_W)$ of Wadge degrees of Borel sets in any zero-dimensional Polish space $X$ remains semi-well-ordered. In contrast, the structure of Wadge degrees in non-zero-dimensional spaces is typically more complicated. Hertling demonstrated this in [35] by showing that there are infinite antichains and infinite descending chains in the structure of

Wadge degrees of $\mathbf{\Delta}^0_2(\mathbb{R})$-sets. This result has been strengthened in [43] to the result that any poset of cardinality $\omega_1$ embeds into $(\mathbf{B}(\mathbb{R}); \leq_W)$.

Schlicht showed in [102] that the structure of Wadge degrees on *any* non zero-dimensional Polish space must contain infinite antichains. Thus, the class of zero-dimensional Polish spaces maybe characterized in terms of Wadge reducibility within the Polish spaces.

Selivanov showed in [91] that the structure of Wadge degrees of finite Boolean combinations of open sets in many $\omega$-algebraic domains is semi-well-ordered, but already for $\mathbf{\Delta}^0_2$-sets the structure contains antichains of size 4. Additional information on the structure of Wadge-degrees in non-zero-dimensional spaces maybe found e.g. in [42, 43, 91].

The mentioned results show that it is not straightforward to extend Wadge hierarchy to quasi-Polish spaces using the Wadge reducibility in those spaces. We return to this question in Sect. 6.

## 4.4 Weak Homeomorphisms Between Quasi-Polish Spaces

As the Wadge reducibility in non-zero-dimensional quasi-Polish spaces is often far from being WQO, one could hope to find natural weaker notions of reducibility that induce semi-well-ordered degree structures. Good candidates are $\leq_\alpha$ and $\leq^W_\alpha$, but before looking on them we briefly discuss here some properties of the corresponding classes of functions. All uncredited results in this section are from [72].

It is a classical result of DST that every two uncountable Polish spaces $X, Y$ are Borel-isomorphic (see e.g. Theorem 15.6 in [65]). The next proposition extends this result to the context of uncountable quasi-Polish spaces and computes an upper bound for the complexity of the Borel-isomorphism. We write $X \simeq_\alpha Y$ to denote that there is a bijection $f : X \to Y$ such that $f \in D_\alpha(X, Y)$ and $f^{-1} \in D_\alpha(Y, X)$. The relation $\simeq^W_\alpha$ is defined in the same way.

**Proposition 4.5** *(1) Let $X, Y$ be two uncountable quasi-Polish spaces. Then $X \simeq_\omega Y$.*
*(2) Every quasi-Polish space is $D^W_4$-isomorphic to an $\omega$-algebraic domain.*
*(3) $\mathcal{N} \simeq^W_2 \omega \sqcup \mathcal{N}$.*
*(4) If $X$ is a $\sigma$-compact quasi-Polish space then $\mathcal{N} \not\simeq^W_2 X$. In particular, $\mathcal{N} \not\simeq^W_2 \mathcal{C}$, $\mathcal{N} \not\simeq^W_2 \mathbb{R}^n$ for every $n < \omega$, and $\mathcal{N} \not\simeq^W_2 \omega^{\leq\omega}$ where $\omega^{\leq\omega}$ is the $\omega$-algebraic domain $(\omega^* \cup \omega^\omega, \sqsubseteq)$ endowed with the Scott topology.*
*(5) $\mathcal{N} \simeq^W_3 \mathcal{C} \simeq^W_3 \omega^{\leq\omega} \simeq^W_3 \mathbb{R}^n$ for every $1 \leq n < \omega$.*

Our next goal is to extend Proposition 4.5 (5) to a wider class of quasi-Polish spaces (see Theorem 4.10). Such generalization will involve the following definition of the (inductive) topological dimension of a space $X$, denoted in this paper by $\dim(X)$ — see e.g. [41, p. 24].

**Definition 4.6** The empty set $\emptyset$ is the only space with *dimension* $-1$, in symbols $\dim(\emptyset) = -1$.

Let $\alpha$ be an ordinal and $\emptyset \neq X$. We say that $X$ has *dimension* $\leq \alpha$, $\dim(X) \leq \alpha$ in symbols, if every $x \in X$ has arbitrarily small neighborhoods whose boundaries have dimension $< \alpha$, i.e. for every $x \in X$ and every open set $U$ containing $x$ there is an open $x \in V \subseteq U$ such that $\dim(\partial V) \leq \beta$ (where $\partial V = cl(V) \setminus V$) for some $\beta < \alpha$.

We say that a space $X$ has *dimension* $\alpha$, $\dim(X) = \alpha$ in symbols, if $\dim(X) \leq \alpha$ and $\dim(X) \not\leq \beta$ for all $\beta < \alpha$.

Finally, we say that a space $X$ has *dimension* $\infty$, $\dim(X) = \infty$ in symbols, if $\dim(X) \not\leq \alpha$ for every ordinal $\alpha$.

It is obvious that the dimension of a space is a topological invariant (i.e. $\dim(X) = \dim(Y)$ whenever $X$ and $Y$ are homeomorphic). Moreover, one can easily check that $\dim(X) \leq \alpha$ (for $\alpha$ an ordinal) if and only if there is a base of the topology of $X$ consisting of open sets whose boundaries have dimension $< \alpha$. Therefore, if $X$ is countably based and $\dim(X) \neq \infty$, then $\dim(X) = \alpha$ for some *countable* ordinal $\alpha$.

*Example 4.7 Finite dimension.*

(1) $\dim(\mathcal{N}) = \dim(\mathcal{C}) = 0$.
(2) $\dim(\mathbb{R}^n) = n$ for every $0 \neq n \leq \omega$.
(3) For $n < \omega$, let $L_n$ be the (finite) quasi-Polish space obtained by endowing $(n, \leq)$ with the Scott (equivalently, the Alexandrov) topology: then $\dim(L_n) = n - 1$.

*Example 4.8 Transfinite dimension.*

(1) The disjoint union $X = \bigsqcup_{0 \neq n \in \omega} [0, 1]^n$ of the $n$-dimensional cubes $[0, 1]^n$ is a Polish space of dimension $\omega$.
(2) $\dim(\omega^{\leq \omega}) = \omega$.
(3) For $\alpha < \omega_1$, let $L_{\alpha+1}$ be the quasi-Polish space obtained by endowing the poset $(\alpha + 1, \leq)$ with the Scott topology. Then $\dim(L_{\alpha+1}) = \alpha$.

*Example 4.9 Dimension $\infty$.*

(1) The Hilbert cube $[0, 1]^\omega$, the space $\mathbb{R}^\omega$ (both endowed with the product topology), and the Scott domain $P\omega$ all have dimension $\infty$.
(2) Let $\mathsf{C}_\infty$ be the quasi-Polish space obtained by endowing the poset $(\omega, \geq)$ with the Scott (equivalently, the Alexandrov) topology. Then $\mathsf{C}_\infty$ is a (scattered) countable space with $\dim(\mathsf{C}_\infty) = \infty$. Hence the space $\mathsf{UC}_\infty = \mathsf{C}_\infty \times \mathcal{N}$, endowed with the product topology, is an (uncountable) quasi-Polish space of dimension $\infty$.

We are ready to formulate the extension of Proposition 4.5.

**Theorem 4.10** *Let $X$ be an uncountable quasi-Polish space.*

*(1) If* $\dim(X) \neq \infty$ *then* $\mathcal{N} \simeq_3^W X$*;*
*(2) If* $\dim(X) = \infty$ *and $X$ is Polish then* $\mathcal{N} \not\preceq_\alpha^W X$ *for every* $\alpha < \omega_1$ *and* $\mathcal{N} \not\preceq_n X$ *for every* $n < \omega$*;*
*(3)* $P\omega \not\preceq_\alpha^W \mathcal{N}$ *for every* $\alpha < \omega_1$ *and* $P\omega \not\preceq_n \mathcal{N}$ *for every* $n < \omega$*. The same result holds when replacing $P\omega$ with any other quasi-Polish space which is universal for (compact) Polish spaces;*
*(4)* $\mathsf{UC}_\infty \simeq_2^W \mathcal{N}$*. Therefore,* $\mathsf{UC}_\infty \not\preceq_\alpha^W X$ *(*$\alpha < \omega_1$*) and* $\mathsf{UC}_\infty \not\preceq_n X$ *(*$n \in \omega$*) for $X$ a Polish space of dimension $\infty$ (e.g.* $X = [0, 1]^\omega$ *or* $X = \mathbb{R}^\omega$*) or* $X = P\omega$*.*

## 4.5 Weak Reducibilities in Quasi-Polish Spaces

Here we show that most of the reducibilities $\leq_\alpha$, $\leq_\alpha^W$ are in fact semi-well-orders on the Borel sets in quasi-Polish spaces. The following result from [72] is an immediate corollary of Theorem 4.10.

**Theorem 4.11** *Let $X$ be an uncountable quasi-Polish space.*

*(1) If* $\dim(X) = 0$ *and* $1 \leq \alpha < \omega_1$ *then* $(\mathbf{B}(X); \leq_\alpha)$ *and* $(\mathbf{B}(X); \leq_\alpha^W)$ *are semi-well-ordered.*
*(2) Assume* $\dim(X) = 0$ *and* $2 \leq \alpha < \omega_1$*. If $X$ is $\sigma$-compact then the quotient-posets of* $(\mathbf{B}(X); \leq_\alpha)$ *and* $(\mathbf{B}(X); \leq_\alpha^W)$ *are isomorphic to the quotient-poset of* $(\mathbf{B}(\mathcal{C}); \leq_W)$*, otherwise they are isomorphic to the quotient-poset of* $(\mathbf{B}(\mathcal{N}); \leq_W)$*.*
*(3) If* $\dim(X) \neq \infty$ *and* $3 \leq \alpha < \omega_1$ *then the quotient-posets of* $(\mathbf{B}(X); \leq_\alpha)$ *and* $(\mathbf{B}(X); \leq_\alpha^W)$ *are isomorphic to the quotient-poset of* $(\mathbf{B}(\mathcal{N}); \leq_W)$*.*
*(4) If $X$ is universal for Polish (respectively, quasi-Polish) spaces and* $3 \leq \alpha < \omega_1$ *then the quotient-posets of* $(\mathbf{B}(X); \leq_\alpha)$ *and* $(\mathbf{B}(X); \leq_\alpha^W)$ *are isomorphic to the quotient-posets of* $(\mathbf{B}([0, 1]^\omega); \leq_W)$ *and of* $(\mathbf{B}(P\omega); \leq_W)$*.*
*(5) If* $\omega \leq \alpha < \omega_1$ *then the quotient-posets of* $(\mathbf{B}(X); \leq_\alpha)$ *and* $(\mathbf{B}(X); \leq_\alpha^W)$ *are isomorphic to the quotient-poset of* $(\mathbf{B}(\mathcal{N}); \leq_W)$*.*

Similar results clearly hold also for $k$-partitions which means that, for most of $\alpha$, the relations $\leq_\alpha$ and $\leq_\alpha^W$ on the Borel $k$-partitions of quasi-Polish spaces are again intimately related to the iterated $h$-QO in Sect. 3.2.

Theorem 4.11 leaves open the question about the structure of degrees under $\leq_2$. The next two results from [72] give some information on this.

**Theorem 4.12** *(1) There are infinite antichains in* $(\mathbf{B}([0, 1]); \leq_2)$*.*
*(2) The quasiorder* $(P(\omega), \subseteq^*)$ *of inclusion modulo finite sets on* $P(\omega)$ *embeds into* $(\mathbf{\Sigma}_2^0(\mathbb{R}^2), \leq_2)$*.*

# 5 Other Reducibilities and Hierarchies

In this section we discuss some reducibilities and hierarchies on objects more complex than sets and $k$-partitions (in particular, on equivalence relations or functions between spaces). They provide useful tools for measuring descriptive complexity of such objects but, unfortunately, in most interesting cases they are far from being WQOs.

## *5.1 Borel Reducibility of Borel Equivalence Relations*

In mathematics one often deals with problems of classifying objects up to some notion of equivalence by invariants. Via suitable encodings, these objects can be viewed as elements of a standard Borel space $X$ and the equivalence turns out to be a Borel equivalence relation $E$ on $X$. In this and the next subsection we briefly discuss (following [8, 31]) some reducibilities on Borel equivalence relations which provide a mathematical framework for measuring the complexity of such classification problems. The most fundamental is probably Borel reducibility defined as follows.

By *standard Borel space* we mean a Polish (equivalently, quasi-Polish) space equipped with its Borel structure. Let $E, F$ be equivalence relations on standard Borel spaces $X, Y$, respectively. We say that $E$ is *Borel reducible* to $F$ (in symbols, $E \leq_B F$) if there is a Borel map $f : X \to Y$ such that $xEy$ iff $f(x)Ff(y)$. We denote by $(BER :\leq_B)$ the QO of Borel equivalence relations with the Borel reducibility.

For any standard Borel space $X$, denote by $X$ also the equality relation on this space, and let $n$ be any such space of finite cardinality $n$. Then we clearly have $1 <_B 2 <_B \cdots <_B \omega < \mathbb{R}$. Let $E_0$ be the *Vitali equivalence relation* on $\mathbb{R}$ defined by: $xE_0y$ iff $x - y \in \mathbb{Q}$. Then we have the following deep result which includes the so called Silver Dichotomy and General Glimm-Effros Dichotomy (due to Harrington-Kechris-Louveau):

**Theorem 5.1** *The chain* $1 <_B 2 <_B \cdots <_B \omega < \mathbb{R} <_B E_0$ *is an initial segment of* $(BER; \leq_B)$.

The QO $(BER; \leq_B)$ is very rich, in particular it has no maximal elements (by the Friedman-Stanley jump theorem). In contrast, the set $CBER$ of countable Borel equivalence relations (those with countable equivalence classes) has the greatest element $E_\infty$ which is equivalent to many natural equivalence relations. For instance, it was recently shown in [68, 69] that the equivalence relations on the Cantor space induced by some natural reducibilities in computability theory (e.g., polynomial-time many-one, polynomial-time Turing, and the arithmetical reducibilities) are Borel equivalent to $E_\infty$.

According to Feldman-Moore theorem, the countable Borel equivalence relations coincide with the orbit equivalence relations $E_G$ induced by Borel actions $G \times X \to X$ of (discrete) countable groups $G$ where $xE_Gy$ iff $g \cdot x = y$ for some $g \in G$. This suggests that the QO $(CBER; \leq_B)$ is rich. Is it a WQO? Unfortunately, this is far from being the case, as the following deep result from [8] demonstrates.

**Theorem 5.2** *The poset of Borel subsets of $\mathcal{N}$ under inclusion can be embedded in the quasiorder $(CBER; \leq_B)$.*

The last result shows that WQO-theory is not directly related to the QO $(CBER; \leq_B)$. To find such a relation, one probably would have to search for natural substructures of $(CBER; \leq_B)$ which form WQOs.

## 5.2 Continuous Reducibility of Borel Equivalence Relations

Although the Borel reducibility from the previous section is quite important due to its deep relation to classical mathematics, it is in a sense too coarse. For instance, it does not distinguish between Polish and quasi-Polish spaces of the same cardinality. There are natural finer versions, the most important of which is continuous reducibility. Though it may be defined for arbitrary quasi-Polish spaces, we briefly discuss it here only for the Baire space.

For equivalence relations $E, F$ on $\mathcal{N}$, $E$ is *continuously reducible* to $F$, in symbols $E \leq_c F$, if there is a continuous function $f$ on $\mathcal{N}$ such that for all $x, y \in \mathcal{N}$, $xEy$ iff $f(x)Ff(y)$. The QO $(ER(\mathcal{N}); \leq_c)$, where $ER(\mathcal{N})$ is the set of all equivalence relations on $\mathcal{N}$, and its substructure $(BER(\mathcal{N}); \leq_c)$ formed by Borel equivalence relations, are extremely complicated, as it follows from Theorem 5.2.

Let $(ER_k; \leq_c)$ (resp. $(BER_k; \leq_c)$) be the initial segment of $(ER(\mathcal{N}); \leq_c)$ (resp. $(BER(\mathcal{N}); \leq_c)$) formed by the set $ER_k$ of equivalence relations which have at most $k$ equivalence classes. We relate this substructure to the structure $(k^{\mathcal{N}}; \leq_0)$ where $\mu \leq_0 \nu$ iff $\mu = \varphi \circ \nu \circ f$ for some continuous function $f$ on $\mathcal{N}$ and for some permutation $\varphi$ of $\{0, \ldots, k-1\}$. Since $\leq_W$ implies $\leq_0$, the results of Sect. 4.2 imply that $(BER_k; \leq_c)$ is WQO. The following is straightforward:

**Proposition 5.3** *For any $2 \leq k < \omega$, the function $\nu \mapsto E_\nu$, where $pE_\nu q$ iff $\nu(p) = \nu(q)$, induces an isomorphism between the quotient-structures of $(k^{\mathcal{N}}; \leq_0)$ and $(ER_k; \leq_c)$.*

Let $\mathcal{B}_k$ be the subset of $k^{\mathcal{N}}$ formed by the $k$-patitions whose components are finite Boolean combinations of open sets. By a result in [34], the quotient-posets of $(\mathcal{B}_k; \leq_W)$ and $(\mathcal{F}_k; \leq_h)$ are isomorphic. It is not hard to modify the proof in [34] to show that the quotient-posets of $(\mathcal{B}_k; \leq_0)$ and $(\mathcal{F}_k; \leq_0)$ are isomorphic.

## 5.3 Hierarchies and Reducibilities of Functions

$k$-Partitions are of course very special functions between spaces, and it is natural to search for natural hierarchies and reducibilities between functions. Some such hierarchies induced by the classical hierarchies of sets are known from the beginning of DST.

For any family $\Gamma$ of pointclasses and any quasi-Polish spaces $X, Y$, a function $f : X \to Y$ is $\Gamma$-*measurable* if $f^{-1}(U) \in \Gamma(X)$ for each open set $U \subseteq Y$. For each $1 \leq \alpha < \omega_1$, let $B_\alpha(X)$ be the class of $\mathbf{\Sigma}^0_\alpha$-measurable functions. The ascending sequence $\{B_\alpha(X)\}$, known as the *Baire hierarchy in X* exhausts all the Borel functions and is important for DST.

Equally natural is the hierarchy $\{D_\alpha(X)\}$ formed by the classes from Sect. 4.1. It is also ascending and exhausts the Borel functions [71]. A pleasant property of this hierarchy is that all of its levels are closed under composition. A problem with both hierarchies is that they are coarse, but it is not clear how to extend the classical hierarchy theory of sets (and $k$-partitions) to these hierarchies of functions (in particular, we are not aware of natural analogues for the Hausdorff–Kuratowski theorem). Also, there is no clear relation to WQO-theory so far.

The Borel hierarchy also induces a natural QO $\leq_B$ on Borel functions which we define only for the Baire space. Associate with any Borel function $f$ on $\mathcal{N}$ the monotone function $d_f$ on $(\omega_1; \leq)$ as follows: $d_f(\alpha)$ is the smallest $\beta$ with $\forall A \in \mathbf{\Sigma}^0_\alpha(\mathcal{N})(f^{-1}(A) \in \mathbf{\Sigma}^0_\beta(\mathcal{N}))$. Now define the QO $\leq_B$ by: $f \leq_B g$ if $\forall \alpha(d_f(\alpha) \leq d_g(\alpha))$. We do not know whether $\leq_B$ is a WQO but we will see this at least for some reasonable sets of partial functions (note that the definition works for the Borel partial functions).

In an attempt to measure the discontinuity of functions, Weihrauch [37, 118] introduced some notions of reducibility for functions on topological spaces. In particular, the following three notions of reducibility between functions $f, g : X \to Y$ on topological spaces were introduced: $f \leq_0 g$ (resp. $f \leq_1 g$, $f \leq_2 g$) iff $f = g \circ H$ for some continuous function $H : X \to X$ (resp. $f = F \circ g \circ H$ for some continuous functions $H : X \to X$ and $F : Y \to Y$, $f(x) = F(x, g(H(x)))$ for some continuous functions $H : X \to X$ and $F : X \times Y \to Y$). In this way we obtain QOs $(Y^X; \leq_i)$, $i \leq 2$, on the set of all functions from $X$ to $Y$.

There are many variations of Weihrauch reducibilities (in particular, for multivalued functions or with computable reducing functions in place of continuous ones), some of which turned out to be useful for understanding the non-computability and non-continuity of interesting decision problems in computable analysis [14, 37] and constructive mathematics [13]. This research is closely related to reverse mathematics where similar problems (including intersting problems about WQOs) are treated by proof-theoretic means [39]. This is now a popular research field, but its relation to WQO-theory is not clear. Nevertheless, some restricted versions of Weihrauch reducibilities are relevant.

The notions of Wehrauch reducibilities are nontrivial even for the case of discrete spaces $Y = k = \{0, \ldots, k-1\}$ with $k < \omega$ points, which brings us back to the $k$-partitions of $X$. In this case, $\leq_0$ coincides with the Wadge reducibility of $k$-partitions already discussed above (for this reason in the next subsection we reserve the notation $\leq_0$ for the QO from the previous subsection). We again concentrate on the case $X = \mathcal{N}$ because for the non-zero dimensional spaces the degree structures are often complicated. Let $\mathcal{B}_k$ be the subset of $k^{\mathcal{N}}$ formed by the $k$-patitions whose components are finite Boolean combinations of open sets. In [34] Hertling proved the following combinatorial description of small fragments of Weihrauch degrees:

**Theorem 5.4** *The quotient-posets of* $(\mathcal{B}_k; \leq_i)$ *and of* $(\mathcal{F}_k; \leq_i)$ *are isomorphic for each* $i = 1, 2$.

For the functions with infinite range, some interesting results about the relation $\leq_1$ on partial functions on the Baire space were obtained in [15]. The relation $\leq_2$ has the disadvantage that it does not refine the above-defined relation $\leq_B$. Some natural easy properties of $\leq_1$ are collected in the next assertion:

**Proposition 5.5** *(1) For every* $A, B \subseteq \mathcal{N}$, $A \leq_W B$ *implies* $c_A \leq_1 c_B$; $c_A \leq_1 c_B$ *implies that* $A \leq_W B$ *or* $\overline{A} \leq_W B$; *and* $c_A \equiv_1 c_{\overline{A}}$, *where* $c_A(p) = 1^\omega$ *for* $p \in A$ *and* $c_A(p) = 0^\omega$ *otherwise.*
*(2) For every* $A, B \subseteq \mathcal{N}$, $id_A \leq_1 id_B$ *iff* $A$ *is a retract of* $B$, *where* $id_A$ *is the identity on* $A$.
*(3) For every pair of Borel partial functions* $f, g$ *on* $\mathcal{N}$, $f \leq_1 g$ *implies* $f \leq_B g$.

Let $C_\infty$ (resp. $C$, $C^*$) be the set of partial continuous functions on $\mathcal{N}$ with closed domain (resp. with closed domain and countable range, resp. with compact domain). In [15] the notion of Cantor-Bendixon rank of a function in $C$ was defined in such a way that $CB(id_F)$ coincides with the classical Cantor-Bendixon rank $CB(F)$ of a closed set $F$. Let $C_\alpha$ be the set of functions with Cantor-Bendixon rank $\alpha$. Although it is open whether $\leq_1$ is WQO on $C$, the following theorem (that collects some results in [15]) shows that it is WQO on some subsets of $C$.

**Theorem 5.6** *(1) The quotient-poset of* $(C^* \leq_1)$ *is a well order of rank* $\omega_1$.
*(2) Any two functions from* $C_\infty$ *with uncountable range are equivalent w.r.t.* $\equiv_1$.
*(3) Let* $Q$ *be a subset of* $C_\infty$ *such that* $(Q \cap C_\alpha; \leq_1)$ *is BQO for all* $\alpha < \omega_1$. *Then* $(Q; \leq_1)$ *is BQO.*
*(4) The relation of being a retract is WQO on* $\Pi^0_1(\mathcal{N})$.

### *5.4 Definability and Decidability Issues*

One may ask why we are interested in characterizing the degrees of $k$-partitions and other degree structures above in terms of "combinatorial" objects like labeled forests. First, any such non-obvious characterization gives new information about the structures under investigation. Second, the combinatorial objects are much easier to handle and explore than the degree structures. In particular, the definability and decidability issues are much easier to study for the labeled forests than for the Wadge degrees. Such issues (e.g. characterizing the automorphism groups or the complexity of first-order theories) are important and principal for understanding the natural degree structures in DST, though they remain much unexplored, to our knowledge. In contrast, this topic for the natural degree structures in computability theory is a central theme.

From the results in Sects. 4 and 3.4 we immediately obtain many corollaries on definability and decidability in the initial segments of the Wadge degrees of $k$-partitions. In particular, Theorem 3.10 implies the following:

**Theorem 5.7** *1. Let $k \geq 3$, $Q$ be any set of Borel $k$-partitions of $\mathcal{N}$ that contains the set $\mathcal{B}_k$ from the previous subsection, and let $\mathbb{A}$ be the quotient-poset of $(Q; \leq_W)$. Then $FO(\mathbb{N}) \leq_m FO(\mathbb{A})$.*
*2. Let $\mathbb{A}$ be the quotient-poset of $(\mathcal{B}_k; \leq_W)$. Then $FO(\mathbb{N}) \equiv_m FO(\mathbb{A})$ and $(\mathbb{A}; \mathbf{0}, \ldots, \mathbf{k-1})$ is biinterpretable with $\mathbb{N}$.*

In computability theory, people actively discuss several versions of the so called biinterpretability conjecture stating that some structures of degrees of unsolvability are biinterpretable (in parameters) with $\mathbb{N}$ (see e.g. [75] and references therein). The conjecture (which seems still open for the most important cases) is considered as in a sense the best possible definability result about degree structures. Item (2) above solves a similar question (even without parameters) for a natural object of DST.

The results at the end of Sects. 5.2 and 3.4 imply the following result from [99]:

**Proposition 5.8** *Let $k \geq 3$ and let $\mathbb{A}$ be the quotient-poset of any initial segment of $(ER(\mathcal{N}); \leq_c)$ that contains the set of equivalence relations with components in $\mathcal{B}_k$. Then the first-order theory of $\mathbb{A}$ is hereditarily undecidable.*

A similar undecidability result for the segments of Weihruch's degrees follows from the results in Sect. 5.3 and Theorem 3.11. Another natural and important question is to study definability and decidability in the fragments of the quotient-poset of $(CBER; \leq_B)$. Unfortunately, we are not aware of any result in this direction.

## 6 Hierarchies in Quasi-Polish Spaces

In this section we discuss hierarchies of sets and $k$-partitions in quasi-Polish spaces mentioned in Sect. 4.3. In particular, we provide set-theoretic descriptions of these hierarchies, which give some new information even for the classical case of the Wadge hierarchy in Baire space.

### *6.1 Hierarchies of Sets*

Here we briefly discuss some notions of the hierarchy theory mentioned in the Introduction which provide a convenient language to discuss various concrete hierarchies. The next definition of a hierarchy of sets was first proposed by Addison [5].

**Definition 6.1** Let $X$ be a set and $\eta$ be an ordinal.

(1) An *$\eta$-hierarchy of sets in $X$* is a sequence $\{H_\alpha\}_{\alpha<\eta}$ of subsets of $P(X)$ such that $H_\alpha \subseteq H_\beta \cap co\text{-}H_\beta$ for all $\alpha < \beta < \eta$, where $co\text{-}H_\beta$ is the class of complements of sets in $H_\beta$.
(2) The classes $H_\alpha$ and $co\text{-}H_\alpha \setminus H_\alpha$ are *non-self-dual levels of* $\{H_\alpha\}$, while the classes $(H_\alpha \cap co\text{-}H_\alpha) \setminus (\bigcup_{\beta<\alpha} H_\beta \cap co\text{-}H_\beta)$ are *self-dual levels of* $\{H_\alpha\}$.
(3) The classes $H_\alpha \setminus co\text{-}H_\alpha, co\text{-}H_\alpha \setminus H_\alpha$, and $(H_\alpha \cap co\text{-}H_\alpha) \setminus (\bigcup_{\beta<\alpha} H_\beta \cap co\text{-}H_\beta)$ are the *components of* $\{H_\alpha\}$.
(4) $\{H_\alpha\}$ *does not collapse* if $H_\alpha \nsubseteq co\text{-}H_\alpha$ for all $\alpha < \eta$.
(5) $\{H_\alpha\}$ is *non-trivial* if $H_\alpha \nsubseteq co\text{-}H_\alpha$ for some $\alpha < \eta$.
(6) $\{H_\alpha\}$ *fits* a QO $\leq$ on $\bigcup_\alpha H_\alpha$ if every non-self-dual level is downward closed and has a largest (up to equivalence) element w.r.t. $\leq$.

Further definitions in this subsection were suggested in [87, 88]. The next one introduces some relations between hierarchies.

**Definition 6.2** Let $\{H_\alpha\}$ and $\{G_\beta\}$ be hierarchies of sets in $X$.

(1) $\{H_\alpha\}$ is a *refinement of* $\{G_\beta\}$ *in a level* $\beta$ if $\bigcup_{\gamma<\beta}(G_\gamma \cup co\text{-}G_\gamma) \subseteq \bigcup_\alpha H_\alpha \subseteq (G_\beta \cap co\text{-}G_\beta)$. Such a refinement is called exhaustive if $\bigcup_\alpha H_\alpha = G_\beta \cap co\text{-}G_\beta$.
(2) $\{H_\alpha\}$ is a (global) *refinement of* $\{G_\beta\}$ if for any $\beta$ there is an $\alpha$ with $H_\alpha = G_\beta$, and $\bigcup_\alpha H_\alpha = \bigcup_\beta G_\beta$.
(3) A hierarchy is *discrete in a given level* if it has no non-trivial refinements in this level. A hierarchy is (globally) discrete if it is discrete in each level.
(4) $\{H_\alpha\}$ *is an extension of* $\{G_\beta\}$ if the sequence $\{G_\beta\}$ is an initial segment of the sequence $\{H_\alpha\}$.
(5) $\{H_\alpha\}$ is *perfect in a level* $\beta$ if $\bigcup_{\gamma<\beta}(H_\gamma \cup co\text{-}H_\gamma) = H_\beta \cap co\text{-}H_\beta$. A hierarchy is (globally) *perfect* if it is perfect in all levels.

For instance, the transfinite Borel hierarchy is an extension of the finite Borel hierarchy, the Borel hierarchy is an exhaustive refinement of the Luzin hierarchy in the first level (the Suslin theorem), the Hausdorff hierarchy over any non-zero level of the Borel hierarchy is an exhaustive refinement of the Borel hierarchy in the next level (the Hausdorff–Kuratowski theorem), all the classical hierarchies in Baire space fit the Wadge reducibility, the Wadge hierarchy is perfect.

Next we discuss a technical notion of a base hierarchy (or just a base). By *$\alpha$-base in $X$* (or just a base) we mean an $\alpha$-hierarchy of sets $\mathcal{L} = \{\mathcal{L}_\beta\}_{\beta<\alpha}$ in $X$ such that each level $\mathcal{L}_\beta$ is a lattice of sets containing $\emptyset, X$ as elements. The 1-base $(\mathcal{L})$ is identified with $\mathcal{L}$. Note that any $(n+1)$-base $(\mathcal{L}_0, \ldots, \mathcal{L}_n)$ may be extended to the $\omega$-base $\{\mathcal{L}_k\}_{k<\omega}$ (or even to a longer base) by setting $\mathcal{L}_k$ to be the class of Boolean combinations of sets in $\mathcal{L}_n$ for all $k > n$. In the sequel we deal mostly with 1-bases, 2-bases, $\omega$-bases and $\omega_1$-bases. A $\sigma$-base is a base every level of which is an upper $\sigma$-semilattice.

Let $\mathcal{L}$ be a 1-base. Using the difference operators $D_\alpha$ from Sect. 2.4, we can define the finitary version $\{D_n(\mathcal{L})\}_{n<\omega}$ of the difference hierarchy over $\mathcal{L}$ with the usual inclusions of levels. If $\mathcal{L}$ is a $\sigma$-base, the infinitary version $\{\mathcal{L}_\alpha\}_{\alpha<\omega_1}$ will also have the usual properties. In particular, if $\mathcal{L}$ is an $\omega$-$\sigma$-base (i.e., a $\sigma$-base which is an $\omega$-hierarchy), we have the inclusion $\bigcup_{\alpha<\omega_1} D_\alpha(\mathcal{L}_n) \subseteq \mathcal{L}_{n+1} \cap co\text{-}\mathcal{L}_{n+1}$ for each $n < \omega$.

Thus, we have a natural refinement of any $\omega$-base $\mathcal{L}$ in any non-zero level. Take for simplicity only the finitary refinements. We can continue the refinement process by adjoining new refinements in any non-discrete level obtained so far. This refinement process, studied in detail in [88], ends (after collecting all the resulting levels together) with the *fine hierarchy over* $\mathcal{L}$ of length $\varepsilon_0$. This hierarchy is a finitary abstract version of the Wadge hierarchy. Several such hierarchies over concrete bases have interesting properties [88, 95]. We do not give all details here because in the next subsection we discuss a more general case. The properties of such hierarchies strongly depend on some structural properties of classes of sets $\mathcal{A} \subseteq P(X)$ (see e.g. [65]).

**Definition 6.3** (1) The class $\mathcal{A}$ has the *separation property* if for every two disjoint sets $A, B \in \mathcal{A}$ there is a set $C \in \mathcal{A} \cap co\text{-}\mathcal{A}$ with $A \subseteq C \subseteq \overline{B}$.
(2) The class $\mathcal{A}$ has the *reduction property* i.e. for all $C_0, C_1 \in \mathcal{A}$ there are disjoint $C'_0, C'_1 \in \mathcal{A}$ such that $C'_i \subseteq C_i$ for both $i < 2$ and $C_0 \cup C_1 = C'_0 \cup C'_1$. The pair $(C'_0, C'_1)$ is called a reduct for the pair $(C_0, C_1)$.
(3) The class $\mathcal{A}$ has the *$\sigma$-reduction property* if for each countable sequence $C_0, C_1, \ldots$ in $\mathcal{A}$ there is a countable sequence $C'_0, C'_1, \ldots$ in $\mathcal{A}$ (called a reduct of $C_0, C_1, \ldots$) such that $C'_i \cap C'_j = \emptyset$ for all $i \neq j$ and $\bigcup_{i<\omega} C'_i = \bigcup_{i<\omega} C_i$.

It is well-known that if $\mathcal{A}$ has the reduction property then the dual class $co\text{-}\mathcal{A}$ has the separation property, but not vice versa, and that if $\mathcal{A}$ has the $\sigma$-reduction property then $\mathcal{A}$ has the reduction property but not vice versa. Nevertheless, if $\mathcal{A}$ has the reduction property then for any finite sequence $(C_0, \ldots, C_n)$ in $\mathcal{C}$ there is a reduct $C'_0, \ldots, C'_n \in \mathcal{C}$ for $(C_0, \ldots, C_n)$ which is defined similarly to the countable reduct above.

The next properties of an $\omega$-base $\mathcal{L}$ imply good properties of the fine hierarchy over $\mathcal{L}$ (and if $\mathcal{L}$ is a $\omega$-base then also of the infinitary version of the fine hierarchy).

**Definition 6.4** Let $\mathcal{L}$ be an $\omega$-base (resp. $\sigma$-base) in $X$.

(1) $\mathcal{L}$ is *reducible* (resp. *$\sigma$-reducible*) if for every $n$ the level $\mathcal{L}_n$ has the reduction (resp. $\sigma$-reduction) property.
(2) $\mathcal{L}$ is *interpolable* (resp. *$\sigma$-interpolable*) if for each $n < \omega$ the class $co\text{-}\mathcal{L}_{n+1}$ has the separation property and $\mathcal{L}_{n+1} \cap \text{co-}\mathcal{L}_{n+1}$ coincides with the class of Boolean combinations of sets in $\mathcal{L}_n$ (resp. with $\bigcup_{\alpha<\omega_1} D_\alpha(\mathcal{L}_n)$).

## 6.2 *Hierarchies of k-Partitions*

Here we extend hierarchies of sets from the previous subsection to hierarchies of $k$-partitions.

The discussion in Sect. 4 suggests that the levels of hierarchies of $k$-partitions should be ordered under inclusion in a more complicated way than for the hierarchies of sets. Accordingly, it is not natural to use ordinals to notate the levels of such hierarchies, as in Definition 6.1. In [98] we looked at the most general case of hierarchies whose levels are notated by an *arbitrary* poset $P$, let us call them $P$-hierarchies. Simple considerations show that only in the case where $P$ is WPO do we have reasonable behaviour in the components of a $P$-hierarchy (they should at least partition the sets covered by the $P$-hierarchy). Accordingly, we stick to hierarchies named by WQOs. For such hierarchies, reasonable extensions of Definitions 6.1 and 6.2 were suggested in [98, 100].

Our experience with classifying $k$-partitions suggests to use the QO $\mathcal{T}_k(\omega)$ and its initial segments as notation system for the finitary versions of hierarchies of $k$-partitions, and the QO $\widetilde{\mathcal{T}}_k(\omega_1)$ and its initial segments as notation system for the infinitary versions.

First we consider the infinitary version of the difference hierarchy (DH) of $k$-partitions (the finitary version is obtained by sticking to the finite trees). The notation system for this hierarchy is $\widetilde{\mathcal{T}}_k$ (for the finitary version — $\mathcal{T}_k$). Let $\mathcal{L}$ be a 1-base in $X$ which is a $\sigma$-base. Let $(T; c) \in \widetilde{\mathcal{T}}_k$, so $T \subseteq \omega^*$ is a well-founded tree and $c : T \to k$.

We say that a $k$-partition $A : X \to k$ is *defined by a $T$-family $\{B_\tau\}_{\tau \in T}$ of $\mathcal{L}$-sets* if $A_i = \bigcup_{\tau \in T_i} \tilde{B}_\tau$ for each $i < k$ where $\tilde{B}_\tau = B_\tau \setminus \bigcup\{B_\sigma \mid \sigma \sqsupset \tau\}$ and $T_i = c^{-1}(i)$. Note that we automatically have $\bigcup_\tau B_\tau = X$. This definition is especially clear for the case when the family $\{B_\tau\}_{\tau \in T}$ is reduced (i.e., $B_{\tau i} \cap B_{\tau j} = \emptyset$ for all distinct $i, j$ with $\tau i, \tau j \in T$) because then $\{\tilde{B}_\tau\}_{\tau \in T}$ is a partition of $X$. Note that any reduced family with $\bigcup_\tau B_\tau = X$ defines a $k$-partition but this fails for the general families.

By the *difference hierarchy of $k$-partitions over $\mathcal{L}$* we mean the family $\{\mathcal{L}(T)\}_{T \in \widetilde{\mathcal{T}}_k}$ were $\mathcal{L}(T)$ is the set of $k$-partitions defined by $T$-families of $\mathcal{L}$-sets. We mention some properties from [100] (see also [98] where the finitary version is considered).

**Proposition 6.5** *(1) If $\mathcal{L}$ has the $\sigma$-reduction property then any level $\mathcal{L}(T)$ of the DH coincides with the set of $k$-partitions defined by the reduced $T$-families of $\mathcal{L}$-sets.*

*(2) If $T \leq_h S$ then $\mathcal{L}(T) \subseteq \mathcal{L}(S)$.*

*(3) Let $f$ be a function on $X$ such that $f^{-1}(A) \in \mathcal{L}$ for each $A \in \mathcal{L}$. Then $A \in \mathcal{L}(T)$ implies $A \circ f \in \mathcal{L}(T)$.*

*(4) The DH of 2-partitions over $\mathcal{L}$ coincides with the DH of sets over $\mathcal{L}$, in particular, $\{\mathcal{L}(T) \mid T \in \widetilde{\mathcal{T}}_2\} = \{D_\alpha(\mathcal{L}), co\text{-}D_\alpha(\mathcal{L}) \mid \alpha < \omega_1\}$.*

If $\mathcal{L}$ is an $\omega$-base in $X$ which is a $\sigma$-base, the considerations above provide some basic information on the DHs of $k$-partitions over any level of $\mathcal{L}$. Obviously, for all $n$ and $T \in \widetilde{\mathcal{T}}_k$ the level $\mathcal{L}(T)$ is contained in the set of $k$-partitions with the components in $\mathcal{L}_{n+1} \cap co - \mathcal{L}_{n+1}$.

Now we extend these DHs of $k$-partitions to the fine hierarchy (FH) of $k$-partitions. Its levels are notated by the iterated $h$-QOs. In [100] we used the initial segment $\widetilde{\mathcal{T}}_k(\omega)$ as the notation system (for the finitary case $\mathcal{T}_k(\omega)$ was used in [98]). Simplifying notation, we stick to the initial segment $\widetilde{\mathcal{T}}_k(2)$, this also means sticking to the 2-base $(\mathcal{L}_0, \mathcal{L}_1)$.

Let $(T; c) \in \widetilde{\mathcal{T}}_k(2)$, so $T \subseteq \omega^*$ is a well-founded tree and $c : T \to \widetilde{\mathcal{T}}_k$, hence for any $\tau \in T$ we have a tree $(S^\tau; c^\tau) = c(T)$ with $c^\tau : S^\tau \to k$. We say that a $k$-partition $A : X \to k$ is *defined by families* $\{B_\tau\}_{\tau \in T}$ and $\{C^\tau_\sigma\}_{\tau \in T, \sigma \in S^\tau}$ *of* $\mathcal{L}$ *-sets* if $B_\tau \in \mathcal{L}_0$, $C^\tau_\sigma \in \mathcal{L}_1$, and $A_i = \bigcup\{\tilde{B}_\tau \cap \tilde{C}^\tau_\sigma \mid \tau \in T, \sigma \in S^\tau_i\}$ for each $i < k$. Again, the definition is much clearer for the case when the families $\{B_\tau\}$ and $\{C^\tau_\sigma\}$ are reduced (for the second family this means $C^\tau_{\sigma i} \cap C^\tau_{\sigma j} = \emptyset$ for all distinct $i, j$ with $\sigma i, \sigma j \in S^\tau$, for each $\tau \in T$) because then $\{\tilde{B}_\tau\}_{\tau \in T}$ is a partition of $X$ and $\{\tilde{B}_\tau \cap \tilde{C}^\tau_\sigma\}_{\sigma \in S^\tau}$ is a partition of $\tilde{B}_\tau$ for each $\tau \in T$. Note that any such reduced family with $\bigcup_\tau B_\tau = X$ defines a $k$-partition but this fails for the general families.

By the *fine hierarchy of k-partitions over* $(\mathcal{L}_0, \mathcal{L}_1)$ we mean the family $\{\mathcal{L}(T)\}_{T \in \widetilde{\mathcal{T}}_k(2)}$ were $\mathcal{L}(T)$ is the set of $k$-partitions defined by $T$-families of $\mathcal{L}$-sets. We mention some properties from [100] (see also [98] where the finitary version is considered). Note that these definitions show that the FH of $k$-partitions is in a sense an iterated version of the DHs.

**Proposition 6.6** *(1) If $\mathcal{L}$ is $\sigma$-reducible then any level $\mathcal{L}(T)$ of the FH coincides with the set of k-partitions defined by the reduced T-families of $\mathcal{L}$-sets.*

*(2) If $T \leq_h S$ then $\mathcal{L}(T) \subseteq \mathcal{L}(S)$.*

*(3) Let $f$ be a function on $X$ such that $f^{-1}(A) \in \mathcal{L}$ for each $A \in \mathcal{L}$. Then $A \in \mathcal{L}(T)$ implies $A \circ f \in \mathcal{L}(T)$.*

*(4) The FH of 2-partitions over $\mathcal{L}$ coincides with the FH of sets over $\mathcal{L}$, in particular, $\{\mathcal{L}(T) \mid T \in \widetilde{\mathcal{T}}_2(2)\}$ coincides with the set of levels $\alpha < \omega_1^{\omega_1}$ of the FH of sets.*

## *6.3 Hierarchies of Sets in Quasi-Polish Spaces*

As we noticed in Sect. 4.3 it is not straightforward to extend Wadge hierarchy to quasi-Polish spaces using the Wadge reducibility in those spaces.

A natural way to do this is to apply the refinement process explained in Sect. 6.1 to the $\omega_1$-base $\mathcal{L}_X = \{\mathbf{\Sigma}^0_{1+\alpha}(X)\}$ in arbitrary quasi-Polish space $X$. Note that this base is interpolable. It is $\sigma$-reducible when $X$ is zero-dimensional but it is not reducible in general because the level $\mathbf{\Sigma}^0_1(X)$ often does not have the reduction property (though the higher levels always have the $\sigma$-reduction property [99]).

At the first step of the process we obtain the classical Hausdorff hierarchies over each level. Further refinements may be done by the construction from the end of the previous subsection for $k = 2$. In [100] this was done for the initial segment $T \in \widetilde{\mathcal{T}}_2(\omega)$ which is semi-linear-ordered with order type $\lambda = sup\{\omega_1, \omega_1^{\omega_1}, \omega_1^{(\omega_1^{\omega_1})}, \ldots\}$. By choosing the trees $T$ with the root label 0, we obtain the increasing sequence of pointclasses $\{\mathbf{\Sigma}_\alpha(X)\}_{\alpha < \lambda}$ which is a good candidate to be the (initial segment of) Wadge hierarchy in $X$.

From results in [67, 116] it is not hard to deduce that for $X = \mathcal{N}$ these classes coincide with the corresponding levels of Wadge hierarchy in Sect. 2.5. This gives an alternative (to those in [67, 116]) set-theoretical characterization of these levels. There is no doubt that this definition maybe extended to $T \in \widetilde{\mathcal{T}}_2(\omega_1)$ (yielding a set-theoretical characterization of the levels of Wadge hierarchy within $\Delta^{\mathbf{0}}_{!}$) and to the higher levels.

Some nice properties of the introduced classes may be obtained using the admissible representations from the end of Sect. 2.3. Let $\delta$ be a total admissible representation of the quasi-Polish space $X$. According to the results in [17, 105], $A \in \mathbf{\Sigma}^{-1,\theta}_{\alpha}(X)$ iff $\delta^{-1}(A) \in \mathbf{\Sigma}^{-1,\theta}_{\alpha}(\mathcal{N})$ for all $\alpha, \theta < \omega_1$, $\theta \geq 1$ (in particular, $A \in \mathbf{\Sigma}^0_\alpha(X)$ iff $\delta^{-1}(A) \in \mathbf{\Sigma}^0_\alpha(\mathcal{N})$ for all $1 \leq \alpha < \omega_1$). In [100] this was extended to the fact that $A \in \mathbf{\Sigma}_\alpha(X)$ iff $\delta^{-1}(A) \in \mathbf{\Sigma}_\alpha(\mathcal{N})$, for all $\alpha < \lambda$. We do believe that this extends to the whole Wadge hierarchy.

As suggested independently in [77, 100], one can also *define* the Wadge hierarchy $\{\mathbf{\Sigma}_\alpha(X)\}_{\alpha<\nu}$ in $X$ by $\mathbf{\Sigma}_\alpha(X) = \{A \subseteq X \mid \delta^{-1}(A) \in \mathbf{\Sigma}_\alpha(\mathcal{N})\}$. One easily checks that the definition of $\mathbf{\Sigma}_\alpha(X)$ does not depend on the choice of $\delta$, $\bigcup_{\alpha<\nu} \mathbf{\Sigma}_\alpha(X) = \mathbf{B}(\mathbf{X})$, $\mathbf{\Sigma}_\alpha(X) \subseteq \mathbf{\Delta}_\beta(X)$ for all $\alpha < \beta < \nu$, and any $\mathbf{\Sigma}_\alpha(X)$ is downward closed under the Wadge reducibility on $X$. This definition is short but gives no real understanding of how the levels look like. Hence, also in this approach the set-theoretic characterization of levels is of principal interest.

### *6.4 Hierarchies of k-Partitions in Quasi-Polish Spaces*

Here we extend the hierarchies of the previous subsection to $k$-partitions. Applying the general definitions of Sect. 6.2 to the $\omega_1$-base $\mathcal{L}_X = \{\mathbf{\Sigma}^0_{1+\alpha}(X)\}$ in a quasi-Polish space $X$ we obain the DHs of $k$-partitions $\{\mathcal{L}_\alpha(X, T)\}_{T \in \widetilde{\mathcal{T}}_k}$, $\alpha < \omega_1$, and the FH of $k$-partitions $\{\mathcal{L}(X, T)\}_{T \in \widetilde{\mathcal{T}}_k(2)}$.

As follows from [100], the results from the previous subsection about the FH extend to $k$-partitions in the following sense: $A \in \mathcal{L}(X, T)$ iff $A \circ \delta \in \mathcal{L}(\mathcal{N}, T)$, for each $T \in \widetilde{\mathcal{T}}_k(2)$. Similarly, for the DHs we have: $A \in \mathcal{L}_\alpha(X, T)$ iff $A \circ \delta \in \mathcal{L}_\alpha(\mathcal{N}, T)$, for all $A \subseteq X$, $\alpha < \omega_1$, and $T \in \widetilde{\mathcal{T}}_k$.

The FH of $k$-partitions $\{\mathcal{L}(\mathcal{N}, T)\}_{T \in \widetilde{\mathcal{T}}_k(2)}$ is related to the Wadge reducibilty of $k$-partitions as follows. As we know from Sect. 4.2, there is a function $\mu : \widetilde{\mathcal{T}}_k(2) \to \Delta^0_3(k^{\mathcal{N}})$ inducing an isomorphism between the quotient-posets of the $\sigma$-join-irreducible elements in $(\mathbf{\Delta}^0_3(k^{\mathcal{N}}); \leq_W)$ and of $(\widetilde{\mathcal{T}}_k(2); \leq_h)$. For any $T \in \widetilde{\mathcal{T}}_k(2)$, we have $\mathcal{L}(\mathcal{N}, T) = \{A \in \mathbf{\Delta}^0_3(k^{\mathcal{N}}) \mid A \leq_W \mu(F)\}$.

## 7 Hierarchies in Computability Theory

In this section we briefly discuss some hierarchies and reducibilities in computability theory. They are important because they provide tools for classifying many interesting decision problems in logic and theoretical computer science. The idea to use reducibilities as a classification tool first appeared in computability theory and later it was borrowed by many other fields.

## *7.1 Preliminaries*

We assume that the reader is familiar with the main notions of computability theory and simply recall some notation and not broadly known definitions. For more details the reader may use any of the many available books on the subject, e.g. [79, 83].

If not specified otherwise, all functions are assumed in this section to be functions on $\omega$, and all sets to be subsets of $\omega$. Thus, for an $n$-ary partial function $\phi$, we have $dom(\phi) \subseteq \omega^n$ and $rng(\phi) \subseteq \omega$. Instead of $(x_1, \ldots, x_n) \in dom(\phi)$ ($(x_1, \ldots, x_n) \notin dom(\phi)$) we sometimes write $\phi(x_1, \ldots, x_n) \downarrow$ (respectively, $\phi(x_1, \ldots, x_n) \uparrow$). We assume the reader to be familiar with the computable partial (c.p.) functions, computable (total) functions and computably enumerable (c.e.) sets. For any $n > 1$, there is a computable bijection $\lambda x_1, \ldots, x_n.\langle x_1, \ldots, x_n\rangle$ (the Cantor coding function) between $\omega^n$ and $\omega$. This fact reduces many considerations to the case of unary functions and predicates.

By a *numbering* we mean any function $\nu$ with $dom(\nu) = \omega$, and by *numbering of a set* $S$ — any numbering $\nu$ with $rng(\nu) = S$. A numbering $\mu$ is *reducible* to a numbering $\nu$ (in symbols $\mu \leq \nu$), if $\mu = \nu \circ f$ for some computable function $f$, and $\mu$ is *equivalent* to $\nu$ ($\mu \equiv \nu$), if $\mu \leq \nu$ and $\nu \leq \mu$. Relate to any numberings $\mu, \nu$ and to any sequence of numberings $\{\nu_k\}_{k<\omega}$ the numberings $\mu \oplus \nu$, and $\bigoplus_k \nu_k$, called respectively the *join of* $\mu$ *and* $\nu$ and the *infinite join of* $\nu_k (k < \omega)$ defined as follows:

$$(\mu \oplus \nu)(2n) = \mu n,\ (\mu \oplus \nu)(2n+1) = \nu n,\ (\bigoplus_k \nu_k)\langle x, y\rangle = \nu_x(y).$$

Let $\{\varphi_n\}$ be the standard numbering of (unary) c.p. functions on $\omega$. We assume the reader to be acquainted with the computations relative to a given set $A \subseteq \omega$ or a function $\xi \in \omega^\omega$ (which in this situation are often called *oracles*). E.g., such computations may be formally defined using Turing machines with oracles. Enumerating all programs for such machines we obtain numberings $\varphi^A$ ($\varphi^\xi$) of all partial functions computable in $A$ (in $\xi$).

A numbering $\nu : \omega \to S$ is *complete w.r.t.* $a \in S$ if for every c.p. function $\psi$ on $\omega$ there is a total computable function $g$ such that $\nu(g(x)) = \nu(\psi(x))$ in case $\psi(x) \downarrow$ and $\nu(g(x)) = a$ otherwise. For any set $S$ and any $a \in S$, define a unary operation $p_a$ on $S^\omega$ as follows: $[p_a(\nu)]n = a$ for $\upsilon(n) \uparrow$ and $[p_a(\nu)]n = \nu\upsilon(n)$ for $\upsilon(n) \downarrow$, where $\upsilon$ is the universal p.c. function $\upsilon(\langle n, x\rangle) = \varphi_n(x)$. These *completion operations* were introduced in [84] as a variant of similar operations from [28]. They are very relevant to fine hierarchies, as the following particular case of results in [84] demonstrates. The notions related to completeness are relativized to a given oracle $A$ in the obvious way.

**Theorem 7.1** *For every* $2 \leq k < \omega$, $(k^\omega; \leq, \oplus, p_0, \ldots, p_{k-1})$ *is a dc-semilattice.*

In case $k = 2$ the QO $(k^\omega; \leq)$ coincides with the QO $(P(\omega); \leq_m)$, where $\leq_m$ is the $m$-reducibility, which is a popular structure of computability theory. Another important QO on $P(\omega)$ is the Turing reducibility. Recall that $A$ is *Turing-reducible*

($T$-reducible) to $B$ (in symbols, $A \leq_T B$) if $A$ is computable in $B$, i.e. $A = \varphi_n^B$ for some $n$. The *Turing jump operator* $A \mapsto A'$ on $P(\omega)$ is defined by $A' = \{n \mid \varphi_n^A(n) \downarrow\}$. For any $n < \omega$, define the $n$-th jump $A^{(n)}$ of $A$ by $A^{(0)} = A$ and $A^{(n+1)} = (A^{(n)})'$.

The arithmetical hierarchy $\mathcal{L} = \{\Sigma^0_{n+1}\}_{n<\omega}$ is an $\omega$-base in $\omega$. It is reducible but not interpolable, it does not collapse and fits the $m$-reducibility. Moreover, $\emptyset^{(n+1)}$ is $m$-complete in $\Sigma^0_{n+1}$ for each $n < \omega$ (see [79]). In the next two subsections we consider the refining process for this base, concentrating for simplicity on the finitary hierarchies. Let $\leq^A$ (resp. $\leq^n$) denote the reducibility of numberings by functions computable in $A$ (resp. in $\emptyset^{(n)}$).

## 7.2 Difference Hierarchies of Sets and k-Partitions

Let $\{\Sigma^{-1,n}_m\}_{m<\omega}$ be the DH over $\Sigma^0_{n+1}$. These hierarchies and their transfinite extensions over the Kleene ordinal notation system where thoroughly investigated in [28, 85]. We recall here only their characterizations in terms of suitable jump operators. Since we will consider several such operators, we give a general notion [87].

By a *jump operator* we mean a unary operation $J$ on $P(\omega)$ such that $A \oplus \overline{A} \leq_m J(A)$ and $J(A)$ is a complete numbering w.r.t. 0 uniformly in $A$ (uniformity in, say, second condition means the existence of a computable sequence $\{g_e\}$ of total computable functions such that, for all $A, e, x$ we have: $\nu(g_e(x)) = \nu(\varphi_e^A(x))$ in case $\varphi_e(x) \downarrow$ and $\nu(g(x)) = 0$ otherwise, where $\nu = J(A) : \omega \to \{0, 1\}$). From the properties of complete numberings it follows that actually we have $A \oplus \overline{A} <_m J(A)$. It is clear that for any set $A$ complete w.r.t. 0 the sequence $\{J^n(A)\}$ of iterates of $J$ starting from the set $A$ is strictly increasing w.r.t. the $m$-reducibility, and the corresponding principal ideals form an $\omega$-hierarchy denoted as $(J, A)$.

As mentioned above, if $TJ$ is the Turing jump then $(TJ, \emptyset)$ is the arithmetical hierarchy. As observed in [28], the operation $mJ(A) = \upsilon^{-1}(A \oplus \overline{A})$, where $\upsilon\langle n, x\rangle = \varphi_n(x)$ is the universal c.p. function, is a jump operator called the *m-jump*. By [28], $(mJ, \emptyset)$ coincides with the Ershov's hierarchy $\{\Sigma^{-1,0}_m\}_{m<\omega}$. Similarly, if we take in place of $mJ$ the $m$-jump relativized to $\emptyset^{(n)}$, for each $n < \omega$, we obtain the DH $\{\Sigma^{-1,n}_m\}_{m<\omega}$ over $\Sigma^0_{n+1}$ [85].

Now consider the DH $\{\Sigma^0_{n+1}(T)\}_{T\in\mathcal{T}_k}$ of $k$-partitions over $\Sigma^0_{n+1}$. This hierarchy can also be characterized in terms of natural operations on $k^\omega$. Namely, let $p^n_i, i < k$, be the relativization of the operation $p_i$ from the end of the previous subsection to $\emptyset^{(n)}$. Define the function $\mu : \mathcal{T}_k \to k^\omega$ just as in Sect. 4.2, only for the finite forests and the finitary joins of $k$-partitions. Let $\mu^n$ be defined similarly but with $p^n_i$ instead of $p_i$. Then from (the relativization of) Theorem 7.1 we obtain:

**Theorem 7.2** *For all $2 \leq k < \omega$ and $n < \omega$, the function $\mu^n$ induces an embedding of the quotient-poset of $(\mathcal{F}_k; \leq_h)$ into that of $(k^\omega; \leq^n)$ (as well of their functional expansions to the signature $\{\oplus, p_0, \ldots, p_{k-1}\}$). Moreover, $\Sigma^0_{n+1}(T) = \{A \mid A \leq \mu^n(T)\} = \{A \mid A \leq^n \mu^n(T)\}$ for each $T \in \mathcal{T}_k$.*

### *7.3 Fine Hierarchies of Sets and k-Partitions*

By the preceding subsection, the DHs over $\mathcal{L} = \{\Sigma^0_{n+1}\}_{n<\omega}$ may be characterized in terms of suitable jump operations. Is there a similar characterization for the FHs? The answer is positive, and actually the FH of sets was first discovered in [85] in this way.

Since the jump-characterization is non-trivial and yields additional information on the FH, we provide some details. Which jump operations to use? Of course, at least the $m$-jumps $J^n_m$ relativized to $\emptyset^{(n)}$, for all $n < \omega$. By the preceding subsection, $(J^n_m, \emptyset)$ is the difference hierarchy over $\Sigma^0_{n+1}$. A wider class of $\omega$-hierarchies is constructed by considering the sets generated from the empty set by all the operations $J^n_m (n < \omega)$, see [85]. It is not hard to check that in this way we obtain a non-collapsing hierarchy with order type $\omega^\omega$. This already shows that these jump operations do not yield the whole FH but only its small fragment.

In order to find a sufficient class of jump operations, we defined in [85] an operation $r : S^\omega \times S^\omega \times k^\omega \to S^\omega$ (where $S$ is a set and $2 \leq k < \omega$) that includes the jump operations above, the Turing jump and many others. We set $r(\mu, \nu, f) = \bigoplus_n p^f_{\nu(n)}(\mu)$. Then $r(\mu, \lambda x.a, f) \equiv p^f_a(\mu)$ for all $a \in S$, hence $r$ generalizes the operations of completion from Sect. 7.1. Note that for $S = k = 2$ the operation $r$ is a ternary operation on sets satisfying $r(\omega, \emptyset, A) \equiv A'$, hence $r$ generalizes also the Turing jump. It induces also several other jump operators. Namely, for any sets $B$ and $C$, if $B$ is complete w.r.t. 0 then $A \mapsto r(A \oplus \overline{A}, B, C)$ is a jump operator. This follows from the definition and the property that if $\nu$ is $f$-complete w.r.t. $a$ then so is also the numbering $r(\mu, \nu, f)$. The last property together with other properties of $r$ generalizing the properties of the completion operations were established in [85]. These properties play a central role in the algebraic proof of the result below that classifies elements of the subalgebra generated by the operations $r$, $\bar{}$ and $\oplus$ from $\emptyset$ within $2^\omega$. As a corollary, we obtain the jump-characterization of the FH of sets over $\mathcal{L}$. For details and much of additional information see e.g. [88, 98].

We conclude this subsection with the brief discussion of the FH $\{\mathcal{L}(T)\}_{T\in\mathcal{T}_k(\omega)}$ of $k$-partitions over $\mathcal{L}$. We define by induction on $m$ the sequence $\{\rho^n_m\}_n$ of functions $\rho^n_m : T \in \mathcal{T}_k(m+1) \to k^\omega$ as follows. Let $\rho^n_0 := \mu^n$ (where $\mu$ is defined before Theorem 7.2) and suppose by induction that $\rho^n_m$, $n \geq 0$, are defined.

Let $(T; c) \in \mathcal{T}_k(m+2)$, so $c : T \to \mathcal{T}_k(m+1)$. Relate to any node $\sigma \in T$ the $k$-partition $\rho_T(\sigma)$ by induction on the rank $rk(\sigma)$ of $\sigma$ in $(T; \sqsupseteq)$ as follows: if $rk(\sigma) = 0$ then $\rho_T(\sigma) := \rho^{n+1}_m(Q)$ where $Q = c(\sigma) \in \mathcal{T}_k(m+1)$; otherwise, $\rho_T(\sigma) := r(\bigoplus\{\rho_T(\sigma n) \mid n < \omega, \sigma n \in T\}, \rho^{n+1}_m(Q), \emptyset^{(n)})$. We set $\rho^n_{m+1}(T) := \rho_T(\varepsilon)$.

From the properties of $r$ in [85, 88] and Proposition 3.2 it follows by induction on $m$ that $\rho^n_m$ induces an embedding of the quotient-poset of $(\mathcal{T}_k(m+1); \leq_h)$ into that of $(k^\omega; \leq^n)$, and $\rho^n_{m+1}$ extends $\rho^n_m$ modulo $\equiv^n$, Therefore, $\rho^n := \bigcup_m \rho^n_m$ induces an embedding of the quotient-poset of $(\mathcal{T}_k(\omega); \leq_h)$ into that of $(k^\omega; \leq^n)$. Set $\rho := \rho^0$ and extend $\rho$ to $\mathcal{T}^\sqcup_k(\omega)$ by $\rho(F) := \bigoplus\{\rho_T(n) \mid n < \omega, (n) \in T\}$ where $T := \{\varepsilon\} \cup F$.

**Theorem 7.3** *For each $2 \leq k < \omega$, the function $\rho$ induces an embedding of the quotient-poset of $(\mathcal{T}_k^{\sqcup}(\omega); \leq_h)$ into that of $(k^\omega; \leq)$ (as well as of their functional expansions). Moreover, $\mathcal{L}(T) = \{A \mid A \leq \rho(T)\}$ for each $T \in \mathcal{T}_k(\omega)$.*

## *7.4 Natural Degrees*

As is well known, the degree structures in computability theory are extremely rich and complicated, including the structures of many-one and Turing degrees. Some of these degrees (e.g., those obtained by iterating the Turing or the $m$-jumps starting from the empty set) are "natural" in the sense that they are equivalent to a lot of sets appearing in mathematical practice (outside the computability theory). It turns out that in fact only a small number of degrees are "natural" in this sense (e.g., no "natural" non-computable set strictly below $\emptyset'$ under Turing reducibility is known).

A main idea of [85] was to find in the rich structure of the $m$-degrees of (hyper-)arithmetical sets a natural easy substructure that contains $m$-degrees of all sets which appear naturally in mathematics. In this search we tried to expand the structure of $m$-degrees with natural jump operations and then look at the degrees generated from the empty set, as explained above. The result was the discovery of the FH of this section which was later characterized set-theoretically [87, 88] as the abstract finitary version of the Wadge hierarchy. Moreover, it was described [87] how to obtain the Wadge hierarchy from the FH using the uniform relativization and taking the "limit" on the oracles.

In parallel, it was formally proved in [11, 104, 108] (using game-theoretic techniques) that the "natural" Turing degrees are, essentially, the iterates of the Turing jump through the transfinite.

Recently [63], a similar result was achieved for the "natural" $m$-degrees, under a precise notion of "naturalness" based on the uniform relativizations. Namely, a function $f : \mathcal{N} \to P(\omega)$ is *uniformly* $(\leq_T, \leq_m)$*-preserving* if, for every $X, Y \in \mathcal{N}$, $X \leq_T Y$ implies $f(X) \leq_m f(Y)$ uniformly, i.e., there is a computable function $u$ on $\omega$ such that, for all $X, Y \in \mathcal{N}$ and $e \in \omega$, the condition "$X \leq_T Y$ via $e$" implies that $f(X) \leq_m f(Y)$ via $u(e)$. The uniformly $(\equiv_T, \equiv_m)$-preserving functions are defined in the same way. It is easy to see that the $m$-degrees of sets complete in the levels of the FH are natural in the sense of [63].

In [63] it was shown that the degree structure of the uniformly $(\equiv_T, \equiv_m)$-order preserving functions under a natural QO of "many-one reducibility on a cone" is isomorphic to the structure of Wadge degrees. Moreover, the result holds for $Q^\omega$ in place of $2^\omega = P(\omega)$ in the definition above, where $Q$ is an arbitrary BQO. The proof heavily uses a generalization of Theorem 3.3. This is another demonstration of the interplay between computability theory and DST.

# 8 Hierarchies in Automata Theory

In this section we discuss some hierarchies and reducibilities arising in automata theory. Automata theory is an important part of computer science with many deep applications. In fact, many results of this extensive field became already a part of the information technology being realized in most of the existing hardware and software systems. At the same time, automata theory remains an area of active research, with many open problems. The theory is naturally divided in two parts devoted to the study of finite and infinite behavior of computing devices. A positive feature of this field is that many important decision problems concerning deterministic finite automata (dfa's) are decidable. Accordingly, much effort is devoted to finding the optimal decision algorithms and to the complexity issues.

Investigation of the infinite behavior of computing devices is of great interest for computer science because many hardware and software concurrent systems (like processors or operating systems) may not terminate. In many cases, the infinite behavior of a device is captured by the notion of $\omega$-language recognized by the device. The most basic notion of this field is that of regular $\omega$-languages, i.e. $\omega$-languages recognized by finite automata. Regular $\omega$-languages play an important role in the theory and technology of specification and verification of finite state systems.

Regular $\omega$-languages were introduced by J.R. Büchi in the 1960s and studied by many people including B.A. Trakhtenbrot, R. McNaughton and M.O. Rabin. The subject quickly developed into a rich topic with several deep applications. Much information and references on the subject may be found e.g. in [78, 111–113, 117]. We assume the reader to be familiar with the standard notions and facts of automata theory which may be found e.g. in [78, 111].

## *8.1 Preliminaries*

If not stated otherwise, $A$ denotes some finite alphabet with at least two letters. Let $A^*$ and $A^+$ be the sets of finite (respectively, of finite non-empty) words over $A$. Sets of words are called languages. We mainly use the logical approach to the theory of regular languages. This is the reason why we mostly deal with subsets of $A^+$ (they correspond to the non-empty structures, the empty structure is excluded because dealing with it in logic is not usual). With suitable changes analogs of the results below hold also for the subsets of $A^*$.

By an *automaton* (over $A$) we mean a triple $\mathcal{M} = (Q, A, f)$ consisting of a finite non-empty set $Q$ of states, the input alphabet $A$ and a transition function $f : Q \times A \to Q$. The transition function is naturally extended to the function $f : Q \times A^* \to Q$ defined by induction $f(q, \varepsilon) = q$ and $f(q, u \cdot x) = f(f(q, u), x)$, where $\varepsilon$ is the empty word, $u \in A^*$ and $x \in A$. A *word acceptor* is a triple $(\mathcal{M}, i, F)$ consisting of an automaton $\mathcal{M}$, an initial state $i$ of $\mathcal{M}$ and a set of final states $F \subseteq Q$.

Such an acceptor recognizes the language $L(\mathcal{M}, i, F) = \{u \in A^* \mid f(i, u) \in F\}$. Languages recognized by such acceptors are called *regular*.

Relate to any alphabet $A = \{a, \ldots\}$ the signatures $\varrho = \{\leq, Q_a, \ldots\}$ and $\sigma = \{\leq, Q_a, \ldots, \bot, \top, p, s\}$, where $\leq$ is a binary relation symbol, $Q_a$ (for any $a \in A$) is a unary relation symbol, $\bot$ and $\top$ are constant symbols, and $p, s$ are unary function symbols. A word $u = u_0 \ldots u_n \in A^+$ may be considered as a structure $\mathbf{u} = (\{0, \ldots, n\}; <, Q_a, \ldots)$ of signature $\sigma$, where $<$ has its usual meaning, $Q_a(a \in A)$ are unary predicates on $\{0, \ldots, n\}$ defined by $Q_a(i) \leftrightarrow u_i = a$, the symbols $\bot$ and $\top$ denote the least and the greatest elements, while $p$ and $s$ are respectively the predecessor and successor functions on $\{0, \ldots, n\}$ (with $p(0) = 0$ and $s(n) = n$).

For a sentence $\phi$ of $\sigma$, set $L_\phi = \{u \in A^+ \mid \mathbf{u} \models \phi\}$. Sentences $\phi, \psi$ are treated as equivalent when $L_\phi = L_\psi$. A language is *$FO_\sigma$-axiomatizable* if it is of the form $L_\phi$ for some first-order sentence $\phi$ of signature $\sigma$. Similar notions apply to other signatures in place of $\sigma$. It is well-known (see e.g. [78, 107]) that the class of $FO_\sigma$-definable languages (as well as the class of $FO_\varrho$-definable languages) coincides with the important class of *regular aperiodic languages* which are also known as *star-free languages*.

By *initial automaton* (over $A$) we mean a tuple $(Q, A, f, i)$ consisting of a dfa $(Q, A, f)$ and an initial state $i \in Q$. Similarly to the function $f : Q \times A^* \to Q$, we may define the function $f : Q \times A^\omega \to Q^\omega$ by $f(q, \xi)(n) = f(q, \xi \restriction n)$. Relate to any initial automaton $\mathcal{M}$ the set of cycles $C_\mathcal{M} = \{f_\mathcal{M}(\xi) \mid \xi \in A^\omega\}$ where $f_\mathcal{M}(\xi)$ is the set of states which occur infinitely often in the sequence $f(i, \xi) \in Q^\omega$.

A *Muller acceptor* has the form $(\mathcal{M}, \mathcal{F})$ where $\mathcal{M}$ is an initial automaton and $\mathcal{F} \subseteq C_\mathcal{M}$; it recognizes the set $L(\mathcal{M}, \mathcal{F}) = \{\xi \in A^\omega \mid f_\mathcal{M}(\xi) \in \mathcal{F}\}$. It is well known that Muller acceptors recognize exactly the *regular $\omega$-languages* called also just regular sets. The class $\mathcal{R}$ of all regular $\omega$-languages is a proper subclass of $\Delta^0_3(A^\omega)$.

## 8.2 *Well Quasiorders and Regular Languages*

A basic fact of automata theory (Myhill-Nerode theorem) states that a language $L \subseteq A^*$ is regular iff it is closed w.r.t. some congruence of finite index on $A^*$ (recall that a congruence is an equivalence relation $\equiv$ such that $u \equiv v$ implies $xuy \equiv xvy$, for all $x, y \in A^*$). In [24] the following version of Myhill-Nerode theorem was established:

**Theorem 8.1** *A language $L \subseteq A^*$ is regular iff it is upward closed w.r.t. some monotone WQO on $A^*$ (where a QO $\leq$ on $A^*$ is* monotone *if $u \leq v$ implies $xuy \leq xvy$, for all $x, y \in A^*$).*

Note that any congruence $\equiv$ of finite index on $A^*$ is a monotone WQO on $A^*$ such that the quotient-poset of $(A^*; \equiv)$ is a finite antichain. As observed in Proposition 6.3.1 of [23], an equivalence on $A^*$ is a congruence of finite index iff it is a monotone WQO.

Associate with any monotone WQO $\leq$ on $A^*$ the class $\mathcal{L}_{\leq}$ of upward closed sets in $(A^*; \leq)$. Then $\mathcal{L}_{\leq}$ is a lattice of regular sets. Clearly, $\mathcal{L}_{\leq}$ is closed under the complement iff $\leq$ is a congruence of finite index. In the literature one can find many examples of monotone QOs $\leq$ for which $\mathcal{L}_{\leq}$ is a finite lattice not closed under the complement (in particular, such examples arise from one-sided Ehrenfeucht-Fraïssé games, see e.g. [96]).

Are there other interesting examples of monotone WQOs? An important example is given by the subword relation $\leq^*$ from Sect. 3.1; in this case $\mathcal{L}_{\leq^*}$ is infinite. As observed independently in [32] and [90], $\mathcal{L}_{\leq^*}$ coincides with $\Sigma_1^{\sigma}$, hence there is a relation to the logical approach to automata theory (see the next subsection for additional details). There are other natural examples, for instance for each $k < \omega$ the following relations $\leq_k$ on non-empty words studied e.g. in [32, 90, 106]: $u \leq_k v$, if $u = v \in A^{\leq k}$ or $u, v \in A^{>k}$, $p_k(u) = p_k(v)$, $s_k(u) = s_k(v)$, and there is a $k$-embedding $f : u \to v$. Here $p_k(u)$ (resp. $s_k(u)$) is the prefix (resp., suffix) of $u$ of length $k$, and the $k$-embedding $f$ is a monotone injective function from $\{0. \dots, |u| - 1\}$ to $\{0. \dots, |v| - 1\}$ such that $u(i) \cdots u(i + k) = v(f(i)) \cdots v(f(i) + k)$ for all $i < |u| - k$. Note that the relation $\leq_0$ coincides with the subword relation.

It would be of interest to have more examples of natural monotone WQOs or maybe even a characterization of a wide class of monotone WQOs, in order to understand which classes of regular languages may be obtained in this way. Theorem 8.1 turned out useful in the study of rewriting systems, serving as a tool to prove regularity of languages obtained by such systems. Many interesting facts on this may be found in [23] and references therein. This interesting direction has a strong semigroup-theoretic flavour.

Another development of Theorem 8.1 was initiated in [76] where some analogues of this theorem for infinite words were found. A QO $\preceq$ on $A^\omega$ is a *periodic extension* of a QO $\leq$ on $A^*$ if $\forall i < \omega (u_i \leq v_i)$ implies $u_0 u_1 \cdots \preceq v_0 v_1 \cdots$ and $\forall p \in A^\omega \exists u, v \in A^* (p \preceq uv^\omega \wedge uv^\omega \preceq p)$. Clearly, every periodic extension of a monotone WQO on $A^*$ is WQO on $A^\omega$. For instance, the subword relation on infinite words is a periodic extension of the subword relation on finite words and is therefore WQO. A basic fact in [76] is the following characterization of regular $\omega$-languages.

**Theorem 8.2** *An $\omega$-language $L \subseteq A^\omega$ is regular iff it is upward closed w.r.t. some periodic extension of a monotone WQO on $A^*$.*

Relate to any monotone WQO $\leq$ on $A^*$ the class $\mathcal{L}_{\leq}^{\omega}$ of upward closed sets in $(A^\omega; \preceq)$, for some periodic extension $\preceq$ of $\leq$. Then $\mathcal{L}_{\leq}^{\omega}$ is a class of regular $\omega$-languages. To our knowledge, almost nothing is known on which classes of regular $\omega$-languages are obtained in this way. It seems natural to explore possible relationships of such classes to the "logical" hierarchies of regular $\omega$-languages which are important but are much less understood than the logical hierarchies of languages in the next subsection. For some information on logical hierarchies of regular languages see e.g. [19, 78] and references therein.

## 8.3 Hierarchies of Regular Languages

We denote by $\Sigma_n^\sigma$ the class of languages that can be axiomatized by a $\Sigma_n^0$-sentence of signature $\sigma$. The classes $\Sigma_n^\rho$ are defined analogously with respect to $\rho$. There is a level-wise correspondence of these classes to the well-known concatenation hierarchies of automata theory (see e.g. [78, 107, 113]).

The $\omega$-bases $\mathcal{L}^\sigma = \{\Sigma_n^\sigma\}$ and $\mathcal{L}^\rho = \{\Sigma_n^\rho\}$ do not collapse, and they are neither reducible nor interpolable (see e.g. [96] and references therein). In [110] a natural reducibility $\leq_{qf\rho}$ by quantifier-free formulas of signature $\rho$ was introduced and studied. This reducibility fits the hierarchy $\mathcal{L}^\rho$.

One can of course consider the refinements of the hierarchies $\mathcal{L}^\sigma$ and $\mathcal{L}^\rho$. Among these, the difference hierarchies of sets were studied in detail, see e.g. [32, 33, 90, 96, 106]. Note that many variants of the mentioned hierarchies and reducibilities on regular languages were also considered in the literature (say, for other signatures or other logics in place of the first-order logic).

The main problems about the mentioned hierarchies concern decidability of the corresponding classes of languages or relations between languages (if the languages are given, say, by recognizing automata). Many such decidability problems turn out to be complicated, in particular the decidability of only lower levels of the mentioned hierarchies is currently known.

The relation of this theme to WQOs was not investigated systematically, though the relation to some WQOs was used in proving decidability of levels of the DHs over $\Sigma_1^\sigma$ (and also for some other natural bases, see [32, 90, 96]). Such algorithms are based on the characterization of $\Sigma_1^\sigma$ in terms of the subword partial order mentioned in the previous subsection. Applicability of WQO-theory to higher levels remains unclear.

Concerning the quantifier-free reducibilities, some interesting structural results were obtained in [96, 110]. The relation $\leq_{qf\rho}$ is not WQO on the regular languages but it is open whether $\leq_{qf\rho}$ is WQO on the regular aperiodic languages (or at least on some reasonable subclasses).

The mentioned "logical" hierarchies of regular languages may be defined a similar way for the regular $\omega$-languages, and it is known that this extension brings many new aspects, in particular one has to deal with the topological issues (see e.g. [19] and references therein). The relation to WQO-theory is not clear.

## 8.4 Hierarchies of ω-Regular Languages and k-Partitions

On the class $\mathcal{R}$ of regular $\omega$-languages there is a natural 2-base $\mathcal{L} = (\mathcal{R} \cap \Sigma_1^0, \mathcal{R} \cap \Sigma_2^0)$. As shown in [89], this base is reducible and interpolable. There is also a natural reducibility $\leq_{DA}$ that fits this hierarchy (namely, the reducibility by the so-called deterministic asynchronous finite transducers [117], i.e., by dfa's with output which may print a word at any step). One can also consider the Wadge reducibility, which is denoted in [117] by $\leq_{CA}$; it does not fit this hierarchy.

In [117] K. Wagner gave in a sense the finest possible topological classification of regular $\omega$-languages which subsumes several hierarchies considered before him. Among his main results are the following:

(1) The QO $(\mathcal{R}; \leq_{CA})$ is semi-well-ordered with the order type $\omega^\omega$.
(2) The $CA$-reducibility coincides on $\mathcal{R}$ with the $DA$-reducibility.
(3) Every level of the hierarchy formed by the principal ideals of $(\mathcal{R}; \leq_{DA})$ is decidable.

In [89] the Wagner hierarchy was related to the Wadge hierarchy and to the author's fine hierarchy, namely it is just the FH of sets over $\mathcal{L}$.

Here we briefly discuss the extension of the Wagner hierarchy to the $\omega$-regular $k$-partitions, the class of which is denoted by $\mathcal{R}_k$. Recall from [97] that a *Muller $k$-acceptor* is a pair $(\mathcal{A}, c)$ where $\mathcal{A}$ is an automaton and $c : C_\mathcal{A} \to k$ a $k$-partition of $C_\mathcal{A}$. Such a $k$-acceptor recognizes the $k$-partition $L(\mathcal{A}, c) = c \circ f_\mathcal{A}$ where $f_\mathcal{A} : A^\omega \to C_\mathcal{A}$ is the function defined in Sect. 8.1 As shown in [97], a $k$-partition $L : A^\omega \to k$ is regular iff it is recognized by a Muller $k$-acceptor. Below we also use the 2-base $\mathcal{M} = (\Sigma^0_1, \Sigma^0_2)$. The main results in [97] maybe summarized as follows:

**Theorem 8.3** *(1) The quotient-posets of* $(\mathcal{T}^\sqcup_k(2); \leq_h)$*,* $(\mathcal{R}_k; \leq_{CA})$ *and* $(\mathcal{R}_k; \leq_{DA})$ *are isomorphic.*
*(2) The relations* $\leq_{CA}$ *and* $\leq_{DA}$ *coincide on* $\mathcal{R}_k$*.*
*(3) Every level* $\mathcal{L}(T)$ *of the FH of $k$-partitions over* $\mathcal{L}$ *is decidable.*
*(4) For each* $T \in \mathcal{T}_k(2)$*,* $\mathcal{L}(T) = \mathcal{R}_k \cap \mathcal{M}(T)$*.*

Item (2) is obtained by applying the Büchi-Landweber theorem on infinite regular games [12]. The proof of item (1) is similar to that in Sect. 4.2, but first we have to define the corresponding operations on $k^{A^\omega}$, where $A = \{0, 1, \ldots\}$ is a finite alphabet. This needs some coding because here we work with the compact Cantor space $A^\omega$ while in Sect. 4.2 with the Baire space where the coding is easier.

For all $i < k$ and $A \in k^{A^\omega}$, define the $k$-partition $p_i(A)$ as follows: $[p_i(A)](\xi) = i$, if $\xi \notin 0^*1X^\omega$, otherwise $[p_i(A)](\xi) = A(\eta)$ where $\xi = 0^n1\eta$.

Next we define unary operations $q_0, \ldots, q_{k-1}$ on $k^{A^\omega}$. To simplify notation, we do this only for the particular case $A = \{0, 1\}$ (the general case is treated similarly). Define a $DA$-function $f : 3^\omega \to 2^\omega$ by $f(x_0x_1\cdots) = \tilde{x}_0\tilde{x}_1\cdots$ where $x_0, x_1 \ldots < 3$ and $\tilde{0} = 110000$, $\tilde{1} = 110100$, $\tilde{2} = 110010$ (in the same way we may define $f : 3^* \to 2^*$). Obviously, $f(3^\omega) \in \mathcal{R} \cap \Pi^0_1(A^\omega)$ and there is a $DA$-function $f_1 : 2^\omega \to 3^\omega$ such that $f_1 \circ f = id_{3^\omega}$. For all $i < k$ and $k$-partitions $A$ of $X^\omega$, define the $k$-partition $q_i(A)$ as follows: $[q_i(A)](\xi) = i$, if $\xi \notin f(3^\omega) \vee \forall p \exists n \geq p(\xi[n, n + 5] = \tilde{2})$, $[q_i(A)](\xi) = A(f_1(\xi))$, if $\xi \in f(2^\omega)$, and $[q_i(A)](\xi) = A(\eta)$ in the other cases, where $\xi = f(\sigma 3\eta)$ for some $\sigma \in 3^\omega$ and $\eta \in 2^\omega$.

Finally, we define the binary operation $+$ on $k^{A^\omega}$ as follows. Define a $DA$-function $g : X^\omega \to X^\omega$ by $g(x_0x_1\cdots) = x_0 0 x_1 0 \cdots$ where $x_0, x_1, \ldots \in X$ (in the same way we may define $g : X^* \to X^*$). Obviously, $g(A^\omega) \in \mathcal{R} \cap \Pi^0_1(A^\omega)$ and there is a $DA$-function $g_1 : A^\omega \to X^\omega$ such that $g_1 \circ g = id_{X^\omega}$. For all $k$-partitions $A, B$ of $X^\omega$, we set: $[A + B](\xi) = A(g_1(\xi))$ if $\xi \in g(X^\omega)$, otherwise $[A + B](\xi) = B(\eta)$, where $\xi = g(\sigma)i\eta$ for some $\sigma \in A^\omega$, $i \in A \setminus \{0\}$ and $\eta \in A^\omega$.

The operations $p_i, q_i, +$ have the same properties as the corresponding operations in Sect. 4.2 (only this time we have no infinite disjoint union, so we speak e.g. about $dc$-semilatices instead of the $dc\sigma$-semilatices). Therefore, defining the functions $\mu, \nu, \rho$ just as in Sect. 4.2 (but for finite trees $T$) we get that $\rho : \mathcal{T}^\sqcup_k(2) \to k^{A^\omega}$ induces the embedding of the quotient-poset of $(\mathcal{T}^\sqcup_k(2); \leq_h)$ into those of $(\mathcal{R}_k; \leq_{CA})$ and $(\mathcal{R}_k; \leq_{DA})$. Moreover, one easily checks that for any $T \in \mathcal{T}_k(2)$, $\rho(T)$ is $CA$-complete in $\mathcal{M}(T)$ and $DA$-complete in $\mathcal{L}(T)$.

That this embedding is in fact an isomorphism, and that items (3), (4) hold, follows from analysing some invariants of the Muller $k$-acceptors $\mathcal{A}$ based on the QOs $\leq_0$ and $\leq_1$ on the set of cycles $C_\mathcal{A}$ which extends a standard technique in the Wagner hierarchy.

## 9 Conclusion

When this paper was under review, Arxiv preprint of [64] appeared which contains important results on Wadge degrees. It gives a characterization of the quotient-poset of
$(Q^*; \leq^*)$ from Theorem 3.3 for every countable BQO $Q$. Unifying notation, we denote $(Q^*; \leq^*)$ as $(\mathbf{B}(Q^\mathcal{N}); \leq_W)$ and call elements of $Q^\mathcal{N}$ $Q$-partitions of $\mathcal{N}$. In fact, the authors of [64] also characterize the quotient-poset of $(\Delta^0_{1+\alpha}(Q^\mathcal{N}); \leq_W)$ for each $\alpha < \omega_1$ where $\Delta^0_{1+\alpha}(Q^\mathcal{N})$ consists of all $A : \mathcal{N} \to Q$ such that $A^{-1}(q) \in \Delta^0_{1+\alpha}(\mathcal{N})$ for every $q \in Q$.

As in Sect. 4.2, the characterization uses a suitable iteration $\widetilde{\mathcal{T}}_\alpha(Q)$ of the introduced in [100] operator $\widetilde{\mathcal{T}}$ on the class of all BQOs starting from $Q$, and an extension $\Omega : \widetilde{\mathcal{T}}^\sqcup_\alpha(Q) \to \Sigma^0_{1+\alpha}(Q^\mathcal{N})$ of our embedding $\mu$. Using an induction on BQO $(\Sigma^0_{1+\alpha}(Q^\mathcal{N}); \leq_W)$ one can show that for every $A \in \Sigma^0_{1+\alpha}(Q^\mathcal{N})$ there is $F \in \widetilde{\mathcal{T}}^\sqcup_\alpha(Q)$ with $\Omega(F) \equiv_W A$; this yields the desired isomorphism of quotient-posets. Thus, the idea and the scheme of proof is the same as in [100, 101] but the proof of surjectivity of $\Omega$ in [64] is quite different from our proof of particular cases. The proof in [64] uses a nice extension of the notion of conciliatory set [22] to $Q$-partitions and a deep relation of this field to some basic facts about Turing degrees.

We conclude this survey with collecting some open questions which seem of interest to the discussed topic:

(1) What are the maximal order types of the concrete BPOs mentioned in the paper (except those which are semi-well-ordered or are already known).
(2) Characterize the maximal order types of computable WPOs of rank $\omega$ (or any other infinite computable ordinal).

(3) Characterize the degree spectra of countable WPOs. In particular, is it true that for any given countable graph there is a countable WPO with the same degree spectrum?
(4) Associate with any WPO $P$ the function $f_P : rk(P) \to \omega$ by: $f_P(\alpha)$ is the cardinality of $\{x \in P \mid rk_P(x) = \alpha\}$. Is there a computable WPO $P$ such that $f_P$ is not computable?
(5) Characterize the finite posets $Q$ for which the first-order theory of the quotient-poset of $\mathcal{T}_Q^{\sqcup}$ is decidable.
(6) Is there a finite poset $Q$ of width $\geq 3$ such that the automorphism groups of $Q$ and of the quotient-poset of $\mathcal{T}_Q^{\sqcup}$ are not isomorphic?
(7) Extend the above-mentioned characterizations of the initial segments of $Q$-partitions of Baire space ($Q$ is BQO) beyond the Borel $Q$-partitions.

**Acknowledgements** I am grateful to Leibniz-Zentrum für Informatik for accepting and taking care of Dagstuhl Seminars 08271, 11411, 15392, 16031 which were important for promoting the topic of this paper.

## References

1. Abdulla, P. A., et al. (2004). Using forward reachability analysis for verification of lossy channel systems? *Formal Methods in System Design*, *25*(1), 39–65.
2. Addison, J. W. (1959). Separation principles in the hierarchies of classical and effective set theory. *Fundamenta Mathematicae*, *46*, 123–135.
3. Addison, J. W. (1962). Some problems in hierarchy theory. In *Recursive function theory*. Proceedings of symposia in pure mathematics (Vol. 5 , pp. 123–130). AMS.
4. Addison, J. W. (1962). The theory of hierarchies. In *Logic, methodology and philosophy of science*. Proceedings of the 1960 international congress, Palo Alto (pp. 26–37).
5. Addison, J. W. (1965). The method of alternating chains. In: *The theory of models* (pp. 1–16). Amsterdam, North Holland.
6. Abramsky, S., & Jung, A. (1994). Domain theory. In *Handbook of logic in computer science* (Vol. 3, pp. 1–168). Oxford.
7. Ash, C. J., & Knight, J. (2000). *Computable structures and the hyperarithmetical hierarchy*. Amsterdam: Elsevier Science.
8. Adams, S., & Kechris, A. S. (2000). Linear algebraic groups and countable Borel equivalence relations. *Journal of the American Mathematical Society*, *13*(4), 909–943.
9. Andretta, A. (2006). More on Wadge determinacy. *Annals of Pure and Applied Logic*, *144*(1–3), 2–32.
10. Andretta, A., & Martin, D. A. (2003). Borel-Wadge degrees. *Fundamenta Mathematicae*, *177*(2), 175–192.
11. Becker, H. (1988). A characterization of jump operators. *Journal of Symbolic Logic*, *53*(3), 708–728.
12. Büchi, J. R., & Landweber, L. H. (1969). Solving sequential conditions by finite-state strategies. *Transactions of the American Mathematical Society*, *138*, 295–311.
13. Brattka, V., & Gherardi, G. (2009). Weihrauch degrees, omniscience principles and weak computability. arXiv:0905.4679.
14. Brattka, V., & Gherardi, G. (2009). Effective choice and boundedness principles in computable analysis. arXiv:0905.4685.
15. Carroy, R. (2013). A quasiorder on continuous functions. *Journal of Symbolic Logic*, *78*(2), 633–648.

16. Cenzer, D. (1991). Polynomial-time versus recursive models. *Annals of Pure and Applied Logic*, *54*(1), 17–58.
17. de Brecht, M. (2013). Quasi-Polish spaces. *Annals of Pure and Applied Logic*, *164*, 356–381.
18. Downey, R. G., & Fellows, M. R. (1998). *Parameterized complexity*. New York: Springer.
19. Diekert, V., & Kufleitner, M. (2011). Fragments of first-order logic over infinite words. *Theory of Computing Systems*, *48*(3), 486–516.
20. de Jongh, D. H. J., & Parikh R. (1977). Well-partial orderings and hierarchies. In *Indagationes mathematicae (Proceedings)* (Vol. 39, No. 3, pp. 95–207).
21. Davey, B. A., & Priestley, H. A. (1994). *Introduction to lattices and order*. Cambridge:
22. Duparc, J. (2001). Wadge hierarchy and Veblen hierarchy, part I. *Journal of Symbolic Logic*, *66*(1), 56–86.
23. de Luca, A., & Varricchio, S. (1999). *Finiteness and regularity in semigroups and formal languages*. Berlin: Springer.
24. Ehrenfeucht, A., Haussler, D., & Rozenberg, G. (1983). On regularity of context-free languages. *Theoretical Computer Science*, *27*, 311–332.
25. Engelking, R. (1989). *General topology*. Berlin: Heldermann.
26. van Engelen, F. (1986). *Homogeneous zero-dimensional absolute Borel sets*. Amsterdam: CWI Tracts.
27. van Engelen, F., Miller, A., & Steel, J. (1987). Rigid Borel sets and better quasiorder theory. *Contemporary Mathematics*, *65*, 199–222.
28. Ershov, Y. L. (1968). On a hierarchy of sets 3 (in Russian). *Algebra i Logika*, *7*(4), 15–47.
29. Ershov, Y. L., & Goncharov, S. S. (1999). *Constructive models*. Novosibirsk: Scientific Book (in Russian, there is an English Translation).
30. Ershov, Y. L., Lavrov, I. A., Taimanov, A. D., & Taitslin, M. A. (1965). Elementary theories. *Uspekhi Matematicheskikh Nauk*, *20*(4), 37–108. (in Russian).
31. Gao, S. (2009). *Invariant descriptive set theory*. Pure and applied mathematics: A series of monographs and textbooks (p. 293). Taylor & Francis Group.
32. Glaßer amd, C. & Schmitz, H., (2001). The Boolean structure of dot-depth one. *Journal of Automata, Languages and Combinatorics*, *6*, 437–452.
33. Glasser, C., Schmitz, H., & Selivanov, V. (2016). Efficient algorithms for membership in Boolean hierarchies of regular languages. *Theoretical Computer Science, 646*(C), 86–108. https://doi.org/10.1016/j.tcs.2016.07.017. ISSN 0304-3975.
34. Hertling, P. (1993, December). Topologische Komplexitätsgrade von Funktionen mit endlichem Bild. Informatik-Berichte 152, 34 pages, Fernuniversität Hagen.
35. Hertling, P. (1996). *Unstetigkeitsgrade von Funktionen in der effektiven Analysis*. Ph.D. thesis, Fachbereich Informatik, FernUniversität Hagen.
36. Hertling, P., & Selivanov, V. (2014). Complexity issues for preorders on finite labeled forests. In V. Brattka, H. Diener, & D. Spreen (Eds.), *Logic, computation, hierarchies* (pp. 165–190). Boston, Berlin: Ontos Publishing, de Gruiter.
37. Hertling, P., & Weihrauch, K. (1994). Levels of degeneracy and exact lower complexity bounds for geometric algorithms. In *Proceedings of the 6th Canadian Conference on Computational Geometry, Saskatoon* (pp. 237–224).
38. Higman, G. (1952). Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society*, *2*(3), 326–336.
39. Hirschfeldt, D. R. (2014). *Slicing the truth—On the computable and reverse mathematics of combinatorial principles*. Lecture notes series (Vol. 28). Institute for Mathematical Sciences, National University of Singapore: World Scientific.
40. Hodges, W. (1993). *Model theory*. Cambridge University Press.
41. Hurewicz, W., & Wallman, H. (1948). *Dimension theory*. Princeton mathematical series (Vol. 4). Princeton, NJ: Princeton University Press.
42. Ikegami, D. (2010). Games in set theory and logic. Ph.D. thesis, University of Amsterdam.
43. Ikegami, D., Schlicht, P., & Tanaka, H. (2012). Continuous reducibility for the real line, preprint, submitted. *Appeared as: Borel subsets of the real line and continuous reducibility, Fundamenta Mathematicae*, *244*(3), 209–241.

44. Jezek, J., & McKenzie, R. (2009). Definability in substructure orderings, IV: Finite lattices. *Algebra Universalis*, *61*(3–4), 301–312.
45. Jezek, J., & McKenzie, R. (2009). Definability in substructure orderings, I: Fifinite semilattices. *Algebra Universalis*, *61*(1), 59–75.
46. Jezek, J., & McKenzie, R. (2009). Definability in substructure orderings, III: Finite distributive lattices. *Algebra Universalis*, *61*(3–4), 283–300.
47. Jezek, J., & McKenzie, R. (2010). Definability in substructure orderings, II: Finite ordered sets. *Order*, *27*(2), 115–145.
48. Kunos, A. (2015). Definability in the embeddability ordering of finite directed graphs. *Order*, *32*(1), 117–133.
49. Kruskal, J. B. (1960). Well-quasi-ordering, the tree theorem, and Vazsonyis conjecture. *Transactions of the American Mathematical Society*, *95*, 210–225.
50. Karandikar, P., Niewerth, M., & Schnoebelen, Ph. (2016). On the state complexity of closures and interiors of regular languages with subwords and superwords. *Theoretical Computer Science*, *610*, 91–107.
51. Karandikar, P., & Schnoebelen, Ph. (2015). Decidability in the logic of subsequences and supersequences. In *Proceedings of FST&TCS 2015, LIPIcs 45* (pp. 84–97). Leibniz-Zentrum für Informatik.
52. Karandikar, P., & Schnoebelen, Ph. (2015). Generalized post embedding problems. *Theory of Computing Systems*, *56*(4), 697–716.
53. Karandikar, P., & Schnoebelen, Ph. (2016). The height of piecewise-testable languages with applications in logical complexity. In *Proceeding of CSL 2016, LIPIcs62* (37:1–37:22). Leibniz-Zentrum für Informatik.
54. Kudinov, O. V., & Selivanov, V. L. (2006). Undecidability in the homomorphic quasiorder of finite labeled forests. In *Proceedings of CiE 2006*. Lecture notes in computer science (Vol. 3988, pp. 289–296). Berlin: Springer.
55. Kudinov, O. V., & Selivanov, V. L. (2007). Definability in the homomorphic quasiorder of finite labeled forests. In *Proceedings of CiE-2007*. Lecture notes in computer science (Vol. 4497, 436–445). Berlin: Springer.
56. Kudinov, O. V., & Selivanov, V. L. (2007). Undecidability in the homomorphic quasiorder of finite labelled forests. *Journal of Logic and Computation*, *17*, 113–1151.
57. Kudinov, O. V., Selivanov, V. L. (2009). A Gandy theorem for abstract structures and applications to first order definability. In *Proceeding of CiE-2009*. LNCS 5635 (pp. 290–299). Berlin: Springer.
58. Kudinov, O., & Selivanov, V. (2009). Definability in the infix order on words. In V. Diekert & D. Nowotka (Eds.), *Proceeding of DLT-2009* (Vol. 5583, pp. 454–465). Lecture notes in computer science. Berlin: Springer.
59. Kudinov, O. V., Selivanov, V. L., & Yartseva, L. V. (2010). Definability in the subwordx order. In F. Ferreira, B. Loewe, E. Mayordomo and L. M. Gomes (Eds.), *Proceeding of CiE-2010*. Lecture notes in computer science (Vol. 6158, pp. 246–255). Berlin: Springer.
60. Kudinov, O. V., Selivanov, V. L., & Zhukov, A. V. (2009). Definability in the h-quasiorder of labeled forests. *Annals of Pure and Applied Logic*, *159*(3), 318–332.
61. Kuske, D. (2006). Theories of orders on the set of words. *Theoretical Informatics and Applications*, *40*, 53–74.
62. Kuratowski, K., & Mostowski, A. (1967). *Set theory*. North Holland.
63. Kihara, T., & Montalban, A. (17 Aug 2016). The uniform Martin's conjecture for many-one degrees. arXiv:1608.05065 v1 [Math.LO] . Accepted by Transactions of AMS. Appeared in: *Transactions of the American Mathematical Social, 370*(12), 9025–9044, 2018.
64. Kihara. T., & Montalban, A. (22 May 2017). On the structure of the Wadge degrees of BQO-valued Borel functions. arXiv:1705.07802 v1 [Math.LO]. Accepted by Transactions of AMS. Appeared in: *Transactions of the American Mathematical Social, 371*(11):7885–7923, 2019.
65. Kechris, A. S. (1995). *Classical Descriptive Set Theory*. New York: Springer.
66. Laver, R. (1971). On Fraïssés order type conjecture. *Ann. of Math.*, *93*(2), 89–111.

67. Louveau, A. (1983). *Some results in the Wadge hierarchy of Borel sets*. Lecture Notes in Mathematics (No. 1019, pp. 28–55).
68. Marks, A. S. (13 Jan 2016). The universality of polynomial time Turing equivalence. arXiv:1601.033431 [math LO].
69. Marks, A. S., Slaman, T., & Steel, J. Martins conjecture, arithmetic equivalence and countable Borel equivalence relations. Appeared in: (2016) In *Ordinal Definability and Recursion Theory: The Cabal Seminar. Lecture notes in logic* (Vol. 3, No.43, pp. 493–520). Cambridge University Press, Cambridge.
70. Montalban, A. (2007). Computable linearizations of well-partial-orderings. *Order*, *24*, 39–48. https://doi.org/10.1007/s11083-007-9058-0.
71. Motto Ros, L. (2009). Borel-amenable reducibilities for sets of reals. *Journal of Symbolic Logic*, *74*(1), 27–49.
72. Motto Ros, L., Schlicht, P., & Selivanov, V. (2015). Wadge-like reducibilities on arbitrary quasi-Polish spaces. *Mathematical Structures in Computer Science*, *25*(8), 1705–1754. https://doi.org/10.1017/S0960129513000339.
73. Nash-Williams, C. St. J. A. (1965). On well-quasi-ordering infinite trees. *Mathematical proceedings of the Cambridge philosophical society* (Vol. 61, pp. 697–720). Cambridge University Press.
74. Nash-Williams, C. St. J. A. (1968). On better-quasi-ordering transfinite sequences. In *Mathematical proceedings of the Cambridge philosophical society* (Vol. 64, p. 273290). Cambridge University Press.
75. Nies, A. (2000). Definability in the c.e. degrees: Questions and results. *Contemporary Mathematics*, *257*, 207–213.
76. Ogawa, M. (2004). Well quasiorders and regular $\omega$-languages. *Theoretical Computer Science*, *324*(1), 55–60.
77. Pequignot, Y. (2015). A Wadge hierarchy for second countable spaces. *Archive for Mathematical Logic*, *54*(5–6), 659683. https://doi.org/10.1007/s00153-015-0434-y.
78. Perrin, D., & Pin, J.-E. (2004). *Infinite words. Volume 141 of pure and applied mathematics*.
79. Rogers, H, Jr. (1967). *Theory of recursive functions and effective computability*. New York: McGraw-Hill.
80. Robertson, N., Seymour, P. D. (2004). Graph minors. XX. Wagners conjecture. *Journal of Combinatorial Theory, Series B, 92*(2), 325–357.
81. Ramanujam, R., & Thinniyam, R. S. (2016). Definability in first order theories of graph orderings. In *Logical foundations of computer science* (pp. 31–348). Springer.
82. Schmidt, D. (1979). *Well-partial orderings and their maximal order types*. Habilitationsschrift: University of Heidelberg.
83. Soare, R. I. (1987) *Recursively enumerable sets and degrees. Perspectives in mathematical Logic. A study of computable functions and computably generated sets*. Berlin: Springer.
84. Selivanov, V. L. (1982). On the structure of degrees of generalized index sets. *Algebra and Logic*, *21*, 316–330.
85. Selivanov, V. L. (1983). Hierarchies of hyperarithmetical sets and functions. *Algebra and Logic*, *22*, 473–491.
86. Selivanov, V. L. (1988). On algorithmic complexity of algebraic systems. *Mathematical Notes*, *44*(5–6), 944–950.
87. Selivanov, V. L. (1992). *Hierarchies, numberings, index sets*. Handwritten notes (300 pp.).
88. Selivanov, V. L. (1995). Fine hierarchies and Boolean terms. *Journal of Symbolic Logic*, *60*(1), 289–317.
89. Selivanov, V. L. (1998). Fine hierarchy of regular $\omega$-languages. *Theoretical Computer Science*, *191*, 37–59.
90. Selivanov, V. L. (2010). *A logical approach to decidability of hierarchies of regular star-free languages*. Lecture notes in computer science (Vol. 2010, pp. 539–550). Berlin: Springer.
91. Selivanov, V. L. (2005). Variations on the Wadge reducibility. *Siberian Advances in Mathematics*, *15*(3), 44–80.

92. Selivanov, V. (2006). Towards a descriptive set theory for domain-like structures. *Theoretical Computer Science*, *365*, 258–282.
93. Selivanov, V. L. (2007). The quotient algebra of labeled forests modulo h-equivalence. *Algebra and Logic*, *46*(2), 120–133.
94. Selivanov, V. L. (2007). Hierarchies of $\Delta^0_2$-measurable $k$-partitions. *Mathematical Logic Quarterly*, *53*, 446–461.
95. Selivanov, V. L. (2008). Fine hierarchies and $m$-reducibilities in theoretical computer science. *Theoretical Computer Science*, *405*, 116–163.
96. Selivanov, V. L. (2009). Hierarchies and reducibilities on regular languages related to modulo counting. *RAIRO Theoretical Informatics and Applications*, *41*, 95–132.
97. Selivanov, V. L., (2011). A fine hierarchy of $\omega$-regular $k$-partitions. In B. Löwe (Eds.), Proceeding of CiE, et al. (2011). *LNCS 6735* (pp. 260–269). Heidelberg: Springer.
98. Selivanov, V. L. (2012). Fine hierarchies via Priestley duality. *Annals of Pure and Applied Logic*, *163*, 1075–1107.
99. Selivanov, V. (2013). Total representations, logical methods in computer. *Science*, *9*(2), 1–30. https://doi.org/10.2168/LMCS-9(2:5)2013.
100. Selivanov, V. L. (2017).Towards a descriptive theory of cb0-spaces. *Mathematical Structures in Computer Science, 28*(8), 1553–1580. https://doi.org/10.1017/S0960129516000177. Earlier version in: ArXiv:1406.3942v1 [Math.GN] 16 June 2014.
101. Selivanov, V. L., Wadge, Extending, & theory to k-partitions. In J. Kari, F. Manea, & Ion Petre (Eds.), Proceeding of CiE,. (2017). *LNCS 10307* (pp. 387–399). Berlin: Springer.
102. Schlicht, P. (2017). Continuous reducibility and dimension of metric spaces. arXiv:1703.10144.
103. Simpson, S.G. (1985). Bqo-theory and Fraïssé conjecture. Chapter 9, Mansfield, R., & Weitkamp, G. (1985). *Recursive aspects of descriptive set theory*. New York: Oxford University Press.
104. Slaman, T. A., & Steel, J. R. (1988). Definable functions on degrees. In *Cabal seminar 81–85*. Lecture notes in mathematics (Vol. 1333, pp. 37–55). Berlin: Springer.
105. Saint Raymond, J. (2007). Preservation of the Borel class under countable-compact-covering mappings. *Topology and Its Applications*, *154*, 1714–1725.
106. Stern, J. (1985). Characterizations of some classes of regular events. *Theoretical Computer Science*, *35*, 163–176.
107. Straubing, H. (1994). *Finite automata, formal logic and circuit complexity*. Boston: Birkhäuser.
108. Steel, J. R. (1982). A classification of jump operators. *Journal of Symbolic Logic*, *47*(2), 347358.
109. Steel, J. (1980). Determinateness and the separation property. *Journal of Symbolic Logic*, *45*, 143–146.
110. Selivanov, V. L., & Wagner, K. W. (2005). A reducibility for the dot-depth hierarchy. *Theoretical Computer Science*, *345*(2–3), 448–472.
111. Trakhtenbrot, B. A., & Barzdin, J. M. (1970). *Finite automata. Behavior and synthesis* (p. 1973). Moscow: Mir (Russian, English translation: North Holland, Amsterdam).
112. Thomas, W. (1990). Automata on infinite objects. *Handbook of theoretical computer science* (Vo. B, pp. 133–191).
113. Thomas, W. (1996). Languages, automata and logic. *Handbook of formal language theory* (Vol. B, pp. 133–191).
114. Tarski, A., Mostowski, A., & Robinson, J. (1953). *Undecidable Theories*. Amsterdam: North Holland.
115. Van Wesep, R. (1976). Wadge degrees and descriptive set theory. *Lecture Notes in Mathematics*, *689*, 151–170.
116. Wadge, W. (1984). Reducibility and determinateness in the Baire space. Ph.D. thesis, University of California, Berkely.
117. Wagner, K. (1979). On $\omega$-regular sets. *Information and Control*, *43*, 123–177.

118. Weihrauch, K. (1992). The degrees of discontinuity of some translators between representations of the real numbers. Technical Report TR-92-050. Berkeley: International Computer Science Institute.
119. Weihrauch, K. (2000). *Computable analysis*. Berlin: Springer.
120. Wires, A. (2012). Definability in the substructure ordering of simple graphs. http://www.math.uwaterloo.ca/awires/SimpleGraphs.pdf.
121. Wires, A. (2016). Definability in the substructure ordering of simple graphs. *Annals of Combinatorics*, *20*(1), 139–176.

# A Combinatorial Bound for a Restricted Form of the Termination Theorem

**Silvia Steila**

**Abstract** We present a combinatorial bound for the H-bounded version of the Termination Theorem. As a consequence we improve the result by Solovay and Ketonen on the relationship between the Paris–Harrington Theorem and the Fast Growing Hierarchy.

## 1 Introduction

A *transition-based program* is a triple $P = (S, I, R)$, where $S$ is the set of *states* of $P$, $I \subseteq S$ is the set of *initial states* and $R \subseteq S \times S$ is the *transition relation* of $P$. A *computation* is a maximal $R$-decreasing sequence of states, which starts in some initial state. The set of *accessible states* (in symbols Acc) is the set of all states which appear in some computation. A program $P$ is *terminating* if its transition relation restricted to the accessible states ($R \cap (\text{Acc} \times \text{Acc})$) is well-founded (i.e., there are no infinite computations).

In [12] Podelski and Rybalchenko characterized the termination of transition-based programs via disjunctively well-founded transition invariants. A *transition invariant* $T$ for $P$ is a superset of the transitive closure of the transition relation of $P$ restricted to the accessible states, namely $T \supseteq R^+ \cap (\text{Acc} \times \text{Acc})$. $T$ is *disjunctively well-founded* if it is the finite union of well-founded relations.

The Termination Theorem by Podelski and Rybalchenko states that a transition-based program $P$ is terminating if and only if there exists a disjunctively well-founded transition invariant for $P$, i.e., there exists a natural number $k$ and $k$-many well-founded relations, whose union contains the transition relation of $P$.

The original proof of the Termination Theorem makes use of Ramsey's Theorem for pairs [13] which is a purely classical result [15]. Therefore the extraction of bounds from that proof is a non-trivial task. As Gasarch [4] pointed out, the methods

S. Steila (✉)
Institut für Informatik, Universität Bern, Neubrückstrasse 10, 3012 Bern, Switzerland
e-mail: steila@inf.unibe.ch; silvia.steila2@gmail.com



P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_11

described are used by real program checkers. Hence the importance of studying bounds for termination.

In this paper a *bound* for a transition-based program $P = (S, I, R)$ is a function which associates to every input of $P$, a natural number greater or equal to the number of steps computed by $P$. By unfolding definitions, it is a function that associates to every $s \in I$, a natural number greater or equal to the length of every $R$-decreasing sequence from $s$.

In [2] bounds are extracted by considering an intuitionistic proof of the Termination Theorem which exploits the definition of inductive well-foundedness. In [18] there is a different intuitionistic proof of the Termination Theorem. As far as we know, no bound analysis based on this argument has been conducted.

In [1] bounds are extracted by using Spector's bar recursion, while in [17] the termination analysis is investigated from the point of view of Reverse Mathematics.

All these approaches highlighted that the class of functions computable by a program for which there exists a disjunctively well-founded transition invariant whose relations have primitive recursive weight functions is exactly the class of primitive recursive functions. Where a *weight function* for a binary relation $R \subseteq S^2$ is a function $f : S \to \mathbb{N}$ such that for any $x,\ y \in S$

$$xRy \implies f(x) < f(y).$$

For short we say that a disjunctively well-founded transition invariant has *height* $\omega$ if it is composed of primitive recursive relations with primitive recursive weight functions.

As shown in [17], this restricted form of the Termination Theorem can be related to some corollary of the Paris–Harrington Theorem for pairs [11] over $RCA_0^*$, namely the Weak Paris–Harrington Theorem. Let us denote by $\mathcal{F}_h$ the $h$-th class of the Fast Growing Hierarchy. This connection, together with a result by Solovay and Ketonen [7, Theorem 6.7] which relates the Fast Growing Hierarchy to the Paris–Harrington Theorem, yields that if we have $k$-many relations whose weight functions are in $\mathcal{F}_h$ then $R$ is bounded by a function in $\mathcal{F}_{k+5+h}$. Hence there is a function $f \in \mathcal{F}_{k+5+h}$, such that, for every state $s$, $f(s)$ is greater or equal of the length of any $R$-decreasing sequence which starts in $s$.

This bound is not optimal, as shown by Figueira et al. [3], who provided instead an optimal bound in $\mathcal{F}_{k+\max\{h-1,1\}}$ by exploiting their bound for the Dickson Lemma. Indeed from any computation of a program which has a transition invariant composed of $k$-many relations with weight functions, we can extract a "bad"-sequence for the well-quasi order $\mathbb{N}^k$ (see Sect. 3). We conjecture that the bound provided in [3] may be formalized in $RCA_0^*$ plus the assumption of the totality of every function in $\mathcal{F}_h$ (in symbols $\mathrm{Tot}(\mathcal{F}_h)$). From that we would extract an improved version of the result by Solovay and Ketonen, but only for the Weak Paris–Harrington Theorem.

This work, extracted from the Phd Thesis of the author [16], provides a combinatorial bound for the H-bounded version of the Termination Theorem, which is proved to be equivalent to the Paris–Harrington Theorem in [17]. This version of the Termination Theorem is based on the notion of H-well-foundedness. We say that a relation is *H-well-founded* if and only if any infinite transitive sequence of elements contains an increasing pair. For transitive relations being H-well founded is equivalent to being well-founded. Note that if $R$ is a well-quasi order then $R$ is H-well-founded, as $R$ is transitive, while the vice versa does not hold.

As a consequence of our bound we get a stronger version of Solovay and Ketonen's result. Namely, for every natural number $k$, $\mathrm{Tot}(\mathcal{F}_{k+5})$ implies the Paris–Harrington Theorem ($\mathrm{PH}^2_k$) for pairs in $k$-many colors over $RCA_0^*$: for every $a \in \mathbb{N}$ there exists a $b \in \mathbb{N}$ such that for every coloring $c : [[a, b]]^2 \to k$ there exists a finite homogeneous set $H \subseteq [a, b]$ such that $\min H < |H|$. We prove that $\mathrm{Tot}(\mathcal{F}_{k+1})$ is enough to derive $\mathrm{PH}^2_k$.

## 2 Background Material

Starting point of this paper is the relation of some restricted forms of the Termination Theorem and some restricted forms of the Paris–Harrington Theorem for pairs, presented in [17].

We work over $RCA_0^*$, which is defined for the language of second order arithmetic enriched with an exponential operation (e.g., [14]). $RCA_0^*$ consists of the basic axioms together with the exponentiation axioms (elementary function arithmetic), $\Delta^0_0$ induction and $\Delta^0_1$-comprehension. Since in $RCA_0$ there exists the exponential function, $RCA_0 \equiv RCA_0^* + \Sigma^0_1$-induction [14, Sect. X.4].

### *2.1 Fast Growing Hierarchy*

We denote with $\mathcal{F}_k$ the usual $k$-class of the Fast Growing Hierarchy [9]. Define

$$\begin{cases} F_0(x) = x + 1, \\ F_{n+1}(x) = {F_n}^{(x+1)}(x). \end{cases}$$

Then $\mathcal{F}_k$ is the closure under limited recursion and substitution of the set of functions defined by constant, projections, sum and $F_h$ for all $h \leq k$. We need also to recall some results by Löb and Wainer [9].

**Proposition 1** *1. For all $k, k', n, m, x, y \in \mathbb{N}$:*

- $x < y \implies F^n_k(x) < F^n_k(y)$;
- $m < n \implies F^m_k(x) < F^n_k(x)$;
- $k < k' \implies F^n_k(x) \leq F^n_{k'}(x)$, *with equality holding only if* $n = x = 0$.

*2. For every $k \in \mathbb{N}$ and for each $f \in \mathcal{F}_k$, there exists $n \in \mathbb{N}$ such that for every $x$ $f(x) \leq F^n_k(x)$.*

Due to these results we directly get the following.

**Corollary 2** *If* $f \in \mathcal{F}_k$ *and* $g \in \mathcal{F}_{k'}$*, for some* $k, k' \in \mathbb{N}$*, then the function* $h(x) = f^{g(x)}(x)$ *is in* $\mathcal{F}_{\max\{k+1,k'\}}$.

*Proof* Let $m, n \in \mathbb{N}$ such that $f(x) \leq F_k^m(x)$ and $g(x) \leq F_{k'}^n(x)$. Therefore

$$h(x) \leq F_k^{mF_{k'}^n(x)}(x) \leq F_k^{mF_{k'}^n(x)}(mF_{k'}^n(x)) \leq F_{k+1}(mF_{k'}^n(x)).$$

□

## 2.2 Paris–Harrington Theorem

The Paris–Harrington Theorem [11] is a strengthened version of finite Ramsey's Theorem which implies the consistency of Peano Arithmetic.

Given $X \subseteq \mathbb{N}$ we denote with $[X]^2$ the complete graph on $X$: any subset $\{x, y\}$ of $X$, for $x$ and $y$ distinct, is an edge of the graph. Given $k \in \mathbb{N}$, any map $c : [\mathbb{N}]^2 \to k$ is called a *coloring* of $[X]^2$ in $k$ colors. If $c(\{x, y\}) = i < k$, then we say that the edge $\{x, y\}$ has color $i$ or that $x$ and $y$ are connected in color $i$.

Given a coloring $c : [X]^2 \to k$, a set $Y \subseteq X$ is *homogeneous* if there exists $i \in k$ such that for every $x, y \in Y$ $c(\{x, y\}) = i$. The Paris–Harrington Theorem states that for any coloring $c : [X]^2 \to k$ over the edges of the complete graph on some infinite set $X \subseteq \mathbb{N}$ in $k$-many colors, there exists a finite homogeneous set $H$ such that $\min H < |H|$. Following the approach of [17], we can slice this statement with respect to the complexity of the set $X$. For every natural number $h$ and $k$ we define

($\mathrm{PH}_k^{h,2}$): Given $f, g : \mathbb{N} \to \mathbb{N}$ such that $g \in \mathcal{F}_{h-1}$ and for all $n \in \mathbb{N}$ $f(n+1) < g(f(n))$, for all coloring $c : [\mathrm{ran}(f)]^2 \to k$, there exists a homogeneous set $H$ for $c$ such that $\min H < |H|$.

The Weak Paris–Harrington Theorem is an immediate corollary of the Paris–Harrington Theorem which turns out to be important, for the sake of termination analysis. A set $Y \subseteq X$ is *weakly homogeneous* if its increasing enumeration $Y = \{y_0 < y_1 < \cdots < y_n < \cdots\}$ is such that there exists $i \in k$ such that for every $n$ $c(\{y_n, y_{n+1}\}) = i$. The Weak Paris–Harrington Theorem states that for any coloring $c : [X]^2 \to k$ over the edges of the complete graph on some infinite set $X \subseteq \mathbb{N}$ in $k$-many colors, there exists a finite weakly homogeneous set $H$ such that $\min H < |H|$. Once again we can consider the restricted version for any natural numbers $h$ and $k$.

($\mathrm{WPH}_k^{h,2}$): Given $f, g : \mathbb{N} \to \mathbb{N}$ such that $g \in \mathcal{F}_{h-1}$ and for all $n \in \mathbb{N}$ $f(n+1) < g(f(n))$, for all coloring $c : [\mathrm{ran}(f)]^2 \to k$, there exists a weakly homogeneous set $H$ for $c$ such that $\min H < |H|$.

Solovay and Ketonen in [7, Theorem 6.7] proved that for any natural number $k$, $(k+5)$-$\mathrm{LRG}^h$ implies $\mathrm{PH}_k^{h,2}$, where $(k+5)$-$\mathrm{LRG}^h$ is a statement equivalent to $\mathrm{Tot}(\mathcal{F}_{k+h+5})$ (for more details see for example [17, Sect. 8]). In particular they show that for every $a, b \in \mathbb{N}$ if $[a, b]$ is $(k+5)$-LRG then for every $c : [[a, b]]^2 \to k$ there exists a finite homogeneous set $H$ such that $\min H < |H|$. Due to the relation with the Fast Growing Hierarchy we get that $\mathrm{Tot}(\mathcal{F}_{k+5})$ implies that for every $a \in \mathbb{N}$ there exists a $b \in \mathbb{N}$ such that for every coloring $c : [[a, b]]^2 \to k$ there exists a finite homogeneous set $H \subseteq [a, b]$ such that $\min H < |H|$. Solovay and Ketonen's result formalizes over $RCA_0^*$.

## 2.3 Restricted Forms of the Termination Theorem

Given a deterministic binary relation $R$ with transition function $t : \mathbb{N} \to \mathbb{N}$ (i.e., $R = \{(t(x), x) : x \in \mathrm{ran}(t) \cup I\}$), we say that $R$ *is the graph of a function*[1] *in* $\mathcal{F}_h$, if there exists $g \in \mathcal{F}_h$ such that for all $x \in S, t(x) < g(x)$. We say that $R$ *has control function in* $\mathcal{F}_h$, if there exists $g \in \mathcal{F}_h$ such that for all $n \in \mathbb{N}$ and any $x \in S$, $t^n(x) < g(n, x)$.

*Remark 3* Note that if $R$ is a graph of a function in $\mathcal{F}_h$ with witness $g$, then $R$ has control function in $\mathcal{F}_{h+1}$. Indeed for any $n \in \mathbb{N}$ and $x \in S$, $t^n(x) < g^n(x)$.

As already mentioned in the introduction, in this paper we will deal with two restricted versions of the Termination Theorem. The first one is the Termination Theorem for weight functions:

($k$-$\mathrm{TT}^h$): Let $R$ be a deterministic binary relation which is a graph of a function in $\mathcal{F}_{h-1}$. If there exists a disjunctively well-founded transition invariant for $R$ composed of $k$-many relations with corresponding weight functions $f_i \in \mathcal{F}_h$, then $R$ is well-founded.

Working in $RCA_0^* + \mathrm{Tot}(F_h)$ we can prove that for every natural number $k$, $k$-$\mathrm{TT}^h$ implies $\mathrm{WPH}_k^{h,2}$ [17, Theorem 7.1]. For the sake of completeness we recall the argument.

**Theorem 4** ($RCA_0^* + \mathrm{Tot}(\mathcal{F}_h)$, [17]) *For any natural number k,*

$$k\text{-}\mathrm{TT}^h \implies \mathrm{WPH}_k^{h,2}.$$

*Proof* Assume by contradiction that there exist $X$ and $c : [X]^2 \to k$ such that, there is no weakly homogeneous set $H$ for $c$ such that $\min H < |H|$. For any $i \in k$, define $R_i$ as follows:

$$x R_i y \iff y < x \wedge x, y \in X \wedge c(\{y, x\}) = i.$$

We claim that $R_i$ has height $\omega$ for any $i \in k$. Indeed, give $f, g : \mathbb{N} \to \mathbb{N}$ such that $X = \{f(i) : i \in \mathbb{N}\}$, $g \in \mathcal{F}_{h-1}$ and for any $n$ $f(n+1) < g(f(n))$, we can define a weight function $f_i : X \to \mathbb{N}$, by limited recursion in $\mathcal{F}_h$:

[1]In our context $R$ is the graph of the function $t$ if $R = \{(t(x), x) : x \in \mathrm{ran}(t) \cup I\}$.

$$f_i(f(n)) = \begin{cases} f(0) & \text{if } n = 0; \\ w(n) & \text{otherwise;} \end{cases}$$

$$w(n) = \min(\{f_i(f(m)) - 1 : m < n \wedge c(\{f(m), f(n)\}) = i\} \cup \{f(n)\})$$

$$f_i(f(n)) \leq f(n).$$

Note that $f \in \mathcal{F}_h$, since for any $n$ we have $f(n) \leq g^n(f(0))$ and $g \in \mathcal{F}_{h-1}$. Note that $f_i$ is a weight function, since if $xR_iy$ then $c(\{y, x\}) = i$ and $y < x$ and so $f_i(x) < f_i(y)$. Moreover for any $x \in X$ we have $f_i(x) \geq 0$. Otherwise there should exist $y_0 > \cdots > y_l$ such that

$$y_0 R_i y_1 \ldots R_i y_l = x,$$

where $l > y_0$, due to the definition of $f_i$ and since $X \subseteq \mathbb{N}$. This is a contradiction since we assumed that there is no weakly homogeneous sets for $c$. Then each $R_i$ has weight function $f_i \in \mathcal{F}_h$.

Therefore, by applying $k$-TT$^h$, $\{(f(n+1), f(n)) : n \in \mathbb{N}\} \subseteq \bigcup\{R_i : i \in k\}$ should be well-founded, but this is a contradiction.

The second version of the Termination Theorem is stronger than the one above and it is based on the notion of H-bound (where H means homogeneous). A *H-bound* for a binary relation $R$ is a function $f : S \to \mathbb{N}$ such that for every $R$-decreasing transitive sequence $\langle a_0, \ldots, a_{l-1}\rangle$, $l \leq f(a_0)$, i.e, any decreasing transitive $R$-sequence starting from $a$ is shorter than $f(a)$.

($k$-TT$^h_H$): Let $R$ be a deterministic binary relation which is a graph of a function in $\mathcal{F}_{h-1}$. If there exists a disjunctively well-founded transition invariant for $R$ composed of $k$-many relations with H-bounds $f_i \in \mathcal{F}_h$, $R$ is well-founded.

Over $RCA_0^* + \mathrm{Tot}(F_h)$, for every natural number $k$, $k$-TT$^h_H$ implies PH$^{h,2}_k$ [17, Theorem 7.3]. Once again we recall the argument of this implication, for the purpose of clarification.

**Theorem 5** ($RCA_0^* + \mathrm{Tot}(\mathcal{F}_h)$, [17])

$$k\text{-TT}^h_H \implies \mathrm{PH}^{h,2}_k.$$

*Proof* Assume that $c : [X]^2 \to k$ has no homogeneous set $H$ such that $\min H < |H|$. Then we define,

$$xR_iy \iff y < x \wedge x, y \in X \wedge c(\{y, x\}) = i.$$

Then let $f_i$ be the identity function on $X$. We claim it is a $H$-bound for any $R_i$. Indeed, let $\langle x_j : j \in l\rangle$ be a decreasing transitive $R_i$-sequence: by unfolding definitions, we have $x_0 < x_1 < \cdots < x_{l-1}$ and $x_b R_i x_a$ for any $0 \leq a < b < l$. If $f_i(x_0) = x_0 < l$ then we obtain a homogeneous set such that its minimum (i.e., $x_0$) is smaller than its cardinality $l$ and this is a contradiction. Then $f_i(x_0) \geq x_0 \geq l$ and this shows that the identity function $f_i$ is a $H$-bound. So, due to $k$-$\mathrm{TT}^h_H$, $\{(f(n+1), f(n)) : n \in \mathbb{N}\} \subseteq \bigcup\{R_i : i \in k\}$ should be well-founded, and this is a contradiction. □

## 2.4 Erdős' Trees

The main ingredient of the combinatorial argument provided in this paper are Erdős' trees. Erdős' trees associated to a given coloring are inspired by the trees used first by Erdős then by Jockusch [6] in their proofs of Ramsey's Theorem for pairs, hence the name.

Given binary relations $R$, $R_0$ and $R_1$ such that $R_0 \cup R_1 \supseteq R^+$, let $\sigma$ be a (finite or infinite) computation of $R$, i.e., some decreasing $R$-sequence. Define $X = \{\sigma(i) \mid i \in |\sigma|\}$[2] and $c : [X]^2 \to 2$ such that $c(\{\sigma(i), \sigma(j)\}) = 0 \iff i < j \wedge \sigma(j) R_0 \sigma(i)$. We say that a finite sequence $t$ is 1-*colored* if for all $i < j < k < |t|$ we have $c(\{t(i), t(j)\}) = c(\{t(i), t(k)\})$. For each $n \in |\sigma|$ we define the finite binary tree $Tr^\sigma_n$ in two colors, with $n+1$-nodes. Here we represent a tree as a set of finite sequences which is closed under initial segments. Informally we add $\sigma(n+1)$ at the end of the longest sequence in $Tr^\sigma_n$ such that its extension with $\sigma(n+1)$ is a 1-coloring. Formally, we define the tree $Tr^\sigma_n$ by induction over $n$:

- $Tr^\sigma_0 = \{\langle\rangle, \langle\sigma(0)\rangle\}$.
- $Tr^\sigma_{n+1} = Tr^\sigma_n \cup \{\langle\sigma(b(0)), \ldots, \sigma(b(|b|-1)), \sigma(n+1)\rangle\}$, where $b$ is the minimal sequence with respect to the lexicographic order of the set $A_{n+1} \subseteq n^{<\omega}$ defined as $t \in A_{n+1}$ if
  - $Tr^\sigma_n$ has a branch $r$ along $t$, i.e., $r = \langle\sigma(t(0)), \ldots, \sigma(t(|t|-1))\rangle \in Tr^\sigma_n$;
  - $r * \langle\sigma(n+1)\rangle = \langle\sigma(t(0)), \ldots, \sigma(t(|t|-1)), \sigma(n+1)\rangle$ is a 1-coloring;
  - $r$ is a maximal branch satisfying the latter property, i.e., for all $z \in n+1$ we have either $r * \langle\sigma(z)\rangle \notin Tr^\sigma_n$ or

$$c(\{\sigma(t(|t|-1)), \sigma(z)\}) \neq c(\{\sigma(t(|t|-1)), \sigma(n+1)\}).$$

Observe that for every $n \in |\sigma|$, $\langle\rangle \in Tr^\sigma_{n-1}$ and $\langle\sigma(n)\rangle$ is 1-colored, because it has a unique element. Therefore the set $A_n$ is not empty, it is totally ordered lexicographically, and it is finite because it is a subset of $Tr^\sigma_n$. Hence there exists a minimal sequence of $A_n$ with respect the lexicographical order. By construction $Tr^\sigma_n$ is binary, since any node of $Tr^\sigma_n$ has at most one descendant for every color.

The Erdős' tree associated to the (finite or infinite) computation $\sigma$ is the tree $\bigcup\left\{Tr^\sigma_n \mid n \in |\sigma|\right\}$. By construction any node in the Erdős' tree is connected with its descendants in the same color.

The case for $k$-many relations is analogous and provides a $k$-branching tree.

[2] If $\sigma$ is infinite put $|\sigma| = \mathbb{N}$.

## 3 Main Result

Due to the equivalence between the Termination Theorem for $k$-many relations with bounds and the Weak Paris–Harrington Theorem, and to Solovay and Ketonen's result we conclude that if we have $(k+5)$-many relations whose weight functions are in $\mathcal{F}_h$ then $R$ is bounded by a function in $\mathcal{F}_{k+5+h}$. But, as already mentioned, this bound is not optimal.

In [3], Figueira et al. provided an optimal bound for programs which admit a disjunctively well-founded transition invariant of height $\omega$ and which have an exponential control on the ranks by using the Dickson Lemma. More precisely the Dickson Lemma states that for every natural number $k$, every infinite sequence $\sigma$ of elements in $\mathbb{N}^k$ is *good*; i.e., for every infinite sequence $\sigma$ of elements in $\mathbb{N}^k$ there exist natural numbers $n < m$ such that $\sigma(n) \leq \sigma(m)$ [8, 10] (where $\leq$ is the componentwise order). Hence for every $k$, $(\leq, \mathbb{N}^k)$ is a well-quasi order. In [3] it is shown that given a control function $f : \mathbb{N} \to \mathbb{N}$ in $\mathcal{F}_h$, there exists a function $L_{k,f}$ in $\mathcal{F}_{k+h-1}$, such that the length of the *bad* (not good) sequences for which there exists a natural number $t$ such that $\forall n \in \mathbb{N}\, \forall i < k\; (\sigma(n)_i < f(n+t))$ is bounded by $L_{k,f}(t)$. If $h = 1$ then the bound is in $\mathcal{F}_k$.

Now assume given a deterministic program $P$ whose transition relation $R$ is a graph of a function in $\mathcal{F}_h$, as in the hypothesis of $k$-$\mathrm{TT}^{h+1}$. Therefore it has control function in $\mathcal{F}_{h+1}$ by Remark 3. Assume also that there exists a transition invariant composed by $k$-many primitive recursive relations $R_0, \ldots, R_{k-1}$ with weight functions $f_0, \ldots, f_{k-1} \in \mathcal{F}_{h+1}$. By mapping each state $s$ to the $k$-tuple $\sigma(s) = \langle f_0(s), \ldots, f_{k-1}(s)\rangle$, any computation $\sigma'$ of $P$ is mapped in a bad sequence. Indeed, by definition of weight function and since for $m < n$ there exists $i \in k$ such that $\sigma'(n) R_i \sigma'(m)$, we have that for any $n < m$ there exists $i \in k$ such that $f_i(\sigma'(n)) < f_i(\sigma'(m))$. Hence $\sigma(\sigma'(m)) \nleq \sigma(\sigma'(n))$. Therefore Figueira et al. provided a bound in $\mathcal{F}_{(h+1)+k-1} = \mathcal{F}_{h+k}$ for $R$.

Since we conjecture that the results above may be formalized in $RCA_0^*+$ $\mathrm{Tot}(\mathcal{F}_{h+1})$ we would extract a proof of $\mathrm{Tot}(\mathcal{F}_{k+\max\{1,h\}}) \implies \mathrm{WPH}_k^{h+1,2}$, by using Theorem 4.

Notice that the argument above does not apply if the functions $f_i$ are H-bounds instead of weight functions. Indeed in this case it is not true that the map $\sigma$ applied to computations of $P$ produces bad sequences.

Here we study an alternative argument for the H-bounded version of the Termination Theorem in order to produce a stronger version of Solovay and Ketonen's result. Indeed, as a consequence, we get that for every natural number $k \geq 2$, $\mathrm{Tot}(\mathcal{F}_{k+\max\{1,h\}}) \implies \mathrm{PH}_k^{\max\{1,h\},2}$ in $RCA_0^*$.

**Theorem 6** ($RCA_0^* + \mathrm{Tot}(\mathcal{F}_h)$) *Let $R, R_0, \ldots, R_{k-1}$, be binary relations on $\mathbb{N}$ for some $k \geq 2$ and that $h \geq 1$. Assume that:*

- *$R$ is a graph of a function in $\mathcal{F}_{h-1}$;*

- *$R_0 \cup \cdots \cup R_{k-1} \subseteq R^+$;*
- *each $R_i$ admits a H-bound in $\mathcal{F}_h$.*

*Then* $\mathrm{Tot}(\mathcal{F}_{k+h})$ *implies that $R$ is well-founded.*

To the aim of proving Theorem 6, let $R$ be the graph of a function $t \in \mathcal{F}_{h-1}$. The proof is developed in two steps: first of all we provide a bound by induction over $k$ (with basic case $k = 2$), then in Sect. 3.3 we analyze the complexity of such bound in order to prove that it belongs to $\mathcal{F}_{k+h}$.

## 3.1 Proof of the Two-Relations Case

Assume that we have two relations $R_0$ and $R_1$ such that $R_0 \cup R_1 \supseteq R^+$ and that $f_0$ and $f_1$ are their H-bounds. Given a state $s$ we want to find a bound on the number of steps we can do from $s$. Define a coloring $c : [\mathbb{N}]^2 \to 2$ such that $c(\{t^i(s), t^j(s)\}) = 0$ if and only if $i < j$ and $t^j(s) R_0 t^i(s)$.

First of all we want to find a bound on the number of steps we can do before finding an element which is connected with $s$ in color 0. We build over the proof of this result by using Erdős' tree. We use also the following property of binary trees by observing that if a binary tree has at least $2^n$ nodes, then it has some branch with $n$ edges (with $n + 1$ nodes). Given a tree $Tr$ and a node $x \in Tr$, let $Tr(x)$ be the restriction of $Tr$ with root $x$.

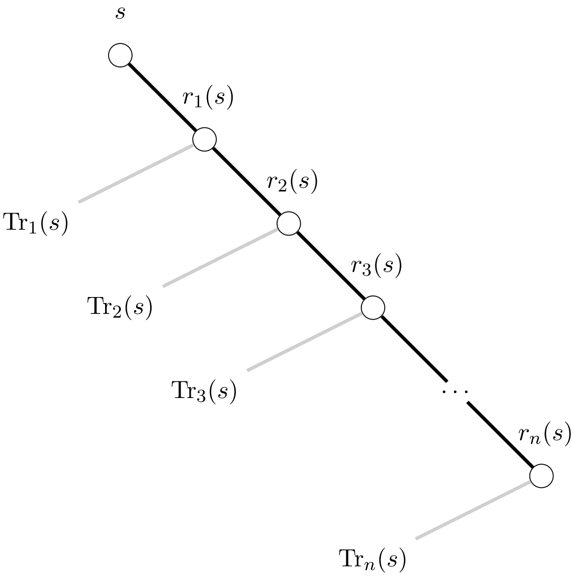
**Lemma 7** *Let $Tr$ be the Erdős' tree which corresponds to the computation which starts in some state.*

1. *Assume that $x \in Tr$ has ancestors $x_0$, $x_1$ in $Tr$ such that there is an edge in color* 0 *between $x_0$ and $x$ and there is an edge in color* 1 *between $x_1$ and $x$. Then $|Tr(x)| < 2^{f_0(x_0)+f_1(x_1)}$.*
2. *As before, assume that $x, x_0, x_1 \in Tr$ are such that $x_0$ is an ancestor in color* 0 *and $x_1$ is an ancestor in color* 1 *of $x$. If we have two subsets $I_0$ and $I_1$ of nodes such that $x_0 \in I_0$ and $x_1 \in I_1$, then $|Tr(x)| < 2^{\max\{f_0(z)|z \in I_0\}+\max\{f_1(z)|z \in I_1\}}$.*

*Proof* 1. Assume $x_0, x_1, x \in Tr$ are such that $x_0$ and $x_1$ are both ancestors of $x$, the former in color 0 and the latter in color 1. Assume by contradiction that $|Tr(x)| = 2^{f_0(x_0)+f_1(x_1)}$, then since $Tr(x)$ is a binary tree there exists a branch with $(f_0(x_0) + f_1(x_1))$-many edges. This means that in this branch there are either $f_0(x_0)$-many edges in color 0 or $f_1(x_1)$-many edges in color 1. Without loss of generality assume that we have $f_0(x_0)$-many edges in color 0. Hence if we consider the first nodes in each of these edges and the last node of the last edge we obtain a homogeneous sequence in color 0 with $f_0(x_0) + 1$-many nodes. Thus, this is a transitive $R_0$-sequence from $x_0$ such whose length is greater than $f_0(x_0)$. This is a contradiction.

2. It follows by the hypotheses and by point (1). □

Let $Tr(s)$ be the Erdős' tree whose root is $s$. For any state $s$, let $r(s)$ be the branch of $Tr(s)$ whose elements are all connected in color 1 with the root and between them. For every $n$ let $r_n(s)$ be the $n$-th node of $r(s)$, and $root_n(s)$ the child in color 0 of $r_n(s)$, if they exist. Let $Tr_n(s) = Tr(root_n(s))$ be the subtree of $Tr$ having root $root_n(s)$. The father of $Tr_n(s)$ is the $n$-th element of $r(s)$. Then $s$ is the 0-th element of $r(s)$.




We want to define a bound for the size of any $Tr_n(s)$. In order to do that we shall apply Lemma 7.1. So for every $Tr_n(s)$ we will find two nodes $x_0$ and $x_1$ in $Tr(s)$ such that they are an ancestor in color 0 and an ancestor in color 1 for any element of $Tr_n(s)$. This will provide the nodes $x_0$ and $x_1$ we need in order to apply Lemma 7.1. If we assume that any element of $Tr(s)$ is connected with $s$ in color 1 we will choose $x_1 = s$, connected to any element of $Tr_n$(s) with color 1. The natural choice for a node $x_0$ connected to any element of $Tr_n(s)$ with color 0 is $r_n(s)$. Therefore we can define a map $b_n^0(s)$ returning an upper bound for $|Tr_n(s)|$. To verify this is a bound we apply Lemma 7.2.

**Definition 8** For any state $s$, define $b_n^0(s)$ by induction on $n$:

$$b_0^0(s) = 1$$
$$b_{n+1}^0(s) = 2^{\max\left\{f_0(t^i(s)) | i \leq \sum_{j=0}^{n} b_j^0(s)\right\} + f_1(s)}.$$

**Proposition 9** *Assume that any element of the Erdős' tree is connected in color* 1 *with* $s$. *Then*

1. *for every* $n \geq 1$, *and for every* $x \in Tr_n(s)$, $x R_1 s$;

2. *for every* $n \geq 1$, $b_n^0(s)$ *is a bound for the size of* $Tr_n(s)$*;*
3. *for every* $n \geq 1$, $r_{n+1}(s) \in \left\{t^i(s) \mid i \leq \sum_{j=0}^{n} b_j^0(s)\right\}$*; i.e., after at most* $\sum_{j=0}^{n} b_j^0(s)$*-many steps we find* $r_{n+1}(s)$*, the* $(n+1)$*-th element of* $r(s)$*.*

*Proof* By induction over $n$. Assume that $n = 1$.

1. It follows since any $x$ is connected to $s$ in color 1.
2. Observe that

$$r_1(s) = t(s) \in \left\{t^i(s) \mid i \leq 1\right\}.$$

   Due to point (1) above we can apply Lemma 7.2 for $I_0 = \left\{t^i(s) \mid i \leq 1\right\}$ and $I_1 = \{s\}$. Hence we get $|Tr_1(s)| < b_1^0(s)$.
3. Due to point (2) above, in $b_0^0(s) + b_1^0(s)$ many steps we find the second element of $r(s)$.

Assume now that the thesis holds for $n$ and we prove it for $n+1$. So assume that $b_n^0(s)$ is a bound for $Tr_n(s)$ and that after $(\sum_{j=0}^{n} b_j^0(s))$-many steps we find the $(n+1)$-th element of $r(s)$.

1. Again it follows since any $x$ is connected to $s$ in color 1.
2. Due to the inductive hypothesis we have

$$r_n(s) \in \left\{t^i(s) \mid i \leq \sum_{j=0}^{n-1} b_j^0(s)\right\}$$

   Again, by using point (1) and this remark, we can apply Lemma 7.2 by putting $I_0 = \left\{t^i(s) \mid i \leq \sum_{j=0}^{n-1} b_j^0(s)\right\}$ and $I_1 = \{s\}$. Hence we get the thesis.
3. Due to point (2), in less than $\sum_{j=0}^{n+1} b_j^0(s)$-many steps we complete each $Tr_i(s)$ for every $i < n+2$ and so, since by hypothesis we are assuming that any element is connected to $s$ in color 1, we are forced to add a new element $r_{n+2}(s)$ of $r(s)$. □

Now we study a bound for the size of the whole Erdős' tree.

**Definition 10** Put

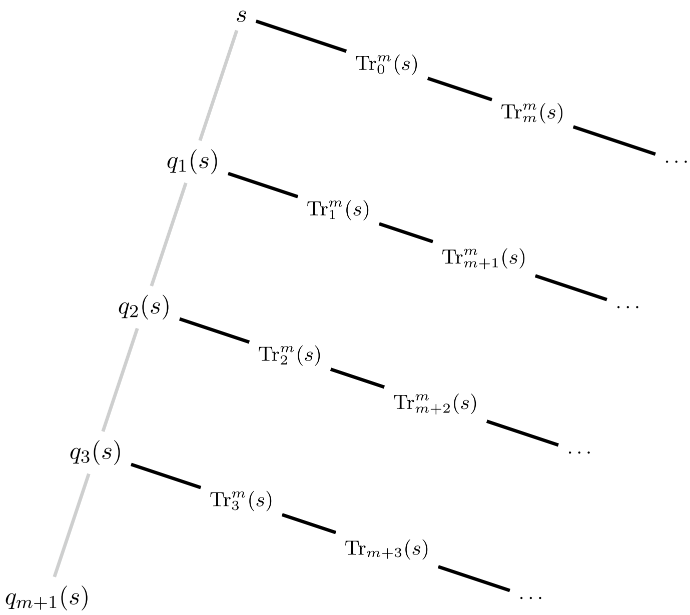
$$b_0^{m+1}(s) = \sum_{i=0}^{m \cdot \max\left\{f_1(t^i(s)) \mid i < b_0^m(s)\right\}} b_i^m(s)$$

$$b_{n+1}^{m+1}(s) = 2^{\max\left\{f_0(t^i(s)) \mid i \leq \sum_{j=0}^{n} b_j^{m+1}(s)\right\}} \cdot 2^{\max\left\{f_1(t^i(s)) \mid i < b_0^m(s)\right\}}$$

$$f_2(s) = b_0^{f_0(s)}(s).$$

Observe that $b_0^1(s)$ is a bound for the number of steps required to find an element which is connected to $s$ in color 0. While we will prove $f_2(s)$ is a bound for the whole computation as guaranteed by the following results.

Let $q(s)$ be the branch of $Tr(s)$ whose elements are all connected in color 0 with the root and between them. Let $q_m(s)$ be the $m$-th node of $q(s)$.




Let $r_n^m(s)$ be the ancestor in color 0 of the root of $Tr_n^m(s)$ and let $\phi : \mathbb{N} \to \left\{r_n^m \mid n \in \mathbb{N}\right\}$ be such that $\phi(l)$ is the $l$-th element of $\left\{r_n^m \mid n \in \mathbb{N}\right\}$ which appears in the computation. As a corollary of the following result we get a bound for $\phi(n)$.

**Proposition 11** *For all $m \in \mathbb{N}$.*

1. *$q_m(s) \in \left\{t^i(s) \mid i < b_0^m(s)\right\}$; i.e., after $b_0^m(s)$-many steps we find $q_m(s)$, the $m+1$-th element of $q(s)$.*
2. *For all $n$, $b_{n+1}^m(s)$ is a bound for $Tr_{\phi(n)}^m(s)$.*

*Proof* By induction on $m$.

- If $m = 0$. It follows by Proposition 9.
- Assume that the induction hypothesis holds for $m$. We prove it for $m + 1$.

  1. Observe that by induction hypothesis for all $j \in m + 1$,

$$f_1(q_j(s)) < \max\left\{f_1(t^i(s)) \mid i < b_0^m(s)\right\}.$$

Therefore $\max\left\{f_1(t^i(s)) \mid i < b_0^m\right\}$ is a bound for the homogeneous 1-branch which starts in $q_j(s)$ for all $j < m + 1$. Thus

$$(m+1)\max\left\{f_1(t^i(s)) \mid i < b_0^m(s)\right\}$$

is a bound for the number of trees of the form $Tr_l^m(s)$. Since by induction hypothesis on point (2) every $Tr^m_{\phi(n)}(s)$ is bounded by $b^m_{n+1}(s)$, the thesis follows.

2. By induction on $n$. If $n = 0$, then by point (1), there exists $i$ such that $t^i(s) = q_{m+1}(s)$ and $b_0^{m+1}(s) \geq i + 1$. Then the root of $Tr^{m+1}_{\phi(0)}(s)$ appears in $b_0^{m+1}(s)$-many steps. Therefore, following the proof of Proposition 9, $b_1^{m+1}(s)$ is a bound for $Tr^{m+1}_{\phi(0)}(s)$.
   Assume that the thesis holds for $n$. By induction hypothesis in $b_n^{m+1}(s)$-many steps we find a bound for $\left|\bigcup\left\{Tr^{m+1}_{\phi(i)}(s) \mid i < n\right\}\right|$. Therefore the root of $Tr^{m+1}_{\phi(n)}(s)$ appears in $b_n^{m+1}(s)$-many steps. Again as shown in Proposition 9, $b^{m+1}_{n+1}(s)$ is a bound for $Tr^{m+1}_{\phi(n)}(s)$. □

**Corollary 12** *$f_2(s)$ is a bound for the number of nodes of the Erdős' tree associated to the computation which starts in $s$.*

*Proof* Since $f_0(s)$ is a bound for the length of $q(s)$, by Proposition 11 we get that $f_2(s)$ is a bound for the size of $Tr(s)$.

This yields that $f_2(s)$ is a bound for the length of the computation of $R$ in $s$, hence $R$ is well-founded. In Sect. 3.3 we analize the complexity of $f_2$, to show that, for every $h > 0$, $f_2 \in \mathcal{F}_{h+2}$.

### 3.2 *Proof of the* $(k+1)$*-Relations Case*

Assume we proved that the bound for $k$-many colors and weight functions in $\mathcal{F}_h$ is given by $f_k \in \mathcal{F}_{h+k}$, we want to prove that the bound for $k+1$-many colors and weight functions in $\mathcal{F}_h$ is in $\mathcal{F}_{h+k+1}$. We define $b_n^m(s)$ as before by putting $f_1 = f_k$ and then a possible bound we obtain is $f_{k+1}(s) = b_0^{f_0(s)}(s)$.

### 3.3 *Complexity*

We analyze the complexity of the bounds. Fix $h > 0$, we prove by induction on $k$ that for every $k > 1$, $f_k \in \mathcal{F}_{h+k}$. Assume that we have $f_0 \in \mathcal{F}_h$, $f_k \in \mathcal{F}_{h+k}$, with $k \geq 1$, in order to prove that $f_{h+k+1} \in \mathcal{F}_{h+k+1}$ (the case $k = 1$ is the basic case of the induction, where $f_k$ is the given $f_1 \in \mathcal{F}_h$).

**Claim** *For every $m$, the function $f_m(n, s) := \sum_{i=0}^{n} b_i^m(s)$ is in $\mathcal{F}_{h+k}$.*

*Proof* We prove, by induction on $m$ that $f_m$ can be defined by limited recursion by using functions in $\mathcal{F}_{h+k}$. By applying Proposition 1.2, let $l_0, l_1, l_2 \in \mathbb{N}$ be such that $t(x) \leq F^{l_0}_{h-1}(x)$, $f_k(x) \leq F^{l_1}_{h+k}(x)$, and $f_0(x) \leq F^{l_2}_h(x)$. Define

$$g_m(x) = F^{l_1}_{h+k}(F^{l_0 b^m_0(s)}_{h-1}(x))$$

$$u(x) = 2F_2(2(F^{l_2+1}_h(l_0(x)).$$

Then we can define $f_m$ by limited recursion:

$$\begin{aligned} f_m(0,s) &= b^m_0(s) \\ f_m(n+1,s) &= f(n,s) + 2^{\max\left\{f_0(t^i(s))|i \le f_m(n,s)\right\}+\max\left\{f_k(t^i(s))|i<b^m_0(s)\right\}} \\ f_m(n,s) &\le u^n(g(s)). \end{aligned}$$

We claim that the third inequality holds. We prove it by induction on $n$. Assume that $f_m(n,s) \le u^n(g(s))$. Then

$$\begin{aligned} f_m(n+1,s) &\le f_m(n,s) + 2^{\max\left\{f_0(t^i(s))|i\le f_m(n,s)\right\}+\max\left\{f_k(t^i(s))|i<b^m_0(s)\right\}} \\ &\overset{\text{(Prop.1.1)}}{\le} f_m(n,s) + 2^{\max\left\{F^{l_2}_h(F^{l_0\,i}_{h-1}(s))|i\le f_m(n,s)\right\}+\max\left\{F^{l_1}_{h+k}(F^{l_0\,i}_{h-1}(s))|i\le b^m_0(s)\right\}} \\ &\overset{\text{(Prop.1.1)}}{\le} f_m(n,s) + 2^{F^{l_2}_h(F^{l_0 f_m(n,s)}_{h-1}(s))+g_m(s)} \\ &\overset{\text{(Ind. Hyp.)}}{\le} f_m(n,s) + 2^{F^{l_2}_h(F^{l_0 u^n(g_m(s))}_h(s))+g_m(s))} \\ &\overset{\text{(Prop.1.1)}}{\le} f_m(n,s) + 2^{2F^{l_2}_h(F^{l_0 u^n(g_m(s))}_{h-1}(s))} \\ &\overset{\text{(Cor.2)}}{\le} f_m(n,s) + F_2(2(F^{l_2}_h(F_h(l_0 \cdot u^n(g_m(s)))))) \\ &\overset{\text{(Ind. Hyp.)}}{\le} 2F_2(2(F^{l_2+1}_h(l_0 \cdot u^n(g_m(s))))) \\ &\overset{\text{(Definition } u)}{=} u(u^n(g_m(s))). \end{aligned}$$

Therefore the third inequality holds. Now assume that $m = 0$. Then $b^0_0(s) = 1 \in \mathcal{F}_{h+k}$, hence $g_0 \in \mathcal{F}_{h+k}$. Thus, by the closure of $\mathcal{F}_{h+k}$ under limited recursion, also $f_0$ is. Now assume by induction hypothesis that $f_m \in \mathcal{F}_{h+k}$. We prove that this holds also for $b^{m+1}_0$.

By definition $b^{m+1}_0(s) = \sum_{i=0}^{m\cdot\max\left\{f_k(t^i(s))|i<b^m_0(s)\right\}} b^m_i(s)$. Hence by exploiting the definition of $f_m$, we have $b^{m+1}_0(s) \le f_m(mF^{l_1}_{h+k}(F^{l_0 b^m_0(s)}_{h-1}(s)), s)$. Since $f_m \in \mathcal{F}_{h+k}$, also $b^{m+1}_0$ is. As above, both $g_{m+1}$ and $f_{m+1}$ belong to the same class.

**Claim** *The function* $f_{k+1}(s) = b^{f_0(s)}_0(s)$ *is in* $\mathcal{F}_{h+k+1}$.

*Proof* Let $u$ be the function defined in the proof of the previous claim. Since $u \in \mathcal{F}_h$, then let $l_3$ be such that $u(x) \le F^{l_3}_{\max\{h,2\}}(x)$ (Proposition 1.2). Define

$$v(x) = F^{l_3}_{\max\{h,2\}}(mF^{l_1}_{h+k}(F_h(l_0x^2)+1)x)$$

Then we have:

$$
\begin{aligned}
f_{m+1}(0,s) &\le f_m(mF_{h+k}^{l_1}(F_{h-1}^{l_0 b_0^m(s)}(s)), s) \\
&\le u^{mF_{h+k}^{l_1}(F_{h-1}^{l_0 b_0^m(s)}(s))}(g_m(s)) \\
&\overset{(\text{Def.}g_m)}{=} u^{mF_{h+k}^{l_1}(F_{h-1}^{l_0 b_0^m(s)}(s))+1}(s) \\
&\overset{(\text{Cor.2})}{\le} F_{\max\{h,2\}}^{l_3}(mF_{h+k}^{l_1}(F_h(l_0 b_0^m(s)s)+1)s) \\
&\overset{(\text{Def.}f_m)}{\le} F_{\max\{h,2\}}^{l_3}(mF_{h+k}^{l_1}(F_h(l_0 f_m(0,s)s)+1)s) \\
&\overset{(\text{Def.}v)}{\le} v(\max\{f_m(0,s), s\}).
\end{aligned}
$$

By induction on $m$ we derive that

$$f_{m+1}(0,s) \le v^m(\max\{f_0(0,s), s\}) = v^m(s).$$

Therefore $f_{k+1}(s) = b_0^{f_0(s)}(s) = f_{f_0(s)}(0,s) \le v^{f_0(s)}(s)$ is in $\mathcal{F}_{\max\{h+k+1,3\}}$, since $v \in \mathcal{F}_{\max\{h+k,2\}}$. □

Since $h \ge 1$ and $k \ge 2$ we have a bound in $\mathcal{F}_{\max\{h+k,3\}} = \mathcal{F}_{h+k}$ for $f_k$.

## 4 Conclusion

The result in this paper guarantees that over $RCA_0^*$+ $\text{Tot}(\mathcal{F}_{k+h+1})$ we can prove $k$-$\text{TT}_H^{h+1}$. By using Theorem 5 we get the following for any natural numbers $h$ and $k \ge 2$.

**Corollary 13** *($RCA_0^*$)* $\text{Tot}(\mathcal{F}_{k+h+1}) \implies \text{PH}_k^{h+1,2}$.

In particular for every natural number $k \ge 1$, $\text{Tot}(\mathcal{F}_{k+1}) \implies \text{PH}_k^{1,2}$. Over $RCA_0^*$ we have:

$$\text{Tot}(\mathcal{F}_3) \ge \text{PH}_2^{1,2} \ge \text{WPH}_k^{1,2}.$$

Figueira et al. in [3] proved that the bound provided for the miniaturization of the Dickson Lemma is optimal; i.e., there are examples of programs with control functions in $\mathcal{F}_{h+1}$ and a transition invariant composed of $k$-many relations with weight functions in $\mathcal{F}_{h+1}$ for which the computations cannot be bounded by a function in $\mathcal{F}_{k+h-1}$.

*Example 14* Consider the following program.

```
while  ( x > 0 and y > 0 )
if(y > 1) (x, y, z) = (x, y - 1, 2 * z)
else      (x, y, z) = (x - 1, 2 * z, 2 * z)
```

A transition invariant for this program is $R_1 \cup R_2$, where

$$R_1 = \left\{ (\langle x', y', z' \rangle, \langle x, y, z \rangle) \mid y > 0 \wedge y' < y \right\}$$
$$R_2 = \left\{ (\langle x', y', z' \rangle, \langle x, y, z \rangle) \mid x > 0 \wedge x' < x \right\}.$$

$R_1$ and $R_2$ are bounded by $F_0$. By [3] and since $R$ is the graph of function in $\mathcal{F}_1$, $R$ is bounded in $\mathcal{F}_{2+1}$. It is straightforward to prove that such bound is optimal since for any $x \geq y > 0$, the computation which starts in $(x, y, 1)$ has length greater than $F_2^{x-2}(y)$.

By Example 14, we know that we cannot have the implication from $\text{Tot}(\mathcal{F}_2)$ to $k$-$\text{TT}^2$ over $RCA_0^*$. Thus $\text{Tot}(\mathcal{F}_2)$ cannot prove $k$-$\text{TT}^2_H$ in $RCA_0^*$.

**Question 15** Does $\text{WPH}_k^{1,2}$ imply $\text{Tot}(\mathcal{F}_{k+1})$ (and therefore $\text{PH}_k^{1,2}$) over $RCA_0^*$?

By Theorem 6 over $RCA_0^*$ for every $h \geq 1$, $\text{Tot}(\mathcal{F}_{k+\max\{h,1\}})$ implies $k$-$\text{TT}^h_H$ (and so $\text{PH}_k^{h,2}$). Moreover, by [3], $\text{Tot}(\mathcal{F}_{k+\max\{h,1\}})$ implies $k$-$\text{TT}^{h+1}$ (and so $\text{WPH}_k^{h+1,2}$). By the example above, we know that our bound for $h = 0$ is strict. We may wonder whether our bound is strict also for $h > 0$, and, in particular, if we can prove a converse for Theorem 6.

**Question 16** Does $\text{Tot}(\mathcal{F}_{k+h})$ imply $\text{PH}_k^{h+1,2}$ over $RCA_0^*$, for every $h > 0$? If it is not the case, does $\text{PH}_k^{h+1,2}$ imply $\text{Tot}(\mathcal{F}_{k+h+1})$?

Actually the version of the Paris–Harrington Theorem we deal with is different from the one provided in [5]. Anyway we can use also the statement of the Paris–Harrington Theorem introduced in [5] in our argument. In particular we have that $\text{Tot}(\mathcal{F}_{k+1})$ implies that for every $a \in \mathbb{N}$ there exists $b \in \mathbb{N}$ such that for every coloring $c : [[a, b]]^2 \to k$ there exists a finite homogeneous set $H \subseteq [a, b]$ such that $\min H < |H|$.

Indeed define the functions $t, f_0, \dots, f_{k-1}$ such that $t(x) = x + 1$ for any $x \in \mathbb{N}$ and $f_0(x) = \dots = f_{k-1}(x) = x$ for any $x \in \mathbb{N}$. Thus we can compute the bound $f_k^*$ given in the proof of Theorem 6. Since all these functions are in $\mathcal{F}_0$, we proved that $f_k^* \in \mathcal{F}_{k+1}$. For every $a$, let $b = f_k^*(a) + 1$, we can choose such $b$ since we assumed $\text{Tot}(\mathcal{F}_{k+1})$. Suppose by contradiction that there exists a coloring $c : [[a, b]]^2 \to k$ without finite homogeneous sets $H \subseteq [a, b]$ such that $\min H < |H|$. Then by using the argument of Theorem 5 we can define for every $i \in k$ the relation:

$$x R_i y \iff c(\{x, y\}) = i \wedge y < x \wedge x, y \in [a, b].$$

The identity function is a H-bound for every $R_i$. Indeed every transitive $R_i$-sequence whose length is greater than the first element would provide an homogeneous set $H$ with $\min H < |H|$. But this is impossible by hypothesis on $c$, hence $R_i$ is H-bounded by the identity function. Let $R = \{(x + 1, x) : x \in [a, b)\}$. By construction, $R$ is bounded by the $f_k^*$ defined above. This is a contradiction since the $R$-decreasing sequence which starts in $a$ has length $f_k^*(a) + 1$.

Therefore we get that for every $a \in \mathbb{N}$ there exists $b \in \mathbb{N}$ (namely $f_k^*(a) + 1$) such that for every coloring $c : [[a, b]]^2 \to k$ there exists a finite homogeneous set $H \subseteq [a, b]$ such that $\min H < |H|$.

Now let $(\text{PH})_n$ be the statement: for every natural number $k$, for every $a \in \mathbb{N}$ there exists $b \in \mathbb{N}$ such that for every coloring $c : [[a, b]]^{n+1} \to k$ there exists a finite homogeneous set $H$ such that $\min H < |H|$. Hájek and Pudlák in [5, Problem 3.37] proposed the following problem: find a reasonably simple proof of $\text{I}\Sigma_1 \vdash (W)_n \implies (\text{PH})_n$, where $(W)_n$ is a principle equivalent to the statement[3] $\forall k \forall \alpha < \omega_{n-1}^k \text{Tot}(F_\alpha)$. The argument above guarantees that $RCA_0^* \vdash \text{Tot}(\mathcal{F}_{k+1}) \implies \text{PH}_k^2$. However, since the proof of the bound is by induction over $k$, we cannot conclude directly that $RCA_0^* \vdash \forall k \text{Tot}(\mathcal{F}_{k+1}) \implies \forall k \text{PH}_k^2$. Hence:

**Question 17** Can we use a similar argument to prove $(W)_1 \implies (\text{PH})_1$ in $\text{I}\Sigma_1$?

**Question 18** Is it possible to generalize our argument to prove $\text{PH}_k^n$ from $\text{Tot}(\mathcal{F}_{\omega_{n-2}^{k+1}})$, for $n > 2$?

**Acknowledgements** I would like to thank my supervisor Stefano Berardi, Alexander Kreuzer, Sylvain Schmitz, Keita Yokoyama and the anonymous referee for their valuable suggestions and insightful remarks.

## References

1. Berardi, S., Oliva, P., & Steila, S. (2015). An analysis of the Podelski–Rybalchenko termination theorem via bar recursion. *Journal of Logic and Computation*, *24*.
2. Berardi, S., & Steila, S. (2015). An intuitionistic version of Ramsey's Theorem and its use in Program Termination. *Annals of Pure and Applied Logic*, *166*(12), 1382–1406.
3. Figueira, D., Figueira, S., Schmitz, S., & Schnoebelen, P. (2011). Ackermannian and primitive-recursive bounds with Dickson's Lemma. In *LICS 2011: Proceedings of the 26th Annual IEEE Symposium on Logic in Computer Science* (pp. 269–278). IEEE Press.
4. Gasarch, W. I. (2015). Proving programs terminate using well-founded orderings, Ramsey's Theorem, and matrices. *Advances in Computers*, *97*, 147–200.
5. Hájek, P., & Pudlák, P. (1998). *Chapter II: fragments and combinatorics*. Perspectives in Mathematical Logic (Vol. 3, pp. 111–144). Springer, Berlin.
6. Jockusch, C. G, Jr. (1972). Ramsey's Theorem and recursion theory. *Journal of Symbolic Logic*, *37*, 268–280.
7. Ketonen, J., & Solovay, R. (1981). Rapidly growing ramsey functions. *Annals of Mathematics*, *113*(2), 267–314.
8. Kruskal, J. B. (1972). The theory of well-quasi-ordering: A frequently discovered concept. *Journal of Combinatorial Theory, Series A*, *13*(3), 297–305.
9. Löb, M. H., & Wainer, S. S. (1970). Hierarchies of number-theoretic functions. *I. Archive for Mathematical Logic*, *13*(1–2), 39–51.
10. Milner, E. C. (1985). Basic WQO- and BQO-theory. In Rival, I. (Ed.), *Graphs and order*. NATO ASI Series (Vol. 147, pp. 487–502). Netherlands: Springer.

[3] For any natural numbers $n$ and $k$ $\omega_0^k = \omega^k$ and $\omega_{n+1}^k = \omega^{\omega_n^k}$. By adding $F_\alpha(x) = F_{\{\alpha\}(x)}(x+1)$ to the definition of Fast Growing Hierarchy we obtain the definition of Schwichtenberg-Wainer used in [5]. We refer to [5] for details.

11. Paris, J. B., & Harrington, L. (1977). A Mathematical incompleteness in Peano arithmetic. In J. Barwise (Ed.), *Handbook of mathematical logic, North-Holland* (pp. 1133–1142). Amsterdam:
12. Podelski, A., & Rybalchenko, A. (2004). Transition invariants. In *Proceedings of the 19th IEEE Symposium on Logic in Computer Science (LICS 2004), 14–17 July 2004, Turku, Finland* (pp. 32–41). IEEE Computer Society.
13. Ramsey, F. P. (1930). On a problem in formal logic. *Proceedings of the London Mathematical Society*, *3*(30), 264–286.
14. Simpson, S. G. (1999). *Subsystems of second order arithmetic*. Perspectives in mathematical logic. Springer.
15. Specker, E. (1971). Ramsey's theorem does not hold in recursive set theory. In *Logic Colloquium '69 (Proceedings of the Summer School and Colloquium in Mathematical Logic, Manchester, 1969)*, North-Holland, Amsterdam (pp. 439–442).
16. Steila, S. (2016). Terminating via Ramsey's Theorem. Ph.D. Thesis, Università degli studi di Torino.
17. Steila, S., & Yokoyama, K. (2016). Reverse mathematical bounds for the Termination Theorem. *Annals of Pure and Applied Logic*, *167*(12), 1213–1241.
18. Vytiniotis, D., Coquand, T., Wahlstedt, D. (2012). Stop when you are almost-full–adventures in constructive termination. In Beringer, L. & Felty, A.P. (Eds.), *Proceedings of the Third International Conference on Interactive Theorem Proving, ITP 2012, Princeton, NJ, USA, August 13–15, 2012*. Lecture Notes in Computer Science (Vol. 7406, pp. 250–265). Springer.

# A Mechanized Proof of Higman's Lemma by Open Induction

**Christian Sternagel**

**Abstract** I present a short, mechanically checked Isabelle/HOL formalization of Higman's lemma by open induction.

## 1 Introduction

In the winter of 2016 a mixed group of scientists met for a week in Dagstuhl, Germany, to discuss the present and future of *Well Quasi-Orders in Computer Science*.[1] Having worked a little on mechanizing results from well-quasi-order theory with the proof assistant Isabelle in the past, I was for a time thinking hard about any new results I could present. Then I remembered a clingy item on my mental to-do list: applying a previous Isabelle formalization of open induction to obtain an alternative mechanization of Higman's lemma. The following exposition is supposed to give an accessible account of my formalization.

The study of *well-quasi-orders* dates back at least to the early 1940s. (And already in the 1970s, a tendency to duplicate work prompted Kruskal to give an introductory overview of well-quasi-orders, including their history, present, and future [10].)

While initially, the goal was mostly to show that certain structures (like pairs, finite words [8], and finite trees [9]) are indeed well-quasi-ordered, later on a significant amount of work was invested into obtaining shorter/simpler/more elegant proofs of known results. A prime example of this kind of work is Nash-Williams' *minimal bad sequence* argument [14] (which allowed him to shorten the previous, rather involved 7-page proof by Kruskal [9] down to a conceptually simple half-page proof).

Despite the elegance of Nash-Williams' proof, some consider its non-constructive nature a major drawback. Therefore, another line of work focuses on constructive proofs of results in well-quasi-order theory [3, 13, 18]. Also, in order to obtain

[1]Dagstuhl Seminar 16031: www.dagstuhl.de/16031.

C. Sternagel (✉)
University of Innsbruck, Technikerstraße 21a, 6020 Innsbruck, Austria
e-mail: christian.sternagel@uibk.ac.at



P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_12

insight into the *computational content* of a proof—in essence, the goal is to obtain an algorithm directly from a proof, a process that is also known as *program extraction*.

A more recent branch of research is dedicated to the mechanization of results from well-quasi-order theory with the help of proof assistants [2, 6, 11, 12, 19, 21]. Such machine-checked proofs, while often hard to establish—are highly trustworthy and have also other advantages, like the possibility to machine-generate verified programs (eliminating the more traditional but potentially error-prone approaches of either manually transforming an existing computer program into a machine-readable specification or manually writing a program adhering to an existing specification).

While in practice investigations of the computational content of a proof and its formalization often go hand in hand, I want to stress that neither is being constructive a prerequisite for a formalization, nor is having a formalization a prerequisite for investigating the computational content of a proof. Which is why I distinguish between these two goals above.

This work is part of an effort towards combining the above three strands of research by simplifying existing proofs, investigating their computational content, and providing corresponding mechanizations. My starting point was an existing formalization of well-quasi-order theory [20] in Isabelle/HOL (by myself), employing the minimal bad sequence argument, together with the idea (of others) that a classically equivalent but more "computational" way of expressing the same kind of reasoning is via a proof method called *open induction* [17], whose computational interpretation was investigated by Berger [1]. My main result is a new mechanized proof of Higman's lemma by open induction.

Below, we repeat the (semi-)formal statement of Higman's well-known result [8, Theorem 4.3] (where $A^*$ denotes the set of finite words over an alphabet $A$).

**Higman's Lemma**. *If* $A$ *is well-quasi-ordered, then so is* $A^*$.

As another new result I provide a mechanized equivalence proof between the classical definition of almost-full relations and a more recent inductive definition due to Vytiniotis et al. [25].

Isabelle is a generic interactive proof assistant. Its most popular incarnation is Isabelle/HOL [15] (for higher-order logic) which is a classical logic based on Church's simply typed lambda calculus and with Hilbert's choice operator built in. Every mechanized proof is ultimately broken down to the handful of basic axioms of HOL—where every single step of this reduction is machine checked—yielding a very high degree of reliability.

*Related Work*. A similar proof, but without an accompanying formalization and yet more involved, was presented in an unpublished manuscript by Geser [7].

*Overview*. I start, in Sect. 2, by recalling the fundamental notions of almost-full relations and well-quasi-orders and also give some basic results. This includes a new proof of the fact that almost-full relations admit a (classically) equivalent inductive definition. Then, in Sect. 3, it is shown how the proof principle of open induction almost naturally arises when searching for a constructive counterpart to proofs by minimal counterexample. Afterwards, in Sect. 4, the stage is set by discharging the prerequisites of open induction one at a time. My main result, a proof of Higman's lemma by open induction, is presented in Sect. 5. Finally, I conclude in Sect. 6.

## 2 Preliminaries

Let us start by recalling the formal definition of well-quasi-orders, based on the, probably less well known, notion of almost-full relations. (The notion of almost-full relations was introduced by Veldman and Bezem [24]; another very accessible exposition is given by Vytiniotis et al. [25].)

**Definition 1** (*Almost-Full Relations and Well-Quasi-Orders*) Let $\sqsubseteq$ be a binary relation with domain $A$. An infinite sequence $a_1, a_2, a_3, \ldots$ of elements in $A$ (or infinite $A$-sequence for short) is ($\sqsubseteq$-)*good* if it contains an "increasing pair," that is, $a_i \sqsubseteq a_j$ for some $i < j$. Infinite $A$-sequences that do not satisfy this condition are called ($\sqsubseteq$-)*bad*. A relation $\sqsubseteq$ is *almost-full* (on $A$) if all infinite $A$-sequences are good. If in addition $\sqsubseteq$ is a quasi-order (on $A$),[2] it is called a *well-quasi-order* (on $A$).

A nice property of almost-full relations is that in combination with transitivity, we obtain well-foundedness for free. Therefore almost-full relations and well-quasi-orders are of special interest for proving termination (of programs, term rewrite systems, etc.; which is also my angle on the subject).

**Lemma 1** *Every transitive extension $\preceq$ of an almost-full relation $\sqsubseteq$ is well-founded.*

*Proof* Assume to the contrary that there is an infinite descending sequence $a_1 \succ a_2 \succ a_3 \succ \ldots$ (where $x \succ y$ iff $x \succeq y$ and $x \not\preceq y$). By transitivity, we obtain $a_i \succ a_j$ for all $i < j$. But then also $a_i \not\sqsubseteq a_j$ for all $i < j$, and thus the sequence above is $\sqsubseteq$-bad, contradicting the assumption that $\sqsubseteq$ is almost-full. □

It turns out that it is often easy to extend results about almost-full relations to well-quasi-orders (remember that the latter differ from the former only by transitivity), which is an indication that "being almost-full" somehow captures the essence of "being a well-quasi-order."

In my initial presentation of Higman's lemma above, the well-quasi-order on finite words was left implicit. Let us amend this omission with the following definition.

**Definition 2** (*Homeomorphic Embedding*) Given a binary relation $\sqsubseteq$, the induced *(homeomorphic) embedding relation* on finite words is defined inductively by the following three clauses (where finite words are constructed from the empty word $[\,]$ together with the binary constructor $\cdot$ which puts a single letter in front of another finite word):

$$\frac{}{[\,] \sqsubseteq^* ys} \qquad \frac{xs \sqsubseteq^* ys}{xs \sqsubseteq^* y \cdot ys} \qquad \frac{x \sqsubseteq y \quad xs \sqsubseteq^* ys}{x \cdot xs \sqsubseteq^* y \cdot ys}$$

[2] In fact, demanding transitivity suffices, since reflexivity is immediate for almost-full relations.

Now, a more explicit version of Higman's lemma is as follows.

**Theorem 1** (Higman's Lemma) *Given a well-quasi-order* $\sqsubseteq$ *on* $A$*, the induced embedding relation* $\sqsubseteq^*$ *is a well-quasi-order on* $A^*$*.*

Incidentally, the above statement is already true when replacing all occurrences of "*a well-quasi-order*" by "*an almost-full relation.*" Which is to say that transitivity does not pose any additional difficulties.

I conclude this section by some (classically) equivalent definitions of almost-full relations. (For well-quasi-orders more equivalences are known.)

**Lemma 2** *Given a relation* $\sqsubseteq$ *on* $A$*, the following statements are equivalent:*

*(1) The relation* $\sqsubseteq$ *is almost-full on* $A$*.*
*(2) Every infinite* $A$*-sequence* $a$ *admits a* $\sqsubseteq$*-homogeneous subsequence, that is, there is a strictly monotone mapping* $\sigma : \mathbb{N} \to \mathbb{N}$ *such that* $a_{\sigma(i)} \sqsubseteq a_{\sigma(j)}$ *for all* $i < j$*.*
*(3) The relation* $\sqsubseteq$ *satisfies the predicate* $\mathsf{af}(\cdot)$ *which is defined inductively by the two clauses:*

$$\frac{\forall x, y \in A.\, x \sqsubseteq y}{\mathsf{af}(\sqsubseteq)} \qquad \frac{\forall x \in A.\, \mathsf{af}(\lambda y\ z.\, y \sqsubseteq z \vee x \sqsubseteq y)}{\mathsf{af}(\sqsubseteq)}$$

Property (3) above, is due to Vytiniotis et al. [25] and gives a nice intuition why such relations are called "almost full": it is possible within a finite number of steps (since the definition is inductive) to turn them into full relations (which is the only base case).

*Proof* Detailed Isabelle/HOL proofs of the above equivalences are available in the *Archive of Formal Proofs* [20, `Almost_Full.thy`]. Their basic outline follows.

For the implication from (1) to (2), consider the infinite 2-colored graph whose vertices are the natural numbers such that $i$ and $j$ are connected by an edge with color 0 if and only if $a_i$ and $a_j$ are related by $\sqsubseteq$ (in either direction) and by an edge with color 1, otherwise. An application of Ramsey's theorem yields an infinite homogeneously colored subgraph. Since an infinite 1-subgraph contradicts the fact that $\sqsubseteq$ is almost-full, an infinite 0-subgraph is obtained. Enumerating the corresponding indices in increasing order yields the desired homogeneous subsequence of $a$. Since above, we employ the classical infinite Ramsey, the implication from (1) to (2) only holds classically.

Also the implication from (2) to (3) only holds classically. For the sake of a contradiction, let us first assume that (3) does not hold, and then construct a counterexample to (2). To this end, let $\mathsf{NAF}_{\preceq}$ denote some $x \in A$ such that $\mathsf{af}(\lambda y\ z.\, y \preceq z \vee x \preceq y)$ does not hold (which is obtained using Hilbert's choice operator in Isabelle/HOL). Then construct an infinite sequence $c^{\preceq}$ such that $c^{\preceq}_1$ is $\mathsf{NAF}_{\preceq}$ and $c^{\preceq}_{i+1}$ is $c^{\preceq'}_i$ with $\preceq' = (\lambda y\ z.\, y \preceq z \vee \mathsf{NAF}_{\preceq} \preceq y)$ for all $i \geq 1$. In the following $c$ abbreviates $c^{\sqsubseteq}$. Now, from the assumption $\neg\, \mathsf{af}(\sqsubseteq)$, it is shown by induction on $n$ that

$$\mathsf{af}\left(\lambda y\ z.\ y \sqsubseteq z \vee \bigvee_{i \leq n} c_i \sqsubseteq y \vee \bigvee_{1 \leq i < j \leq n} c_i \sqsubseteq c_j\right)$$

does not hold for any $n$, contradicting the fact that $c$ admits an infinite homogeneous subsequence.

The proof of the implication from (3) to (1) proceeds by an easy rule-induction according to the definition of $\mathsf{af}(\cdot)$. □

## 3 From Minimal Counterexamples to Open Induction

Before I give the necessary prerequisites for proving Higman's lemma, let us discuss how *open induction* enters the picture. Since, ideally we want to have a simple, constructive, and formalized proof, I thought it a good idea to start from Nash-Williams proof (which is way simpler than any other proof I am aware of). His proof proceeds by contradiction: assuming that there is a bad sequence, he then argues that there is a minimal one, and finally constructs an even smaller bad sequence. This corresponds to a proof by contradiction assuming a minimal counterexample (which I will call proof "by minimal counterexample" in the remainder)

$$((\exists m.\ \neg P(m) \wedge (\forall x < m.\ P(x))) \rightarrow \bot) \rightarrow \forall x.\ P(x)$$

where $P$ is the property we want to prove, $>$ is an "appropriate" order, and $m$ denotes a minimal counterexample.

The above formula is (classically) equivalent to

$$(\forall x.\ (\forall y < x.\ P(y)) \rightarrow P(x)) \rightarrow \forall x.\ P(x)$$

which, for well-founded $>$, denotes well-founded induction. This is a desirable alternative, since the outermost proof structure is now constructive, while the inner proof structure stays the same.

So at least superficially it seems that it should be possible to prove whatever we can prove by minimal counterexample, also by well-founded induction. The problem, however, is that the order on infinite sequences that Nash-Williams used is not well-founded. Indeed no suitable well-founded order on infinite sequences immediately suggests itself.

Raoult [17] introduced a viable alternative to well-founded induction in the form of *open induction*, a variation of well-founded induction that exchanges well-foundedness of the order by two other prerequisites. To begin with, the order has to be *downward complete*.

**Definition 3** (*Chains and Downward Completeness*) Let $\sqsubseteq$ be a relation with domain $A$. A $\sqsubseteq$*-chain* $C$ is a totally ordered subset of $A$, that is, for all $c, d \in C$

we have either $c \sqsubseteq d$ or $d \sqsubseteq c$. The relation $\sqsubseteq$ is called *downward complete* if every non-empty $\sqsubseteq$-chain has a greatest lower bound in $A$.

Moreover, open induction is only valid for proving *open* properties.

**Definition 4** (*Open Properties*) A property $P$ is ($\sqsubseteq$-)open if for every non-empty $\sqsubseteq$-chain $C$ it holds that whenever some greatest lower bound $g$ of $C$ satisfies $P(g)$, then $P(x)$ also holds for some $x \in C$.

**Theorem 2** (Open Induction) *Let $\sqsubseteq$ be a downward complete quasi-order on $A$ and $P$ be an $\sqsubseteq$-open property, then the principle of* open induction *reads as follows*

$$(\forall x \in A.\, (\forall y \in A.\, y \sqsubset x \to P(y)) \to P(x)) \to \forall x \in A.\, P(x)$$

*where $x \sqsubset y$ abbreviates $x \sqsubseteq y \wedge y \not\sqsubseteq x$.* □

Here, I state open induction as a theorem, since its correctness has been formalized by Mizuhito Ogawa and myself in Isabelle/HOL (basically by an appeal to Zorn's lemma; the development is available in the *Archive of Formal Proofs* [16]).

At this point, three ingredients are still missing before we can actually apply open induction to prove Higman's lemma. First, we need to fix the property of infinite sequences we want to prove (which must of course be a property which implies that $\sqsubseteq^*$ is almost-full). Second, we need to provide an appropriate (that is, downward complete) order on infinite sequences. And third, we have to make sure that the chosen property is open with respect to the chosen order.

## 4 Setting the Stage: An Open Property and an Appropriate Order

The idea to apply open induction to well-quasi-order theory dates back to Raoult [17]. I am not aware of any actual execution of this idea until the work of Geser [7], who chose a rather complicated order on infinite sequences after arguing that the much simpler lexicographic extension of the proper suffix relation on finite words would make it impossible to use the induction hypothesis (moreover, he tried to prove a variation on Lemma 2(2), namely that every infinite sequence contains an infinite ordered subsequence, instead of Lemma 2(1)).

It turns out, that simply using the lexicographic extension of the proper suffix relation to infinite sequences yields a simpler proof than Geser's initial attempt.

**Definition 5** (*Lexicographic Extension to Infinite Sequences*) Let $\prec$ be a relation with domain $A$. Then the *lexicographic extension* of $\prec$ to infinite $A$-sequences $a$ and $b$ is given by $a \prec_{\mathsf{lex}} b$ iff $a_k \prec b_k$ and $\forall i < k.\, a_i = b_i$ for some $k$.

The following construction will provide a greatest lower bound for each non-empty $\prec_{\mathsf{lex}}$-chain (and is actually the same one I also used to obtain *minimal bad* *sequences* in some Nash-Williams-style proofs I formalized [20] and could therefore be reused).

**Definition 6** (*Minimal Infinite Sequences*) Let $\prec$ be a well-founded partial order with domain $A$, $C$ be a non-empty set of infinite $A$-sequences, and $a$ be an infinite $A$-sequence. Then the set $E_k^a$ of sequences in $C$ that are equal to $a$ up to, but not including, position $k$ is defined by $E_k^a = \{b \in C. \forall i < k. a_i = b_i\}$. Now, a *lexicographically* $\prec$-minimal *sequence* is constructed inductively as follows:

$$\mu_i = \min_{\prec}\{a_i \mid a \in E_i^{\mu}\}$$

That is to say that the $i$th element of $\mu$ is a $\prec$-minimal element of the $i$th "column" of sequences in $E_i^{\mu}$. This construction is well-defined, since obtaining the $i$th element of $m$ only requires access to elements of $m$ whose positions are strictly smaller and $E_i^{\mu}$ is non-empty for all $i$.

**Lemma 3** *Given a well-founded partial order* $\prec$ *and a non-empty* $\prec_{\mathsf{lex}}$*-chain* $C$*, the infinite sequence* $\mu$ *is a greatest lower bound of* $C$*.*

*Proof* Let $C$ be a non-empty $\prec_{\mathsf{lex}}$-chain. Let us first establish that $\mu$ is a lower bound of $C$. To this end, let $a$ be an arbitrary infinite sequence in $C$. If $\mu = a$ we are done. Otherwise, $a \neq \mu$ and thus there is some position $k$ at which $a$ and $\mu$ differ for the first time, that is, $a_i = \mu_i$ for all $i < k$. Then $a \in E_k^{\mu}$ and hence $a_k \in \{b_k \mid b \in E_k^{\mu}\}$. But then $\mu_k \prec a_k$ since we have $a_k \neq \mu_k$, $a_k \prec \mu_k$ is impossible by construction of $\mu$, and $C$ is a $\prec_{\mathsf{lex}}$-chain.

It remains to be shown that $\mu$ is greater than or equal (with respect to $\prec_{\mathsf{lex}}$) to any other lower bound $\ell \neq \mu$ of $C$. Again, take the least $k$ such that $\ell_k \neq \mu_k$ (thus $\ell_i = \mu_i$ for all $i < k$). Now, obtain an infinite sequence $a \in E_{k+1}^{\mu}$, that is, $a_i = \mu_i$ for all $i \leq k$ (which as always possible, since $E_i^{\mu}$ is non-empty for all $i$). But then also $a \in C$. Now remember that $\ell_k \neq \mu_k$. Then, $\ell_k \prec a_k$ (since $a_k = \mu_k$, $a \in C$, and $\ell$ is a lower bound of $C$) and thus $\ell \prec_{\mathsf{lex}} \mu$. □

**Corollary 1** *The lexicographic extension* $\prec_{\mathsf{lex}}$ *is downward complete for every well-founded partial order* $\prec$*.*

Below, I will use the *(proper) suffix relation* $\lhd$ as base order, which is a well-founded partial order given by $xs \lhd ys$ iff $ys$ is obtained by taking some non-empty finite word $zs$ and appending $xs$ (or in words: $xs$ *is a proper suffix of* $ys$).

Now that we have an appropriate order on infinite sequences we still have to fix a property and show that it is open. The property of infinite sequences I will use in my proof of Higman's lemma below, is "being good" (thus, in contrast to Geser, I am using Lemma 2(1) as the defining property of almost-full relations).

That "being good" is an open property is proved in the following way: first I show that every property of infinite sequences that only depends on a finite initial segment (that is, an open property with respect to the product topology of the discrete topology) is $\prec_{\mathsf{lex}}$-open, then I use the fact that "being good" is such a property.

**Lemma 4** *Let $\prec$ be a well-founded partial order. Then every property $P$ that only depends on a finite initial segment is also $\prec_{\mathsf{lex}}$-open.*

*Proof* Assume that $C$ is a non-empty $\prec_{\mathsf{lex}}$-chain with greatest lower bound $g$ such that $P(g)$. Then also $P(\mu)$, since for antisymmetric relations greatest lower bounds are unique and thus $g = \mu$. Since $P$ only depends on some finite initial segment, there is an $n$ such that $P(a)$ for all $a \in E_n^{\mu}$. Moreover, such an $a$ exists, since $C$ is non-empty. □

Obviously, "being good" only depends on a finite initial segment (once we found $i < j$ with $a_i \sqsubseteq a_j$, every extension of the initial segment of $a$ consisting of its first $j$ elements, is also good). Thus we obtain the following result as a corollary.

**Corollary 2** *For any well-founded partial order $\prec$, being good is an $\prec_{\mathsf{lex}}$-open property for arbitrary relations.*

You might wonder whether the reverse of Lemma 4 also holds, that is, whether being $\prec_{\mathsf{lex}}$-open coincides with being open in the product topology. The answer is "no," as shown by the following counterexample.

Consider the domain $A$ consisting of two arbitrary disjoint elements $x$ and $y$. Moreover, take $\prec$ to be the empty relation, which trivially is a well-founded partial order and let $P(a)$ be the property that all elements of $a$ are equal to $y$. Then $P$ is trivially $\prec_{\mathsf{lex}}$-open, but it does not only depend on a finite initial segment (for each $n$, the sequence that differs from $a$ only at position $n$, where it has the value $x$, does not satisfy $P$).

## 5 The Proof via Open Induction

Finally, we are ready for proving Higman's lemma by open induction (the Isabelle/HOL formalization of the proof below is available in the *Archive of Formal Proofs* [20, `Higman_OI.thy`]; it might be interesting to note that the formalization is about the same size as the informal proof below).

*Poof of Theorem* 1. By assumption $\sqsubseteq$ is almost-full on $A$. Since the suffix relation $\lhd$ is a well-founded partial order, its lexicographic extension $\mathsf{lex}$ is downward complete by Corollary 1. Together with the fact that being $\sqsubseteq^*$-good is an open property (Corollary 2), this means—according to Theorem 2—that we can apply open induction in order to prove that every infinite $A$-sequence $a$ is $\sqsubseteq^*$-good (which is to say that $\sqsubseteq^*$ is almost-full on $A^*$).

By induction hypothesis (IH) any infinite $A$-sequence $b \lhd_{\mathsf{lex}} a$ is $\sqsubseteq^*$-good. If $a$ contains the empty word then it is trivially good. Thus we concentrate on the case where for each $i \geq 1$ we have $a_i = h_i \cdot t_i$, that is, $a_i$ consists of a head (letter) $h_i \in A$ and a tail (word) $t_i \in A^*$. Since $\sqsubseteq$ is almost-full on $A$, we obtain an infinite increasing subsequence of $h$ by Lemma 2(2): $h_{\sigma(1)} \sqsubseteq h_{\sigma(2)} \sqsubseteq h_{\sigma(3)} \sqsubseteq \cdots$. We form a new infinite $A$-sequence $a'$ by extending the finite initial segment $a_1, a_2, a_3, \ldots, a_{\sigma(1)-1}$

of $a$ by the infinite $A$-sequence $t_{\sigma(1)}, t_{\sigma(2)}, t_{\sigma(3)}, \ldots$. Then, by construction of $a'$ we have $a' \lhd_{\mathsf{lex}} a$ and thus we obtain an increasing pair $a'_i \sqsubseteq^* a'_j$ for some $i < j$ by IH. We conclude by an exhaustive case analysis on the positioning of $i$ and $j$ within $a'$:

- If $j < \sigma(1)$, then $a_i = a'_i \sqsubseteq a'_j = a_j$ and thus $a$ is good.
- If $i < \sigma(1) \leq j$, then $a_i = a'_i \sqsubseteq^* a'_j = t_{\sigma(j-\sigma(1)+1)} \sqsubseteq^* a_{\sigma(j-\sigma(1)+1)}$. Moreover, $i < \sigma(j - \sigma(1) + 1)$ and thus $a$ is good.
- If $\sigma(1) \leq i$ then $t_{\sigma(i-\sigma(1)+1)} = a'_i \sqsubseteq^* a'_j = t_{\sigma(j-\sigma(1)+1)}$. Which trivially implies $a_{\sigma(i-\sigma(1)+1)} \sqsubseteq^* a_{\sigma(j-\sigma(1)+1)}$. Moreover $\sigma(i - \sigma(1) + 1) < \sigma(j - \sigma(1) + 1)$ and thus $a$ is good. □

## 6 Conclusions and Future Work

I have given a short and mechanically checked proof of Higman's lemma by open induction that highlights the computational content of this well-known lemma through the computational interpretation [1] of open induction.

Here is a short roadmap locating the presented results in my AFP entry (where I indicate lemma names by `typewriter` font):

- `Almost_Full.thy`:
  - `af_trans_extension_imp_wf` (Lemma 1)
  - `almost_full_on_imp_homogeneous_subseq` (Lemma 2: (1) ⟹ (2))
  - `almost_full_on_imp_af` (Lemma 2: (2) ⟹ (3))
  - `af_imp_almost_full_on` (Lemma 2: (3) ⟹ (1))
- `Open_Induction.thy`:
  - `open_induct_on` (Theorem 2)
- `Higman_OI.thy`:
  - `higman` (Theorem 1)
  - `glb_LEX_lexmin` (Lemma 3)
  - `dc_on_LEXEQ` (Corollary 1)
  - `pt_open_on_imp_open_on_LEXEQ` (Lemma 4)
  - `open_on_good` (Corollary 2)

My Isabelle/HOL mechanization of Higman's lemma has already been used by myself and others (although for these applications it is not important *how* the result is obtained):

- Felgenhauer and van Oostrom employ simplification orders to obtain well-foundedness of a complex induction order [5, Theorem 23]. The corresponding proof was formalized in Isabelle/HOL by Felgenhauer [4], building on top of my formalization of Higman's lemma (more precisely `higman` is used to prove `wf_greek_less`).
- Wu et al. [27] use it to obtain a formalization of the fact that: *For every language A, the languages of sub- and superstrings of A are regular.* More precisely, Higman's lemma is used to proof the auxiliary result that, given a language $A$, the set of its minimal elements (with respect to the *subword order*, a special case of $\sqsubseteq^*$, where $\sqsubseteq$ is fixed to equality) is finite. Again, this result is available in the AFP [26] (where the proof of `Closures2.subseq_good` indirectly depends on `higman`).
- My initial goal was to apply mechanized results from well-quasi-order theory to simplify well-foundedness proofs in IsaFoR [23] (an Isabelle Formalization of Rewriting), for example, to obtain well-foundedness of the Knuth-Bendix order (KBO) "for free." However, at the moment it is not clear whether the generalized variant of KBO of IsaFoR [22] which is required to certify generated proofs by several different automated termination provers is a simplification order at all (and frankly, I doubt it).

  Anyway, I reused my formalization of Higman's lemma to also formalize Kruskal's tree theorem [20, 21], which constitutes the theoretical bases for the well-foundedness of *simplification orders* as used in term rewriting.

**Acknowledgements** I am grateful to Mizuhito Ogawa, who introduced me to open induction and the idea of applying it to formalize Higman's lemma in the first place. Thanks goes also to the organizers and participants of the delightful Dagstuhl seminar on *Well Quasi-Orders in Computer Science*, for fruitful discussions and providing a nice atmosphere. I also want to thank Bertram Felgenhauer for giving me a nudge when I was stuck in the proof of Lemma 2(3).
Finally, I am indebted to an anonymous reviewer who pointed out many blemishes in my first draft and provided constructive (pun intended) criticism. Her or his feedback helped to improve this work and lead to a generalization (Lemma 4) of an earlier result (Corollary 2).
This work is supported by FWF (Austrian Science Fund) project P27502.

## References

1. Berger, U. (2004). A computational interpretation of open induction. In *19th IEEE symposium on logic in computer science (LICS 2004)* (pp. 326–334). https://doi.org/10.1109/LICS.2004.1319627.
2. Berghofer, S. (2004). A constructive proof of Higman's lemma in Isabelle. In Berardi, S., Coppo, M., & Damiani, F. (Eds.), *Types for proofs and programs: international workshop, TYPES 2003, Torino, Italy, April 30–May 4, 2003, Revised selected papers* (pp. 66–82). Berlin: Springer. https://doi.org/10.1007/978-3-540-24849-1_5.
3. Coquand, T., & Fridlender, D. (1993). A proof of Higman's lemma by structural induction. Unpublished manuscript.
4. Felgenhauer, B. (2015). Decreasing Diagrams II. *Archive of Formal Proofs*. https://isa-afp.org/entries/Decreasing-Diagrams-II.shtml. Formal proof development.

5. Felgenhauer, B., & van Oostrom, V. (2013) Proof orders for decreasing diagrams. In van Raamsdonk, F. (Ed.), *24th International Conference on Rewriting Techniques and Applications (RTA 2013). Leibniz International Proceedings in Informatics (LIPIcs)* (Vol. 21, pp. 174–189). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. https://doi.org/10.4230/LIPIcs.RTA.2013.174.
6. Fridlender, D. (1998). Higman's lemma in type theory. In Giménez, E., & Paulin-Mohring, C. (Eds.), *Types for proofs and programs: international workshop TYPES'96 Aussois, France, December 15–19, 1996 selected papers* (pp. 112–133). Berlin: Springer. https://doi.org/10.1007/BFb0097789.
7. Geser, A. (1996, April). *A proof of Higman's lemma by open induction*. Technical Report MIP-9606, Universitt Passau.
8. Higman, G. (1952). Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society*, *s3-2*(1), 326–336. https://doi.org/10.1112/plms/s3-2.1.326.
9. Kruskal, J. B. (1960). Well-quasi-ordering, the Tree Theorem, and Vazsonyi's conjecture. *Transactions of the American Mathematical Society*, *95*(2), 210–225. https://doi.org/10.1090/S0002-9947-1960-0111704-1.
10. Kruskal, J. B. (1972). The theory of well-quasi-ordering: a frequently discovered concept. *Journal of Combinatorial Theory, Series A*, *13*(3), 297–305. https://doi.org/10.1016/0097-3165(72)90063-5.
11. Larchey-Wendling, D. (2015). A mechanized inductive proof of Kruskal's tree theorem.
12. Murthy, C., & Russell, J. R. (1989, October) *A constructive proof of Higman's lemma*. Technical Report TR 89-1049, Cornell University, Ithaca, NY 14853-7501. http://hdl.handle.net/1813/6849.
13. Murthy, C. R., & Russell, J. R. (1990). A constructive proof of Higman's lemma. In *Proceedings of the Fifth Annual Symposium on Logic in Computer Science (LICS '90), Philadelphia, Pennsylvania, USA, June 4–7, 1990* (pp. 257–267). https://doi.org/10.1109/LICS.1990.113752.
14. Nash-Williams, C. S. J. A. (1963). On well-quasi-ordering finite trees. *Mathematical Proceedings of the Cambridge Philosophical Society*, *59*(4), 833–835. https://doi.org/10.1017/S0305004100003844.
15. Nipkow, T., Paulson, L. C., Wenzel, M. (2002). *Isabelle/HOL-a proof assistant for higher-order logic*. Lecture Notes in Computer Science (Vol. 2283). Springer. https://doi.org/10.1007/3-540-45949-9.
16. Ogawa, M., & Sternagel, C. (2012, November). Open induction. *Archive of Formal Proofs*. https://isa-afp.org/entries/Open_induction.shtml. Formal proof development.
17. Raoult, J.-C. (1988). Proving open properties by induction. *Information Processing Letters*, *29*(1), 19–23. https://doi.org/10.1016/0020-0190(88)90126-3.
18. Richman, F., & Stolzenberg, G. (1993). Well quasi-ordered sets. *Advances in Mathematics*, *97*(2), 145–153. https://doi.org/10.1006/aima.1993.1004.
19. Schwichtenberg, H., Seisenberger, M., & Wiesnet, F. (2016). Higman's lemma and its computational content. In Kahle, R., Strahm, T., & Studer, T. (Eds.), *Advances in proof theory* (pp. 353–375). Cham: Springer. https://doi.org/10.1007/978-3-319-29198-7_11.
20. Sternagel, C. (2012, April). Well-quasi-orders. *Archive of Formal Proofs*. https://isa-afp.org/entries/Well_Quasi_Orders.shtml. Formal proof development.
21. Sternagel, C. (2014). Certified Kruskal's tree theorem. *Journal of Formalized Reasoning*, *7*(1), 45–62. https://doi.org/10.6092/issn.1972-5787/4213.
22. Sternagel, C., & Thiemann, R. (2013). Formalizing Knuth-Bendix orders and Knuth-Bendix completion. In *24th international conference on rewriting techniques and applications*, Dagstuhl. LIPIcs (Vol. 21, pp. 287–302). https://doi.org/10.4230/LIPIcs.RTA.2013.287.
23. Thiemann, R., & Sternagel, C. (2009). Certification of termination proofs using CeTA. In *22nd international conference on theorem proving in higher-order logics (TPHOLs 2009)*. Lecture Notes in Computer Science (Vol. 5674, pp. 452–468). Springer. https://doi.org/10.1007/978-3-642-03359-9_31.
24. Veldman, W., & Bezem, M. (1993). Ramsey's theorem and the pigeonhole principle in intuitionistic mathematics. *Journal of the London Mathematical Society*, *s2-47*(2), 193–211. https://doi.org/10.1112/jlms/s2-47.2.193.

25. Vytiniotis, D., Coquand, T., Wahlstedt, D. (2012). Stop when you are almost-full. In Beringer, L. & Felty, A. (Eds.), *3rd international conference on interactive theorem proving (ITP 2012)* (pp. 250–265). Berlin: Springer. https://doi.org/10.1007/978-3-642-32347-8_17.
26. Wu, C., Zhang, X., Urban, C. (2011, August). The Myhill-Nerode theorem based on regular expressions. *Archive of Formal Proofs*. https://isa-afp.org/entries/Myhill-Nerode.shtml. Formal proof development.
27. Wu, C., Zhang, X., & Urban, C. (2014). A formalisation of the Myhill-Nerode theorem based on regular expressions. *Journal of Automated Reasoning*, *52*(4), 451–480. https://doi.org/10.1007/s10817-013-9297-2.

# Well-Partial Orderings and their Maximal Order Types

**Diana Schmidt**

**Abstract** Combinatorial theorists have for some time been showing that certain partial orderings are well-partial-orderings (*w.p.o.*'s). De Jongh and Parikh showed that w.p.o.'s are just those well-founded partial orderings which can be extended to a well-ordering of maximal order type; we call the ordinal thus obtained the *maximal order type* of the w.p.o. In this paper we calculate, in terms of a system of notations due to Schütte [24], the maximal order types of the w.p.o.'s investigated in Higman [11], and give upper bounds for the maximal order types of the w.p.o.'s investigated in Kruskal [13] and Nash-Williams [16]. As a by-product and an application of de Jongh and Parikh's work, we give new and easier proofs of Higman's, Kruskal's and Nash–Williams' theorems that the partial orderings considered are indeed w.p.o.'s. We also apply our results to the theory of ordinal notations.

## 1 Introduction

A natural question connected with well-orderings and their order types is: Given a well-ordering $\leqslant^+$ of a set S about which only a limited amount of information is available, under what conditions will this information yield a non-trivial upper bound for the order type of (S, $\leqslant^+$) ("non-trivial" means "lower than the obvious upper bound obtained by considering the cardinality of S")? A special case of this question is: Given a well-ordering $\leqslant^+$ of S and a partial ordering $\leqslant$ of S such that $\leqslant^+$ is an extension of $\leqslant$, under what conditions is there a non-trivial upper bound for the order type of (S, $\leqslant^+$) which depends only on (S, $\leqslant$)?

Diana Schmidt (✉)
Hochschule Heilbronn, Max-Planck-Str. 39, D-74081 Heilbronn, Germany
e-mail: diana.schmidt@hs-heilbronn.de

 
P. M. Schuster et al. (eds.), *Well-Quasi Orders in Computation, Logic, Language and Reasoning*, Trends in Logic 53,
https://doi.org/10.1007/978-3-030-30229-0_13

The answer to the second question turns out [12] to lie in a notion which has. interested combinatorial theorists for a quarter of a century: A sufficient (and, for countable S, necessary) condition is that (S, $\leqslant$) be a *well-partial-ordering* (w.p.o.). The most suggestive definition of this notion is that (S, $\leqslant$) is a w.p.o. if and only if it is a partial ordering such that every extension of $\leqslant$ to a linear ordering of S is in fact a well-ordering of S. De Jongh and Parikh's work shows that for each w.p.o. (S, $\leqslant$) the least upper bound of the order types of those (S, $\leqslant^+$) such that $\leqslant^+$ extends $\leqslant$ to a well-ordering on S is actually attained by some $\leqslant^+$, and is therefore non-trivial. For this reason we shall call this least upper bound the *maximal order type* of (S, $\leqslant$). De Jongh and Parikh's work moreover provides tools for estimating the maximal order types of particular w.p.o.'s. These, in turn, can yield upper bounds for the order types of complicated well-orderings, namely any which are extensions of the given w.p.o.'s. See Schmidt [23] for an example of such an application.

In this paper we shall use de Jongh and Parikh's methods to obtain upper bounds for the maximal order types of some well-partial orderings. (We shall also, as a by-product, give new proofs of the fact that these are indeed w.p.o.'s – proofs which, we think, are easier than most of those available so far.) The well-partial orderings whose maximal order types we shall estimate are:

(i) The set of all $\leqslant$ n-branching finite structured trees with labels in an arbitrary w.p.o. set, ordered by the relation "is homeomorphically embeddable into" – this is equivalent to the w.p.o. considered in Higman [11];
(ii) the set of *all* finite structured trees with labels in an arbitrary w.p.o. set, ordered by the relation "is homeomorphically embeddable into" – this was shown to be a w.p.o. in Kruskal [13];
(iii) the set of all those (finite or transfinite) sequences of elements of an arbitrary w.p.o. set which contain only finitely many different members; ordered by the relation "is majorised (in the sense of the w.p.o.) term-by-term by a subsequence of" – this was shown to be a w.q.o. (which is more-or-less the same as a w.p.o. - see p. 1) in Nash–Williams [16].

Section 2 contains the necessary preliminaries; Sect. 3 the statements and proofs of the results described above; and Sect. 4 applications of these results to the theory of ordinal notations. The results of Sects. 3 and 4 yield characterisations and/or closure properties of some ordinals which also turn up naturally in proof theory. It is therefore quite possible that these results could give rise to new insights into the proof theory of the relevant deductive systems. However, gleaning such insights would definitely be a non-trivial task; for the moment we have contented ourselves with simply pointing out (Sect. 4.3) which of the ordinals obtained here as maximal order types are significant in proof theory. We have attached an Appendix, consisting of part of [5], giving a summary of those countable ordinals which have so far been shown to be relevant to proof theory.

# 2 Preliminaries

## 2.1 Well-Quasi-Orderings – Definitions and Results

### 2.1.1 Definitions

A *quasi-ordering* (*q.o.*) is a pair $(X, \leqslant)$, where $X$ is a set and $\leqslant$ is a transitive and reflexive binary relation on $X$. A *partial ordering* (*p.o.*) is a quasi-ordering in which $\leqslant$ is also anti-symmetric. Any quasi-ordering $(X, \leqslant)$ may be regarded as a partial ordering of the set $X/\cong$, where $\cong$ is the equivalence relation on $X$ defined by: $x \cong y \Leftrightarrow x \leqslant y$ and $y \leqslant x$.

For any q.o. $(X, \leqslant)$ and any $x, y \in X$ we write $x < y$ for ($x \leqslant y$ and $y \nleqslant x$). A *linear ordering* is a partial ordering $(X, \leqslant)$ in which any two elements of $X$ are $\leqslant$-comparable (i.e. for any $x, y \in X$ at least one of $x \leqslant y$, $y \leqslant x$ holds). A p.o. $(X, \leqslant)$ is *well-founded* if and only if every nonempty subset $Y$ of $X$ contains at least one minimal element (w.r.t. $\leqslant$), and *well-ordered* if and only if every nonempty subset $Y$ of $X$ contains precisely one minimal element. Every well-ordering is a linear ordering.

A *well-quasi-ordering* (*w.q.o.*) is a quasi-ordering $(X, \leqslant)$ such that there is no sequence $\langle x_i | i \in \omega\rangle$ of elements of $X$ satisfying: $x_i \nleqslant x_j$. for all $i < j$.

A *well-partial-ordering* (*w.p.o.*) is a well-quasi-ordering which is also a partial ordering. It is not very hard to show (see Higman [11], Wolk [27]) that the following conditions are all necessary and sufficient for a partial ordering $(X, \leqslant)$ to be a w.p.o.:

(i) Every extension of $\leqslant$ to a linear ordering on $X$ is a well-ordering.
(ii) $(X, \leqslant)$ is well-founded and $X$ contains no infinite subset whose elements are pairwise $\leqslant$-incomparable.
(iii) Every nonempty subset of $X$ contains at least one, but not infinitely many minimal elements (w.r.t. $\leqslant$).
(iv) Any sequence $\langle x_i | i \in \omega\rangle$ of elements of $X$ contains an infinite subsequence $\langle x_{i_j} | j \in \omega\rangle$ such that $x_{i_j} \leqslant x_{i_{j+1}}$ for each $j \in \omega$.

Note that every w.q.o. is a w.p.o. modulo the equivalence relation given above; thus conditions (i)–(iv), and any other theorems about w.p.o.'s, can be translated into statements about w.q.o.'s, and vice versa. In the following we shall sometimes talk in terms of w.p.o.'s and sometimes in terms of w.q.o.'s, depending on which happens to be more convenient in the context.

Well-quasi-orderings and well-partial-orderings have been studied for quite a while, mainly by (infinitary) combinatorial theorists. We shall not go into the history and development of the concept here, as this is well covered in Kruskal's [14] survey.[1] Suffice it to say that many, often difficult theorems have been proved which show that certain quasi-orderings are well-quasi-orderings; for example, Kruskal's theorem that the set of all finite trees with labels in a w.q.o. set is well-quasi-ordered by the

[1] But see also the note on p. 8.

relation of homeomorphic embedding, and Nash–Williams' theorem that the class of all well-ordered sequences of elements of a well-ordered set is well-quasi- ordered (in fact, better-quasi-ordered—cf. Kruskal [14]) by the relation defined in Sect. 3.3.

Recently, de Jongh and Parikh [12] struck out in a new direction by showing that, if $(X, \leqslant)$ is a w.p.o. and $\alpha$ the supremum of all ordinals which are order types of $X$ with respect to some extension of $\leqslant$ to a linear ordering (see property (i) above of w.p.o.'s), then there is an extension of $\leqslant$ to a linear ordering on $X$ which actually has order type $\alpha$. This means that each w.p.o. is associated with an ordinal – its maximal order type – in a meaningful way, and this ordinal can be used to prove assertions about w.p.o.'s by transfinite induction. In fact, the theorem of de Jongh and Parikh also yields a tool for calculating the maximal order types of given w.p.o.'s.

### 2.1.2 Results on w.p.o.'s (de Jongh and Parikh)

We summarize de Jongh and Parikh's results below, referring the reader to [12] for proofs.

**Definition 2.1** If $(X, \leqslant)$ is a w.p.o., we define its *maximal order type* by

$$o(X, \leqslant) = \sup \{\alpha | \alpha \text{ is the order type of } (X, \leqslant^+) \text{ for some extension } \leqslant^+ \text{ of } \leqslant \text{ to a linear ordering}\};$$

for any $x \in X$,

$$L_{(X, \leqslant)}(x) = \{y | y \in X \;\&\; x \nleqslant y\}.$$

$L_{(X, \leqslant)}(x)$ contains just those elements of $X$ which could be below $x$ in any linear ordering extending $\leqslant$.

$$l_{(X, \leqslant)}(x) = o(L_{(X, \leqslant)}(x), \leqslant \restriction L_{(X, \leqslant)}(x)).$$

$l_{(X, \leqslant)}(x)$ is the highest possible position of $X$ in any linear ordering extending $\leqslant$.

If $(X, \leqslant)$ is a w.q.o., we shall write $o(X, \leqslant)$ for $o(X', \leqslant')$, where $(X', \leqslant')$ is the associated w.p.o., and $l_{(X, \leqslant)}(x)$ for $l_{(X', \leqslant')}(x)$.

We shall write $o(X)$ for $o(X, \leqslant)$ and omit the subscript $(X, \leqslant)$ whenever this causes no confusion.

**Theorem 2.2** *If $(X, \leqslant)$ is a w.p.o., then there is an extension $\leqslant^+$ of $\leqslant$ to a linear ordering such that $(X, \leqslant^+)$ has order type $o(X, \leqslant)$.*

**Corollary 2.3** *If $(X, \leqslant)$ is a w.q.o. and $x \in X$, then $l_{(X, \leqslant)}(x) < o(X, \leqslant)$.*

**Corollary 2.4** *If $(X, \leqslant)$ is a w.q.o. and $o(X, \leqslant)$ is a limit ordinal, then $o(X, \leqslant) = \sup_{x \in X} l_{(X, \leqslant)}(x)$.*

*Definition.* If $(X, \leqslant)$ is a w.q.o., $M \subseteq X$ is said to he a *majorising subset* of $X$ if for every $x \in X$ there is a $y \in M$ such that $x \leqslant y$.

(This notion is slightly different from de Jongh and Parikh's, but – we think – often easier to work with.)

**Corollary 2.5** *If $(X, \leqslant)$ is a w.q.o., $o(X, \leqslant)$ a limit ordinal and $M$ a majorising subset of X, then* $o(X, \leqslant) = \sup_{x \in M} l_{(X, \leqslant)}(x)$.

**Theorem 2.6** *If $(X, \leqslant_X)$ and $(Y, \leqslant_Y)$ are w.p.o.'s (resp. w.q.o.'s), then so is $(X \dot\cup Y, \leqslant_X \dot\cup \leqslant_Y)$ (this was known before) and $o(X \dot\cup Y) = o(X) \# o(Y)$, where $\#$ denotes natural sum (see [2] for a definition of $\#$) and $\dot\cup$ disjoint union.*

**Theorem 2.7** *If $(X, \leqslant_X)$ and $(Y, \leqslant_Y)$ are w.p.o.'s (resp. w.q.o.'s), then so is $(X \times Y, \leqslant_X \times \leqslant_Y)$ (this was known before) and $o(X \times Y) = o(X) \otimes o(Y)$, where $\otimes$ denotes natural product (see [2] for a definition of $\otimes$), and $\times$ cartesian product.*

### 2.1.3 The Maximal Order Type of $(X^*, \leqslant^*)$

The following theorem is proved for finite $o(X)$ in [12] and will be proved for all $o(X)$ in a forthcoming paper of de Jongh. We give a somewhat simpler proof here as an easy illustration of the methods which are to be used in Sect. 3.

**Definition 2.8** Let $(X, \leqslant)$ be a w.p.o. (resp. w.q.o.) Define $(X^*, \leqslant^*)$ as follows:
$X^*$ is the set of all finite sequences of elements of $X$ (including the empty sequence $<>$); $x_1 \cdots x_m \leqslant^* y_1 \cdots y_n$ iff there is a strictly monotonic function $f : \{1, \ldots, m\} \to \{1, \ldots, n\}$ such that $x_i \leqslant y_{f(i)}$ for all $i \leqslant m$.
It is known that $(X^*, \leqslant^*)$ is also a w.p.o. (resp. w.q.o.).

**Theorem 2.9** (de Jongh and Parikh) *If $(X, \leqslant)$ is a w.q.o. and $X$ is nonempty, then $o(X^*, \leqslant^*) \leqslant \omega^{\omega^{(o(X, \leqslant))'}}$, (see also note on p. 358) where for any ordinal $\alpha$*

$$\alpha' = \begin{cases} \alpha - 1 & \text{if } \alpha \text{ is finite and nonzero,} \\ \alpha + 1 & \text{if there are } n < \omega \text{ and } \delta \text{ such that } \omega^\delta = \delta \text{ and } \alpha = \delta + n, \\ \alpha & \text{otherwise.} \end{cases}$$

*Proof* By transfinite induction on $o(X, \leqslant)$.
$o(X, \leqslant) = 1$: Then clearly $o(X^*) = \omega = \omega^{\omega^0}$.
$o(X, \leqslant) > 1$: In the following we shall denote $\omega^{\omega^{(o(X, \leqslant))'}}$ by $\beta$.
It is easy to see that $o(X^*, \leqslant^*)$ is a limit ordinal.
Now note that $X^*$ can be generated by the following inductive definition:

(1) $<> \in X^*$.
(2) $w \in X^*$ and $x \in X \Rightarrow xw \in X^*$

(where $xw$ denotes the sequence obtained by attaching $x$ to the beginning of the sequence $w$).

Accordingly, by Corollary 2.4, it is sufficient to prove that:

**Proposition 1** $l(<>) < \beta$;

**Proposition 2** $x \in X \;\&\; w \in X^* \;\&\; l(w) < \beta \Rightarrow l(xw) < \beta$.

*Proof of Proposition 1*: $<> \leqslant^* v$ for all $v \in X^*$; hence $L(<>)$ is empty, and $l(<>) = 0 < \beta$.

*Proof of Proposition 2*: Suppose that $w \in X^*$, $x \in X$ and $l(w) < \beta$.

Now let $v \in L(xw)$. Then either $v$ is a string of elements of $L_{(X,\leqslant)}(x)$; or $v$ contains a leftmost $y$ such that $x \leqslant y$, and then that part of $v$ which lies to the right of $y$ must be an element of $L_{(X^*,\leqslant^*)}(w)$ (for if this were not the case we should have $xw \leqslant^* v$ after all). In other words, $L_{(X^*,\leqslant^*)}(xw)$ may be regarded as a subset of

$$L = (L_{(X,\leqslant)}(x))^* \dot{\cup} \left[ (L_{(X,\leqslant)}(x))^* \times X \times L_{(X^*,\leqslant^*)}(w) \right].$$

And, what is more important, the order relation $(\leqslant^* \dot{\cup} (\leqslant^* \times \leqslant \times \leqslant^*))$ on $L_{(X^*,\leqslant^*)}$ $(xw)$ induced by this way of looking at $L_{(X^*,\leqslant^*)}(xw)$ is a subrelation of $\leqslant^*$. Hence any extension of $\leqslant^*$ on $L(xw)$ is also an extension of the order relation on the subset of $L$ corresponding to $L(xw)$. Hence

$$\begin{aligned} l(xw) &= o(L_{(X^*,\leqslant^*)}(xw)) \leqslant o(L) \\ &= o(L_{(X,\leqslant)}(x))^* \# \left[ o(L_{(X,\leqslant)}(x))^* \divideontimes o(X) \divideontimes l_{(X^*,\leqslant^*)}(w) \right] \\ &\quad \text{by Theorems 6 and 7.} \end{aligned}$$

But $l_{(X,\leqslant)}(x) < o(X)$ by Corollary 2.3; hence $o(L_{(X,\leqslant)}(x))^* \leqslant \omega^{\omega^{(l(x))'}} < \omega^{\omega^{(o(X))'}} = \beta$ by induction hypothesis;

$o(X) < \omega^{\omega^{(o(X))'}} = \beta$ by the definition of $(o(X))'$; and $l_{(X^*,\leqslant^*)^{(w)}} < \beta$ by hypothesis.

Hence, since all ordinals of the form $\omega^{\omega^r}$ are closed under both $\#$ and $\divideontimes$ (see [2]), $l(xw) < \beta$, q.e.d.

We have spelt the proof of Theorem 2.9 out in detail to make the essentials of the method as clear as possible; when we use similar arguments in Sect. 3 we shall express them more briefly.

**Lemma 2.10** *Let $(X, \leqslant)$ be a q.o., $M$ a majorising subset of $X$. If $(L(m), \leqslant \restriction L(m))$ is a w.q.o. for each $m \in M$, then $(X, \leqslant)$ is also a w.q.o.*

*Proof* Suppose $(X, \leqslant)$ is not a w.q.o. Then, by characterisation (ii) of w.p.o.'s, there is a set $\{x_i | i \in \omega\} \subseteq X$ such that either $x_j < x_i$. for all $i < j \in \omega$ or $(x_i \nleq x_j \;\&\; x_j \nleq x_i)$ for all i $< j \in \omega$.

But in either case, $x_0 \nleqslant x_i$ for all $i > 0$; hence, if $m \in M$ satisfies $x_0 \leqslant m$, then $m \nleqslant x_i$ for all $i > 0$; i.e. $\{x_i | 0 < i \in \omega\} \subseteq L(m)$ and $\{x_i | 0 < i \in \omega\}$ satisfies the above condition; hence $(L(m), \leqslant\restriction L(m))$ is not a w.q.o., contradiction.

**Corollary 2.11** *If* $(X, \leqslant)$ *is a w.q.o., then so is* $(X^*, \leqslant^*)$.

*Proof* This fact is not new; it clearly also follows from Theorem 2.9 for countable $X$, since it is easy to show that $(X^*, \leqslant^*)$ is well-founded and by Theorem 2.9 $X^*$ contains no infinite subset whose elements are pairwise $\leqslant^*$-incomparable (for if it did then every countable ordinal would be the order type of some well-ordering on this infinite subset; hence every countable ordinal would be equalled or exceeded by the order type of $X^*$ under some extension of $\leqslant^*$ to a linear ordering; hence $o(X^*, \leqslant^*)$ would be the first uncountable ordinal, which would contradict Theorem 2.9). However, Corollary 2.11 follows from the *proof* of Theorem 2.9 in the following way:

In the proof of Theorem 2.9 we show by induction w.r.t. $o(X)$ that $o(L(w)) < \beta$ for each $w \in X^*$. But we could equally well forget the ordinal bound $\beta$ and show by induction w.r.t. $o(X)$ that $(L(w), \leqslant\restriction L(w))$ is a w.q.o. for each $w \in X^*$, by simply replacing each assertion of the form

$$o(- - -) < \gamma$$

in the proof by the assertion

$$- - - \text{ is a w.q.o.}$$

It would then follow by Lemma 2.10 that $(X^*, \leqslant^*)$ is also a w.q.o.

All the proofs in Sect. 3 which give bounds for the maximal order types of certain w.q.o.'s can be adapted in the same way to become proofs that the given q.o. is indeed a w.q.o. The proofs of well-quasi-orderedness obtained in this way are, we think, simpler than those in the literature. Note that they all depend essentially on Corollary 2.3 (in order to make the induction hypothesis applicable), and hence on de Jongh and Parikh's Theorem 2.2.

### 2.1.4 A Note on the Minimal Order Type of a w.p.o.

It may occur to the reader to ask, for a given w.p.o. $(X, \leqslant)$, what is the *least* ordinal which is the order type of $(X, \leqslant^+)$ for some extension $\leqslant^+$ of $\leqslant$ to a well-ordering. The remarks below show that this ordinal is identical with the height of $(X, \leqslant)$ as a partial ordering if this is a limit ordinal, and differs from the height of $(X, \leqslant)$ at most by a finite ordinal otherwise. By contrast, note that the w.p.o.'s in Sects. 3.1 and 3.2 have height $\omega$ and maximal order types which are much larger than – if, for example, the $X_i$ are all taken to be singletons.

*Question* Is there any (non-trivial) relationship between the height of a w.p.o. and its maximal order type (note that, by [27], the height of a w.p.o. is equal to the length of its longest chain)? A related question is: Is it true that the maximal order type of a recursive w.p.o. is always a recursive ordinal?

We now prove the assertions above about the minimal order type of a w.p.o.

*Definitions* Let $(X, \leqslant)$ be a well-founded partial ordering. For $x \in X$, we define the *height* of $x$ w.r.t. $\leqslant$ by transfinite recursion w.r.t. $\leqslant$ as follows:

$$ht(x) = \text{strict sup}\{ht(y) | y \in X \text{ and } y < x\}.$$

We then define

$$\text{height of } (X, \leqslant) = \sup\{ht(x) |\ x \in X\}.$$

Thus the height of $(X, \leqslant)$ is the least ordinal into which $X$ can be embedded in a $\leqslant$-preserving manner.
Now let $f : X \to \alpha$ be any 1–1 mapping of $X$ into some ordinal $\alpha$ (assuming the Axiom of Choice). We define $\leqslant^+$ by:

$$x \leqslant^+ y \Leftrightarrow x \in X \ \& \ y \in X \ \& \ (ht(x) < ht(y) \text{ or } [ht(x) = ht(y) \ \& \\ \& \ f(x) < f(y)]).$$

Then $\leqslant^+$ is an extension of $\leqslant$ to a well-ordering of $X$.
Moreover, note that if $ht(x) = ht(y)$ then either $x = y$ or $x$ and $y$ are $\leqslant$-incomparable. Thus, since $(X, \leqslant)$ is a w.p.o., there are only finitely many elements of $X$ of any given height. Thus it is easy to show that, for any $x \in X$,
$ht(x)$ w.r.t. $\leqslant^+ \ < \ (ht(x)$ w.r.t. $\leqslant) + \omega$.
Hence if the height of $(X, \leqslant)$ is a limit ordinal then
order type of $(X, \leqslant^+) =$ height of $(X, \leqslant)$;
and if not then
height of $(X, \leqslant) \leqslant$ order type of $(X, \leqslant^+) <$ (height of $(X, \leqslant)) + \omega$.

*Note* to Kruskal's [14] survey: A further important early paper on well-partial orderings is that of [7]. Although Carruth does not use the notion of a w.p.o., he in effect proves that the disjoint union and the cartesian product of two well-orderings are w.p.o.'s, and calculates the maximal order types of these w.p.o.'s as given in Theorems 2.6 and 2.7 below. We are grateful to Hilbert Levitz for drawing our attention to this paper.

*Note* to Theorem 2.9: It is easy to prove by transfinite induction on $o(X, \leqslant)$ that the inequality can also be reversed; the proof of the successor case can be taken from the proof of Theorem 3.11 in [12], and we leave the proof of the limit case to the reader.

## 2.2 Sequences and Trees Over a q.o. Set

This section contains the definitions of the sets and relations (which are w.q.o.'s and) whose maximal order types will be studied in Sect. 3.

**Definition 2.12** (*Sequences*) Let $(X, \leqslant)$ be a w.q.o. For any ordinal $\alpha$,

$$S_\alpha(X) = \{f \mid f : \beta \to X \text{ for some } \beta < \alpha\};$$

i.e. $S_\alpha(X)$ is the set of all sequences of elements of $X$ of length less than $\alpha$.

$S^F_\alpha(X) = \{f \mid f : \beta \to X$ for some $\beta < \alpha$ and range $(f)$ is finite$\}$; i.e. $S^F_\alpha(X)$ contains just those elements of $S_\alpha(X)$ in which only finitely many different elements of $X$ occur. Of course, $S^F_\alpha(X) = S_\alpha(X)$ for all $\alpha \leqslant \omega$.

*Examples* $S_0(X)$ is empty, $S_1(X)$ contains just the empty sequence, $S_2(X)$ is isomorphic to $(X \dot\cup \{0\})$ and $S_\omega(X)$ to $X^*$.

Now if $(X, \leqslant)$ is a q.o., then $\leqslant$ induces a q.o. $\leqslant_S$ on $S_\alpha(X)$ in a natural way:

$$\begin{aligned} f \leqslant_S g \Leftrightarrow\, & f : \beta \to X,\ g : \gamma \to X \text{ and there is a strictly monotonic} \\ & \text{function } h : \beta \to \gamma \text{ such that} \\ & f(\delta) \leqslant g(h(\delta)) \text{ for all } \delta < \beta. \end{aligned}$$

In other words, $f \leqslant_S g$ if and only if there is a subsequence of $g$ which $\leqslant$-majorises $f$ term by term.

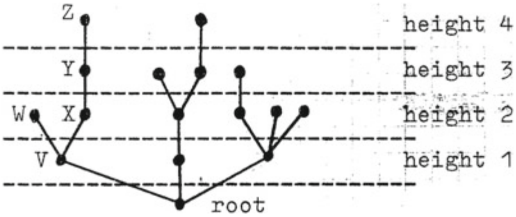
Note that, even if $(X, \leqslant)$ is a p.o., $(S_\alpha(X), \leqslant_S)$ is in general only a q.o. – for example, the following two sequences of $S_{\omega+1}(\{0, 1\})$ are $\leqslant_S$-equivalent but not equal:

01010101 . . .    10101010 . . .

Thus, whereas in almost all theorems in this paper 'w.q.o,' can be replaced throughout by 'w.p.o.', this is not true of the theorem below.

Nash–Williams [16] has proved that, if $(X, \leqslant)$ is a w.q.o., then $(S^F_\alpha(X), \leqslant_S)$ is a w.q.o. for each $\alpha$; we shall give another proof of this result in Sect. 3, 3.3. Nash–Williams [17] showed that, if $(X, \leqslant)$ is a well-ordering, then $(S_\alpha(X), \leqslant_S)$ is a w.q.o. (in fact, a b.q.o.—see [18]).

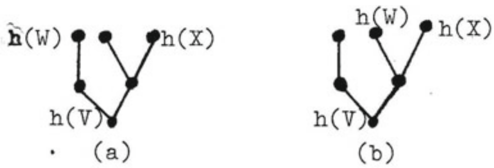
**Definition 2.13** (*Trees*) A *tree* is a graph in which any two distinct vertices are connected by precisely one path. A *rooted tree* is a tree with one distinguished vertex, called its *root*. Since the root is connected to every vertex by precisely one path, we may define the height of any vertex as the length of this path. Then any rooted tree may be represented in layers by putting all vertices with the same height in the same layer, to look roughly like a common or garden tree:




If $\tau$ is a rooted tree and $V$ one of its vertices, the *immediate successors of* $V$ are those vertices which are connected to $V$ by an edge of $\tau$ and have height greater than the height of $V$; for example, the immediate successors of $V$ in the picture are $W$ and $X$. The *successors* of $V$ are defined as follows: The relation 'is a successor of' is the transitive closure of the relation 'is an immediate successor of'. Thus the successors of $V$ in the picture are $W$, $X$, $Y$, $Z$. Note that the relation 'is a successor of' is a well-founded partial ordering on the vertices of $\tau$. A *structured tree* is a rooted tree $\tau$ together with a relation which, for each vertex $V$ of $\tau$, well- orders the set of $V$'s immediate successors. We shall call this relation the *horizontal relation* of the structured tree. Let $X$ be any set. A *structured tree with labels in* $X$ (*structured tree over* $X$) is a structured tree $\tau$ together with a mapping $f$ from the set of vertices of $\tau$ into $X$; if $V$ is a vertex of $\tau$, $f(V)$ is called the *label* of $V$. In diagrams, we shall indicate labels by writing their names beside the appropriate vertex.

**Definition 2.14** (*The relation* $\leqslant_T$) Let $\tau_1$, $\tau_2$ be two rooted trees. $\tau_1$ is *homeomorphically embeddable* into $\tau_2$ iff there is a function $h$ (a *homeomorphic embedding* of $\tau_1$ into $\tau_2$) from the set of vertices of $\tau_1$ into the set of vertices of $\tau_2$ such that

(i) $h$ preserves the 'successor' ordering; i.e. if $W$ is a successor of $V$ in $\tau_1$ then $h(W)$ is a successor of $h(V)$ in $\tau_2$;
(ii) for any vertex $V$ of $\tau_1$ and any distinct immediate successors $W$ and $X$ of $V$, the paths from $h(V)$ to $h(W)$ and from $h(V)$ to $h(X)$ have no vertices except $h(V)$ in common; i.e. they pass through distinct immediate successors of $h(V)$. In diagram (a) this condition is satisfied and in diagram (b) it is not:




If $\tau_1$, $\tau_2$ are two structured trees, then $\tau_1$ is homeomorphically embeddable into $\tau_2$ by $h$ iff $h$ satisfies (i) and (ii) and also preserves the horizontal relation in the following sense:

If $V$ is a vertex of $\tau_1$ and $W$ and $X$ are distinct immediate successors of $V$ such that $(W, X)$ is in the horizontal relation of $\tau_1$, then $(P, Q)$ is in the horizontal relation of $\tau_2$, where $P$ is the immediate successor of $h(V)$ on the path to $h(W)$ and $Q$ the immediate successor of $h(V)$ on the path to $h(X)$.

If $\tau_1$, $\tau_2$ are two structured trees with labels in $X$, where $(X, \leqslant)$ is a q.o., then

$\tau_1 \leqslant_T \tau_2$ iff there is a function $h$ which homeomorphically embeds $\tau_1$ into $\tau_2$ (disregarding the labels) and which also satisfies label of $V \leqslant$ label of $h(V)$ for each vertex $V$ of $\tau_1$.

Kruskal [13] has proved that, if $(X, \leqslant)$ is a w.p.o. (w.q.o.), then $\leqslant_T$ is a w.p.o. (w.q.o.) on the set of all *finite* structured trees with labels in $X$. From this it follows that the homeomorphic embedding relation well-partially orders the set of all finite trees. In fact, Nash–Williams [17] has shown that the homeomorphic embedding relation well-quasi-orders the class of all trees. Section 3.2 will yield an alternative proof of Kruskal's result together with the relevant maximal order types. Our proof of Kruskal's result is, we think, simpler than Kruskal's, but not as easy as the very elegant proof in Nash–Williams [15].

## *2.3 Systems of Notations for Ordinals*

First, a few preliminary remarks: We shall work in Zermelo- Fraenkel set theory with the Axiom of Choice in an informal way, as expounded, for example, in Halmos [9]. Additional information about ordinals may be gleaned from Bachmann [2]. We shall assume knowledge only of the ordinal functions +, $\cdot$ and exponentiation, and of the natural sum and product $\#$ and $\otimes$. All the reader needs to know about $\#$ and $\otimes$ is that they are binary functions on the ordinals such that, for any ordinals $\alpha, \beta, \gamma$

$$\alpha, \beta < \omega^{\omega^{\gamma}} \Rightarrow \alpha \# \beta, \alpha \otimes \beta < \omega^{\omega^{\gamma}}.$$

We shall always identify each ordinal with the set of all its predecessors; thus $\alpha < \beta$ is synonymous with $\alpha \in \beta$ if $\alpha, \beta$ are ordinals; and the remark above could be rephrased as "$\omega^{\omega^{\gamma}}$ is closed under $\#$ and $\otimes$". $\omega$ will denote the smallest infinite ordinal and $\Omega$ the smallest uncountable one.

As the ordinals which are going to emerge as maximal order types in Sect. 3 are much bigger than those for which generally accepted notation is available (e.g. ($\omega + 2, \omega^{\omega 3}, \varepsilon_0$ etc.), we shall introduce here a system of notations for ordinals in terms of various ordinal functions, due to Schütte [24]. The Bachmann-Feferman-Aczel-Bridge systems are stronger and much better known these days, but we shall use Schütte's system because the ordinals which will arise in this paper can be expressed more neatly in it. We shall provide a translation into the Bachmann notation below.

**Definition 2.15** (Schütte) A *Klammersymbol* (*KS*) is a configuration $\begin{pmatrix} \alpha_0 \dots \alpha_n \\ \beta_0 \dots \beta_n \end{pmatrix}$ $(n \geqslant 0)$, where the $\alpha_i, \beta_i$ are ordinals and $\beta_0 < \dots < \beta_n$. If $A$ and $B$ are KS, then

$A = B$ if and only if $A$ and $B$ contain exactly the same columns except (possibly) for columns of the form $\begin{smallmatrix}0\\ \beta\end{smallmatrix}$.

We define a lexicographic well-ordering $\prec$ on the KS as follows :

If $A = \begin{pmatrix}\alpha_0 \dots \alpha_n\\ \beta_0 \dots \beta_n\end{pmatrix}$ and B = $\begin{pmatrix}\gamma_0 \dots \gamma_n\\ \beta_0 \dots \beta_n\end{pmatrix}$, then $A \prec B$ if and only if $\alpha_i < \gamma_i$ for the largest $i \leqslant n$ such that $\alpha_i \neq \gamma_i$.

Now, for each KS $A$, we define an ordinal $fA$ by induction on $A$ with respect to $\prec$:

If $A = \begin{pmatrix}\alpha_0 & \alpha_1\\ 0 & 1\end{pmatrix} \prec \binom{1}{2}$, then $fA = \omega^{\omega^{-1+(\alpha_0 \# \alpha_1)}}$ (where $(-1+\alpha) = \alpha$ if $\alpha$ is infinite or zero; $\beta$ if $\alpha = \beta + 1$ and $\alpha$ is finite); if $A \geqq \binom{1}{2}$, where $A = \begin{pmatrix}\alpha_0 & \alpha_1 \dots \alpha_n\\ 0 & \beta_1 \dots \beta_n\end{pmatrix}$, then $fA$ is the $\alpha_0$-th common solution $\alpha$ of all equations $f\begin{pmatrix}\gamma & \alpha_1^* \dots \alpha_n\\ \beta_1^* & \beta_1 \dots \beta_n\end{pmatrix} = \gamma$ $(\alpha_1^* < \alpha_1, \beta_1^* < \beta_1)$.

*Examples* $f\begin{pmatrix}\alpha & 1\\ 0 & 2\end{pmatrix} = \varepsilon_\alpha$, the $\alpha$-th solution $\gamma$ of the equation $\omega^\gamma = \gamma$; $f\binom{2}{2} = \Gamma_0$, the ordinal defined in [8].

This notation system is taken from [24] with a minor alteration (in the value of $fA$ for $A \prec \binom{1}{2}$), and a minor generalisation (the fact that the ordinals which appear in a KS do not have to be finite or countable – the reader may check that Schütte's proofs go through just the same, and that the cardinality of $f\begin{pmatrix}\alpha_0 \dots \alpha_n\\ \beta_0 \dots \beta_n\end{pmatrix}$ is always equal to the maximum of the cardinalities of $\omega$ and $\alpha_0, \dots, \alpha_n, \beta_0, \dots, \beta_n$, provided, of course, that all the $\alpha_i$ are nonzero).

**Theorem 2.16** (Schütte) *For each $n \in \omega$ and all ordinals $\alpha_0, \dots, \alpha_n, \beta_0 < \dots < \beta_n$, writing $A = \begin{pmatrix}\alpha_0 \cdots \alpha_n\\ \beta_0 \cdots \beta_n\end{pmatrix}$,*

*(a) $fA$ is defined;*
*(b)* $\max\{\alpha_0, \dots, \alpha_n\} \leqslant fA$; *if* $\alpha_0 \neq 0$ *then* $\max\{\alpha_1, \dots, \alpha_n\} < fA$;
*(c) if $A \geqq \binom{1}{2}$, $fA = \alpha_0$, $\gamma < \alpha_q$ and $\gamma_1, \dots, \gamma_{q-1} < fA$, then*

$$f\begin{pmatrix}\alpha_0 & \gamma_1 & \dots & \gamma_{q-1} & \gamma & \alpha_{q+1} \cdots \alpha_n\\ \beta_0 & \beta_1 & \cdots & \beta_{q-1} & \beta_q & \beta_{q+1} \cdots \beta_n\end{pmatrix} = fA;$$

*(d) if $\gamma < \alpha_q$ and $\gamma_0, \dots, \gamma_{q-1}, \alpha_{q+1}, \dots, \alpha_n < fA$ and $A \geqq \binom{1}{2}$, then*

$$f\begin{pmatrix}\gamma_0 \dots \gamma_{q-1} & \gamma & \alpha_{q+1} \dots \alpha_n\\ \beta_0 \dots \beta_{q-1} & \beta_q & \beta_{q+1} \dots \beta_n\end{pmatrix} < fA;$$

*(e) if $\alpha_0 = \sup_{\alpha \in M} \alpha$ and $A \geqq \binom{1}{2}$, then*

$$f\begin{pmatrix}\alpha_0 \cdots \alpha_n \\ \beta_0 \cdots \beta_n\end{pmatrix} = \sup_{\alpha \in M} f\begin{pmatrix}\alpha\ \alpha_1 \cdots \alpha_n \\ \beta_0\ \beta_1 \cdots \beta_n\end{pmatrix}.$$

We now adapt $f$ a little to make sure that $\alpha_i < f\begin{pmatrix}\alpha_0 \cdots \alpha_n \\ \beta_0 \cdots \beta_n\end{pmatrix}$ for each $i \leqslant n$:

If $A = \begin{pmatrix}\alpha_0\ \alpha_1 \cdots \alpha_n \\ 0\ \beta_1 \cdots \beta_n\end{pmatrix}$, then

$$f^+A = \begin{cases} f\begin{pmatrix}\alpha_0 + 1\ \alpha_1 \ldots \alpha_n \\ 0 \quad \beta_1 \ldots \beta_n\end{pmatrix} & \text{if there is an } \alpha \text{ and an } m < \omega \\ & \text{such that } \alpha_0 = \alpha + m \text{ and} \\ & f\begin{pmatrix}\alpha\ \alpha_1 \cdots \alpha_n \\ 0\ \beta_1 \cdots \beta_n\end{pmatrix} = \alpha, \\ fA \text{ otherwise}. & \end{cases}$$

The following results then follow from Theorem 2.16:

**Theorem 2.17** (Schütte) *If* $A = \begin{pmatrix}\alpha_0 \cdots \alpha_n \\ \gamma_0 \cdots \gamma_n\end{pmatrix}$ *and* $B = \begin{pmatrix}\beta_0 \ldots \beta_n \\ \gamma_0 \ldots \gamma_n\end{pmatrix}$ *are Klammersymbols and* $\binom{1}{2} \leqq B$*, then*

*(a)* $\max(\beta_0, \ldots, \beta_n) < f^+B$;
*(b)* $A \prec B \ \&\ \max(\alpha_0, \ldots, \alpha_n) < f^+B \Rightarrow f^+A < f^+B$.

In fact, (a) and (b) could be used to define $f^+B$ for each Klammersymbol $B \succeq \binom{1}{2}$ by transfinite recursion on $B$ w.r.t.$\prec$, by taking $f^+B$ to be the least ordinal of the form $\omega^{\omega^\gamma}$ satisfying (a) and (b).

Note that, for any ordinals $\beta_0 < \cdots < \beta_n$, it follows from Theorem 2.16 that $f\begin{pmatrix}\alpha_0 \cdots \alpha_n \\ \beta_0 \cdots \beta_n\end{pmatrix}$ is weakly monotonic in $\alpha_0, \ldots, \alpha_n$ (i.e. increases or remains fixed if one of the $\alpha_i$ is increased and the others remain fixed), and from Theorem 2.17 that $f^+\begin{pmatrix}\alpha_0 \cdots \alpha_n \\ \beta_0 \cdots \beta_n\end{pmatrix}$ is strictly monotonic in $\alpha_0, \ldots, \alpha_n$ (i.e. increases if one of the $\alpha_i$ is increased and the others remain fixed). We shall often use these facts in the following chapters.

We shall need the following easy

**Lemma 2.18** *For any Klammersymbol* $A$, $fA$ *and* $f^+A$ *are closed under* $\#$ *and* $\otimes$.

*Proof* See Bachmann [2] for a proof that all ordinals of the form $\omega^{\omega^\gamma}$ are closed under $\#$ and $\otimes$, and note (proof by transfinite induction on $A$ w.r.t. $\prec$) that $fA$ has this form for every Klammersymbol $A$.

For the reader more accustomed to the Bachmann notation — see Bachmann [1] — we mention that, for countable or finite $\alpha_0, \ldots, \alpha_n, \beta_1, \ldots, \beta n$ such that $\binom{1}{2} \leqq \begin{pmatrix}\alpha_0\ \alpha_1 \cdots \alpha_n \\ 0\ \ \beta_1 \cdots \beta_n\end{pmatrix}$, $f\begin{pmatrix}\alpha_0 \cdots \alpha_n \\ 0 \cdots \beta_n\end{pmatrix}$ is approximately $\phi_{\Omega^{\beta_n}\cdot\alpha_n + \cdots + \Omega^{\beta_1}\cdot\alpha_1}(\alpha_0)$.

The reader who is daunted or bored by the ordinal notations may still benefit from the rest of this paper by adapting the theorems and proofs as indicated in the proof of Corollary 2.11 to obtain new proofs of theorems on well-quasi-orderedness.

## 3 The Maximal Order Types of Some Well-Quasi-Ordered Sets of Sequences and Trees

Kruskal [13] has proved that, if $(X, \leqslant)$ is a w.q.o., then the set of finite trees with labels in $X$ is well-quasi-ordered by $\leqslant_T$ (see Definitions 2.13 and 2.14). Higman [11] had previously proved the same result for the set of finite structured trees with labels in $X$ in which the number of immediate successors of each vertex is bounded. The aim of the next two sections is to calculate the maximal order types of these w.q.o.'s and, incidentally (see Corollary 2.11), to give new proofs of Higman's and Kruskal's results.

### *3.1 The Set of all $\leqslant n$-Branching Finite Structured Trees with Labels in a w.q.o. Set*

**Definition 3.1** A tree is *finite* if it has finitely many vertices. Suppose $(X_i, \leqslant_i)$ $(i = 0, \ldots, n)$ are w.q.o.'s. Then, writing $X = \bigcup_{i \leqslant n} X_i$ and $\leqslant = \bigcup_{i \leqslant n} \leqslant_i$, we define $T(X_0, \ldots, X_n)$ to be the set of all finite structured trees $\tau$ with labels in $X$ such that, for each vertex $V$ of $\tau$, $V$ has at most $n$ immediate successors and if $V$ has $i$ immediate successors then the label of $V$ is in $X_i$. The relation $\leqslant_T$ on $T(X_0, \ldots, X_n)$ is as defined in 2.14.

*Examples* $(T(X_0), \leqslant_T)$ is isomorphic to $(X_0, \leqslant_0)$, since $T(X_0)$ contains only trees with just one vertex. $(T(X_0, X_0), \leqslant_T)$ is isomorphic to $(X_0^*, \leqslant_0^*)$. $T(X_0$, empty set, $X_0)$ contains just all finite binary trees with labels in $X_0$.

**Theorem 3.2** *If all the $(X_i, \leqslant_i)$ are w.q.o.'s, then*

$$\mathrm{o}(T(X_0, \ldots, X_n), \leqslant_T) \leqslant f^+\begin{pmatrix} -1 + \mathrm{o}(X_0) & \mathrm{o}(X_1) \ldots \mathrm{o}(X_n) \\ 0 & 1 \ldots n \end{pmatrix}$$

*(where $-1 + \alpha = \alpha - 1$ if $\alpha$ is finite and nonzero; $\alpha$ otherwise).*

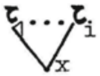
*Proof* By transfinite induction on $\begin{pmatrix} -1 + \mathrm{o}(X_0) & \mathrm{o}(X_1) \ldots \mathrm{o}(X_n) \\ 0 & 1 \ldots n \end{pmatrix}$ w.r.t. $\prec$. We shall write $T(X)$ for $T(X_0, \ldots, X_n)$ and $A$ for $\begin{pmatrix} -1 + \mathrm{o}(X_0) & \mathrm{o}(X_1) \ldots \mathrm{o}(X_n) \\ 0 & 1 \ldots n \end{pmatrix}$.

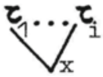
$A \prec \binom{1}{2}$: Then $(T(X), \leqslant_T)$ is isomorphic to $(X_0 \times X_1^*, \leqslant_0 \times \leqslant_1^*)$. Hence, by Theorems 2.7 and 2.9,

$$\begin{aligned} o(T(X), \leqslant_{\mathrm{T}}) &\leqslant o(X_0) \otimes \omega^{\omega^{(o(X_1))'}} \\ &\leqslant \omega^{\omega^{-1+o(X_0)}} \otimes \omega^{\omega^{(o(X_1))'}} \\ &\leqslant \omega^{(\omega^{-1+o(X_0)} \# \omega^{(o(X_1))'})} \\ &\leqslant \omega^{\omega^{((-1+o(X_0)) \# o(X_1))'}} \\ &= f^+ \begin{pmatrix} -1 + o(X_0) & o(X_1) \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

$A \geqq \binom{1}{2}$: Note that $T(X)$ can be defined inductively as follows:

(1) For any $x \in X_0$, $.x \in T(X)$ (where $.x$ denotes the tree with just one vertex whose label is $x$);
(2) for $0 < i \leqslant n$, any $x \in X_i$ and any $\tau_1, \ldots, \tau_i \in T(X)$

 ∈ T(X),

where  is the tree with $x$ as the label of its root obtained by taking the roots of $\tau_1, \ldots, \tau_i$ (in that order) as the immediate successors of its root. We shall denote this tree by $x(\tau_1, \ldots, \tau_i)$ for convenience' sake.

Since $T(X)$ can be defined inductively by (1) and (2) above, and since $o(T(X), \leqslant_T)$ is clearly a limit ordinal, by Corollary 2.4 it is sufficient to prove that

(a) for any $x \in X_0$, $l(.x) < f^+A$;
(b) for $0 < i \leqslant n$, any $x \in X_i$ and any $\tau_1, \ldots, \tau_i \in T(X)$ such that $l(\tau_j) < f^+A$ for all $j \leqslant i$,

$$l(x(\tau_1, \ldots, \tau_i)) < f^+A.$$

*Proof of (a)*: By transfinite induction on $o(X_0)$: $o(X_0) = 1$: Then $X_0$ contains just one element $x$, so $L_{X_0}(x)$ must be empty; hence so is $L_{T(X)}(.x)$; i.e. $l(.x) = 0 < f^+A$. $o(X_0) > 1$: $L(.x)$ contains only those trees which have no label $y$ such that $x \leqslant y$; hence, in particular,

$$L(.x) \subseteq T(L(x), X_1, \ldots, X_n).$$

But $l(x) < o(X_0)$ by Corollary 2.3. Hence, by induction hypothesis,

$$\begin{aligned} l(.x) &\leqslant f^+ \begin{pmatrix} -1 + l(x) & o(X_1) \cdots o(X_n) \\ 0 & 1 \cdots n \end{pmatrix} \\ &< f^+ \begin{pmatrix} -1 + o(X_0) & o(X_1) \ldots o(X_n) \\ 0 & 1 \ldots n \end{pmatrix} \qquad \text{by Theorem 2.17.} \end{aligned}$$

*Proof of (b)*: Suppose that $0 < i \leqslant n$, $x \in X_i$, $\tau_1, \ldots, \tau_i \in T(X)$ and $l(\tau_j) < f^+ A$ for all $j \leqslant i$.

Now let $\tau$ be an element of $L(x(\tau_1, \ldots, \tau_i))$, and let $V$ be any vertex of $\tau$. The subtree $\tau_V$ consisting of $V$ and all its successors must also be an element of $L(x(\tau_1, \ldots, \tau_i))$. But this is only possible if one of the following three conditions is satisfied:

(i) $V$ has $k$ immediate successors and the label of $V$ is in $X_k$, for some $k \neq i$;

(ii) $V$ has $i$ immediate successors and the label of $V$ is in $X_i \cap L(x)$;

(iii) $V$ has $i$ immediate successors, which are the roots of subtrees $\tau'_1, \ldots, \tau'_i$ respectively of $\tau$, and $x \leqslant_i y$, where $y$ is the label of $V$. Then $\tau'_j \in L(\tau_j)$ must hold for at least one $j \leqslant i$. In this case, we could regard $V$ as a vertex with $(i-1)$ immediate successors (i.e. the roots of $\tau'_1, \ldots, \tau'_{j-1}, \tau'_{j+1}, \ldots, \tau'_i$) and regard $\tau'_j$ as part of $V$'s label, where $V$'s label would now be $(y, \tau'_j)$. Thus $V$ could be regarded as a vertex with $(i-1)$ immediate successors and a label in $(X_i \times \dot{\bigcup}_{j \leqslant i} L(\tau_j))$.

Now, for any $\tau \in L(x(\tau_1, \ldots, \tau_i))$, *every* vertex $V$ of $\tau$ must satisfy one of (i)–(iii); but then, given the convention established in (iii),

$$\tau \in T\left(X_0, \ldots, X_{i-2}, \left[X_{i-1} \dot{\cup} \left(X_i \times \dot{\bigcup_{j \leqslant i}} L(\tau_j)\right)\right], L(x), X_{i+1}, \ldots, X_n\right).$$

We shall write $L = X_{i-1} \dot{\cup} (X_i \times \dot{\bigcup}_{j \leqslant i} L(\tau_j))$.

Hence $L(x(\tau_1, \ldots, \tau_i))$ may be identified with a subset of $T(X_0, \ldots, X_{i-2}, L, L(x), X_{i+1}, \ldots, X_n)$ and, what is more important, the q.o. induced on $L(x(\tau_1, \ldots, \tau_i))$ by this correspondence is a subset of $\leqslant_{\mathrm{T}} \restriction L(x(\tau_1, \ldots, \tau_i))$; hence

$$l(x(\tau_1, \ldots, \tau_i)) \leqslant \mathrm{o}(T(X_0, \ldots, X_{i-2}, L, L(x), X_{i+1}, \ldots, X_n), \leqslant_T).$$

But $o(L(x)) < \mathrm{o}(X_i)$ by Corollary 2.3; hence

$$\begin{pmatrix} -1 + o(X_0) & o(X_1) & \cdots & o(X_{i-2}) & o(L) & o(L(x)) & o(X_{i+1}) & \cdots & o(X_n) \\ 0 & 1 & \cdots & i-2 & i-1 & i & i+1 & \cdots & n \end{pmatrix} \prec$$
$$\prec \begin{pmatrix} -1 + o(X_0) & o(X_1) & \cdots & o(X_n) \\ 0 & 1 & \cdots & n \end{pmatrix}.$$

Hence, by induction hypothesis,

$$o(T(X_0, \ldots, X_{i-2}, L, L(x), X_{i+1}, \ldots, X_n)) \leqslant$$
$$\leqslant f^+ \begin{pmatrix} -1 + o(X_0) & o(X_1) & \cdots & o(L) & l(x) & \cdots & o(X_n) \\ 0 & 1 & \cdots i-1 & i & \cdots & n \end{pmatrix}.$$

But, by Theorem 2.17 (a), $-1 + o(X_0), o(X_1), \ldots, o(X_n) < f^+A$.

Moreover, by hypothesis, $l(\tau_j) < f^+A$ for all $j \leqslant i$.

Hence, by Theorems 2.6, 2.7 and Lemma 2.18, $o(L) < f^+A$.

But then we may apply Theorem 2.17 (b) to show that

$$f^+\begin{pmatrix} -1+o(X_0) & o(X_1) & \cdots & o(L) & l(x) & \cdots & o(X_n) \\ 0 & 1 & \cdots i-1 & i & \cdots & n \end{pmatrix} < f^+A,$$

from which it follows that $l(x(\tau_1, \ldots, \tau_i)) < f^+A$, q.e.d.

*Remark* The bounds of Theorem 2.2 are best possible for $A \geqq \binom{1}{2}$ – this will be proved in Sect. 4.1.

## *3.2 The Set of All Finite Structured Trees with Labels in a w.q.o. Set*

In this section we shall calculate upper bounds for the maximal order types of the w.q.o.'s in Kruskal's theorem [13]. It turns out to be more convenient to consider sets of trees a little different from those in Sect. 3.1: Instead of the label of a vertex determining the number of its immediate successors, the label only gives a strict upper bound to the number of immediate successors. Thus, in the course of estimating the maximal order types of the w.q.o.'s we are really interested in, we shall obtain further results about the sets of trees, with labels in a w.q.o. set, in which the number of successors of each vertex is bounded. These results do not, as far as I can see, easily imply or follow from the results of Sect. 3.1.

**Definition 3.3** Suppose $(X_\mathrm{i}, \leqslant_i)$ $(i = 0, \ldots, n)$ are w.q.o.'s such that the $X_i$ are pairwise disjoint and $0 < \alpha_0 < \cdots < \alpha_n \leqslant \omega$ are ordinals. Then, writing $X = \bigcup_{i \leqslant n} X_\mathrm{i}$ and $\leqslant = \bigcup_{i \leqslant n} \leqslant_i$, we define $T\begin{pmatrix} X_0 & \cdots & X_n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$ to be the set of all finite structured trees $\tau$ with labels in $X$ such that, for each vertex $V$ of $\tau$, if the label of $V$ is in $X_i$ then $V$ has fewer than $\alpha_i$ immediate successors.

*Examples* $T\begin{pmatrix} X_0 & \cdots & X_n \\ 1 & \cdots & n+1 \end{pmatrix} = T(X_0 \cup \ldots \cup X_n, X_1 \cup \cdots \cup X_n, \ldots, X_n)$. $T\begin{pmatrix} X_0 \\ \omega \end{pmatrix}$ is just the set of all finite trees with labels in $X_0$.

**Theorem 3.4** *If all the $(X_\mathrm{i}, \leqslant_i)$ are w.q.o.'s, the $X_\mathrm{i}$ are pairwise disjoint and $0 \leqslant \alpha_0 < \cdots < \alpha_n \leqslant \omega$, then*

$$o\left(T\begin{pmatrix} X_0 & \cdots & X_n \\ 1+\alpha_0 & \cdots & 1+\alpha_n \end{pmatrix}, \leqslant_T\right) \leqslant f^+\begin{pmatrix} o(X_0) & \cdots & o(X_n) \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}.$$

*Proof* By transfinite induction on $\begin{pmatrix} o(X_0) & \cdots & o(X_n) \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix} w.r.t. \prec$.

We shall write $T(X)$ for $T\begin{pmatrix} X_0 & \cdots & X_n \\ 1+\alpha_0 & \cdots & 1+\alpha_n \end{pmatrix}$ and $A$ for $\begin{pmatrix} o(X_0) & \cdots & o(X_n) \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$.

$A \prec \binom{1}{2}$: Then $(T(X), \leqslant_T)$ is isomorphic to $(B, \leqslant^*)$ for some subset B of $(X_0 \dot{\cup} X_1)^*$. Hence, by Theorem 2.9,

$$\begin{aligned} o(T(X), \leqslant_T) &\leqslant \omega^{\omega^{(o(X_0 \dot{\cup} X_1))'}} \\ &= \omega^{\omega^{(o(X_0) \# o(X_1))'}} \qquad \text{by Theorem 2.6} \\ &= f^+ \begin{pmatrix} o(X_0) & o(X_1) \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

$A \geqq \binom{1}{2}$: Note that $T(X)$ can be defined inductively as follows:

(1) For any $x \in X$, $.x \in T(X)$;
(2) for any $i \leqslant n$, any $x \in X_i$, any $m < 1+\alpha_i$ and any

$$\tau_1, \dots, \tau_m \in T(X), \quad x(\tau_1, \dots, \tau_m) = \underset{x}{\tau_1 \cdots \tau_m} \in T(X).$$

Hence, since $o(T(X), \leqslant_T)$ is clearly a limit ordinal, by Corollary 2.4 it is sufficient to prove that

(a) for any $x \in X$, $l(.x) < f^+A$;
(b) for any $i \leqslant n$, any $x \in X_i$, any m $< 1+\alpha_i$ and any $\tau_1, \dots, \tau_m \in T(X)$ such that $l(\tau_j) < f^+A$ for all $j \leqslant m$, $l(x(\tau_1, \dots, \tau_m)) < f^+A$.

*Proof of* (a): Suppose $x \in X_i$. Then no element of $L(.x)$ can have any label $y$ such that $x \leqslant y$; in other words,

$$L(.x) \subseteq T\begin{pmatrix} X_0 & \cdots & X_{i-1} & L(x) & X_{i+1} & \cdots & X_n \\ 1+\alpha_0 & \cdots & 1+\alpha_{i-1} & 1+\alpha_i & 1+\alpha_{i+1} & \cdots & 1+\alpha_n \end{pmatrix}.$$

But then, by induction hypothesis,

$$\begin{aligned} l(.x) &\leqslant f^+ \begin{pmatrix} oX_0 & \cdots & oX_{i-1} & l(x) & oX_{i+1} & \cdots & oX_n \\ \alpha_0 & \cdots & \alpha_{i-1} & \alpha_i & \alpha_{i+1} & \cdots & \alpha_n \end{pmatrix} \\ &< f^+A \text{ by Theorem 2.17.} \end{aligned}$$

*Proof of* (b): Suppose that $i \leqslant n$, $x \in X_i$, $m < 1+\alpha_i$ and $1(\tau_j) < f^+A$ for all $j \leqslant m$. We shall assume w.l.o.g. that there is a $k < i$ such that $m = 1+\alpha_k$ (if not, just add an empty $X_k$).

Now let $\tau$ be an element of $L(x(\tau_1, \dots, \tau_m))$, and let $V$ be any vertex of $\tau$. The subtree of $\tau_V$ of $\tau$ consisting of $V$ and all its successors must also be an element of $L(x(\tau_1, \dots, \tau_m))$. But this is only possible if at least one of the following four conditions is satisfied:

(i) The label of $V$ is not in $X_i$;
(ii) the label of $V$ is in $X_i \cap L(x)$;
(iii) the label of $V$ is some $y_i \geqslant x$, but $V$ has fewer than $m = (1 + \alpha_k)$ immediate successors;
(iv) the label of $V$ is some $y_i \geqslant x$ and $V$ has at least $m$ immediate successors, i.e. $\tau_V = y(\tau'_1, \dots, \tau'_q)$, where $q \geqslant m$, but there is no subsequence $\langle i_1, \dots, i_m \rangle$ of $\langle 1, \dots, q \rangle$ such that $\tau_j \leqslant_T \tau'_{i_j}$ for all $j \leqslant m$. This means that if we define

$i_1$ to be the smallest $i$ such that $\tau_1 \leqslant_T \tau'_i$,
$i_2$ to be the smallest $i > i_1$ such that $\tau_2 \leqslant_T \tau'_i$,
...
$i_m$ to be the smallest $i > i_{m-1}$ such that $\tau_m \leqslant_T \tau'_i$,

then $i_m$ will certainly not be defined, and some of the other $i_j$ may not be either. But, supposing for example that $i_{m-1}$ is defined,

$$\begin{aligned} &\tau'_1, \dots, \tau'_{i_1 - 1} \in L(\tau_1), \\ &\tau'_{i_1+1}, \dots, \tau'_{i_2-1} \in L(\tau_2), \\ &\dots \\ &\tau'_{i_{m-1}+1}, \dots, \tau'_q \in L(\tau_m). \end{aligned}$$

In this case, we could regard $V$ as a vertex with fewer than $m$ immediate successors (i.e. the roots of $\tau'_{i_1}, \dots, \tau'_{i_{m-1}}$) and treat the trees growing out of all the other immediate successors as if they were part of V's label. In other words (assuming for the moment that $i_{m-1}$ is defined) we could treat V as a vertex whose immediate successors are the roots of $\tau'_{i_1}, \dots, \tau'_{i_{m-1}}$ and whose label is

$$(y, \tau'_1, \dots, \tau'_{i_1-1}, \tau'_{i_1+1}, \dots, \tau'_{i_2-1}, \dots, \tau_{i_{m-1}+1} + 1', \dots, \tau'_q).$$

This label could be regarded as an element of

$$X_i \times (L(\tau_1))^* \times (L(\tau_2))^* \times \cdots \times (L(\tau_m))^*.$$

In general, $i_{m-1}, i_{m-2}, \dots, i_1$ may not be defined; so in general we shall regard $V$ as a vertex with fewer than $m$ $(= 1 + \alpha_k)$ immediate successors and a label in

$$X_i \times \dot{\bigcup_{j=1}^{m}} ((L(\tau_1))^* \times \cdots \times (L(\tau_j))^*)$$

(we shall denote this set by $L$).

Now, for any $\tau \in L(x(\tau_1, \dots, \tau_m))$, *every* vertex $V$ of $\tau$ must satisfy at least one of (i)–(iv); but then, given the convention established in (iv),

$$\tau \in T \begin{pmatrix} \cdots & X_k \dot{\cup} X_i \dot{\cup} L & \cdots & L(x) \cdots \\ \cdots & 1 + \alpha_k & \cdots & 1 + \alpha_i \cdots \end{pmatrix},$$

where the dots indicate columns identical with the corresponding ones in $\begin{pmatrix} X_0 \cdots X_n \\ 1+\alpha_0 \cdots 1+\alpha_n \end{pmatrix}$.

Hence $L(x(\tau_1, \ldots, \tau_m))$ may be identified with a subset of $T\begin{pmatrix} \cdots & X_k \dot{\cup} X_i \dot{\cup} L & \cdots & L(x) & \cdots \\ \cdots & 1+\alpha_k & \cdots & 1+\alpha_i & \cdots \end{pmatrix}$, and, what is more important, the q.o. induced on $L(x(\tau_1 \ldots \tau_m))$ by this identification is a subset of $\leqslant_T \restriction L(x(\tau_1, \ldots, \tau_m))$ hence

$$l(x(\tau_1, \ldots, \tau_m)) \leqslant o\left(T\begin{pmatrix} \cdots & X_k \dot{\cup} X_i \dot{\cup} L & \cdots & L(x) & \cdots \\ \cdots & 1+\alpha_k & \cdots & 1+\alpha_i & \cdots \end{pmatrix}, \leqslant_T\right).$$

But $o(L(x)) < o(X_i)$ by Corollary 2.3; hence

$$\begin{pmatrix} \cdots & o(X_k \dot{\cup} X_i \dot{\cup} L) & \cdots & o(L(x)) & \cdots \\ \cdots & \alpha_k & \cdots & \alpha_i & \cdots \end{pmatrix} \prec \begin{pmatrix} o(X_0) & \cdots & o(X_n) \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}.$$

Hence, by induction hypothesis,

$$o\left(T\begin{pmatrix} \cdots & X_k \dot{\cup} X_i \dot{\cup} L & \cdots & L(x) & \cdots \\ \cdots & 1+\alpha_k & \cdots & 1+\alpha_i & \cdots \end{pmatrix}\right) \leqslant$$
$$\leqslant f^+\begin{pmatrix} o(X_0) & \cdots & o(X_i \dot{\cup} X_k \dot{\cup} L) & \cdots & l(x) & \cdots & o(X_n) \\ \alpha_0 & \cdots & \alpha_k & \cdots & \alpha_i & \cdots & \alpha_n \end{pmatrix}.$$

But, by Theorem 2.17(a),

$$o(X_0), \ldots, o(X_n) < f^+\begin{pmatrix} o(X_0) & \ldots & o(X_n) \\ \alpha_0 & \ldots & \alpha_n \end{pmatrix} = f^+ A.$$

Moreover, by hypothesis $l(\tau_j) < f^+ A$ for all j $\leqslant$ m.
But then

$$o(L(\tau_j)^*) = f^+\begin{pmatrix} L(\tau_j) \\ 1 \end{pmatrix} \text{ by Theorem 2.9}$$
$$< f^+ A \text{ by Theorem 2.17b).}$$

Hence, by finitely many applications of Theorems 2.6, 2.7 and Lemma 2.18, $o(X_k \dot{\cup} X_i \dot{\cup} L)$
$< f^+ A$.

But then we may apply Theorem 2.17 (b) to show that

$$f^+\begin{pmatrix} o(X_0) & \cdots & o(X_k \dot{\cup} X_i \dot{\cup} L) & \cdots & l(x) & \cdots & o(x_n) \\ \alpha_0 & \cdots & \alpha_k & \cdots & \alpha_i & \cdots & \alpha_n \end{pmatrix} < f^+ A,$$

from which it follows that $l(x(\tau_1, \ldots, \tau_m)) < f^+ A$, q.e.d.

### *3.3 Finitary Sequences of Elements of a w.q.o. Set*

Nash–Williams [16] proved that, if $(X, \leqslant)$ is a w.q.o., then the class of all those sequences of elements of $X$ which contain only finitely many different terms is well-quasi- ordered by the q.o. induced on it by $\leqslant$. In our notation (see Definition 2.12), if $(X, \leqslant)$ is a w.q.o. then so is $(\mathrm{S}^{\mathrm{F}}_{\alpha}(X), \leqslant_{\mathrm{S}})$ for each ordinal $\alpha$. The aim of this section is to provide a more transparent proof of this result via de Jongh and Parikh's theorem and to give upper bounds for the maximal order type of $(\mathrm{S}^{\mathrm{F}}_{\alpha}(X), \leqslant_{\mathrm{S}})$ in terms of $o(X)$.

*Notation* We shall denote $\mathrm{o}(\mathrm{S}^{\mathrm{F}}_{\alpha}(X), \leqslant_{\mathrm{S}})$ by $\mathrm{o}^{F}_{\alpha}(X)$.
The upper bounds we shall obtain for the $o^{F}_{\alpha}(X)$ are in general by no means the best possible; e.g. Fit is easy to see that $o^{F}_{\omega+1}(X) \leqslant \omega^{\omega^{n-1}}.(2^n - 1) + 1$ if $o(X) = n < \omega$. Moreover, in general, $o^{F}_{\alpha}(X)$ depends not only on $o(X)$ but also on how far $(X, \leqslant)$ is from being a well-ordering.

*Example* Suppose that $o(X, \leqslant) = 2$.
(a) If $(X, \leqslant)$ is a well-ordering, say $X = \{0, 1\}$, where $0 < 1$, then any $f \in S^{F}_{\omega+1}(X)$ which contains infinitely many occurrences of 1 satisfies $g \leqslant_S f$ for all $g \in S^{F}_{\omega+1}(X)$. Thus

$$o^{F}_{\omega+1}(X) = o(\{f | f \in S^{F}_{\omega+1}(X) \text{ and } f \text{contains 1 only finitely often}\}, \leqslant_S) + 1.$$

But any $f \in S^{F}_{\omega+1}(X)$ containing 1 only finitely often is either a finite sequence or a finite sequence followed by infinitely many 0's. Hence it is fairly easy to see that $o^{F}_{\omega+1}(X) = (\omega^{\omega} \# \omega^{\omega}) + 1 = \omega^{\omega}.2 + 1$.

(b) If $(X, \leqslant)$ is not a well-ordering, say $X = \{0, 1\}$ where 0 and 1 are $\leqslant$-incomparable, then for any $f \in S^{F}_{\omega+1}(X)$ $g \leqslant_S f$ for all $g \in S^{F}_{\omega+1}(X)$ if and only if $f$ contains infinitely many occurrences of both 0 and 1.

Thus $o^{F}_{\omega+1}(X) = o(\{f | f \in S^{F}_{\omega+1}(X) \text{ and } f \text{ does not contain both 0 and 1 only infinitely often}\}, \leqslant_S) + 1.$

But any $f \in S^{F}_{\omega+1}(X)$ which does not contain both 0 and 1 infinitely often is either a finite sequence or a finite sequence followed by infinitely many 0's or a finite sequence followed by infinitely many 1's. Hence it is fairly easy to see that $o^{F}_{\omega+1}(X) = (\omega^{\omega} \# \omega^{\omega} \# \omega^{\omega}) + 1 = \omega^{\omega}.3 + 1$.

Note that this situation (i.e. that $o^{F}_{\alpha}(X)$ does not depend only on $\alpha$ and $o(X)$) differs from that in Sect. 3.1 – there the upper bound calculated for the maximal order type of $T(X_0, \ldots, X_n)$ is attained even when the $X_i$ are well-ordered (see Theorem 4.9 of Sect. 4), and hence a fortiori when they are not.

After these remarks we proceed to prove the results mentioned above.

**Lemma 3.5** *If $f \in S_\alpha^F(X)$, where $\alpha$ is either a limit ordinal or the successor of a limit ordinal, then there is a $g \in S_\alpha^F(X)$ such that $f \leqslant_S g$ and $g$ consists of a sequence $X_1 \dots X_n$ ($n \geqslant 1$) repeated $\beta$ times for some $\beta < \alpha$.*

*Proof* Let $x_1, \dots, x_n$ be the elements of $X$ which occur in $f$. If $f$ has length $\beta < \alpha$ and $g$ is the sequence obtained by repeating $X_1 \dots X_n$ $\beta$ times, then clearly $f \leqslant_S g$. Now $g$ has length $m.\beta$. But under the conditions of the lemma. $\beta < \alpha \Rightarrow \beta < \alpha$ hence $g \in S_\alpha^F(X)$.

**Theorem 3.6** *If $(X, \leqslant)$ is a w.q.o., then so is $(\mathrm{S}_\alpha^{\mathrm{F}}(X), \leqslant_{\mathrm{S}})$ for any ordinal $\alpha$.*

*Proof* By transfinite induction on $\alpha$ with subsidiary transfinite induction on $o(X)$.

$\alpha = 0$: Then $\mathrm{S}_\alpha^F(X)$ is empty.

$\alpha = 1$: Then $\mathrm{S}_\alpha^F(X)$ contains just the empty sequence.

$\alpha = 2$: Then $\mathrm{S}_\alpha^F(X)$ contains just all elements of $X$ and the empty sequence.

$o(X) = 0$: Then $\mathrm{S}_\alpha^F(X)$ contains at most the empty sequence.

$\alpha = \beta + \gamma$, *where* $\beta + 1, \gamma < \alpha$: Then $\mathrm{S}_\alpha^F(X)$ may be identified with a subset of $\mathrm{S}_{\beta+1}^F(X) \times S_\gamma^F(X)$, since each sequence of length less than $\alpha$ may be split up into a sequence of length less than or equal to $\beta$ followed by a sequence of length less than $\gamma$. Moreover, the q.o. thus induced on $\mathrm{S}^F(X)$ by the q.o. $\leqslant_S \times \leqslant_S$ on $\mathrm{S}_{\beta+1}^F(X) \times S_\gamma^F(X)$ is a subset of $\leqslant_s \restriction S^F(X)$. Thus, by Theorem 2.7 and the induction hypothesis, $\mathrm{S}_\alpha^F(X), \leqslant_S$ is a w.q.o.

$\alpha \in \{\beta, \beta + 1\}$, *where* $\beta > 1$ *is* closed under addition and is hence *a limit ordinal*: By Lemma 2.10 and Lemma 3.5, it is sufficient to prove that, for each $\gamma < \alpha$ and each $f$ consisting of some sequence $x_1 \dots x_n (x_1, \dots, x_n \in X)$ repeated $\gamma$ times, $(L(f), \leqslant_S)$ is a w.q.o.

Now suppose $g \in L(f)$. Then either $g$ has length less than $\gamma$, i.e. $g \in S_\gamma^F(X)$; or $g$ has length at least $\gamma$ but does not contain any $\gamma$-sub-sequence of elements $h_\delta (\delta < \gamma)$ of $\mathrm{S}_\alpha^F(X)$ such that $x_1 \dots x_n \leqslant_S h_\delta$ for each $\delta < \gamma$. In this case we may regard $g$ as a sequence of fewer than $\gamma$ elements of $(L_{S_\alpha^F(X)}(x_1 \cdots x_n) \times X)$ followed by one more element of $L_{S_\alpha^F(X)}(x_1 \dots x_n)$, and the ordering induced by this identification is no stronger than the old one. But $L_{S_\alpha^F(X)}(x_1 \dots x_n) \subseteq L_{x_1 \dots x_n}$, where $L_{x_1 \dots x_n}$ denotes

$$\dot{\bigcup_{i=0}^{n}} \left[S_\alpha^F(L(X_1)) \times X \times \dots \times X \times S_\alpha^F(L(X_i))\right].$$

Hence $\leqslant_S$ on $L(f)$ is isomorphic to an extension of the obvious ordering on a subset of $\mathrm{S}_\gamma^F(X) \dot{\cup} (S_\gamma^F(L_{x_1 \dots x_n} \times X) \times L_{x_1 \dots x_n})$.

But, for each $i \leqslant n$, $\mathrm{S}_\alpha^F(L(x_i))$ is well-quasi-ordered by $\leqslant_S$ by subsidiary induction hypothesis and Corollary 2.3; hence $L_{x_1 \dots x_n}$ and $L_{x_1 \dots x_n} \times X$ are well-quasi-ordered by the obvious ordering by Theorems 2.6 and 2.7. Hence so are $\mathrm{S}_\gamma^F(X)$ and $\mathrm{S}_\gamma^F(L_{x_1 \dots x_n} \times X)$ by the main induction hypothesis. Hence, by Theorems 2.6 and 2.7, $(L(f), \leqslant_S)$ is a w.q.o., q.e.d.

*Remark 1* In the literature, the term 'w.q.o.' is often applied to proper *classes* together with a binary relation, and not just to sets. For example, if we define $\mathrm{S}^F(X)$ as the class of *all* (well-ordered) sequences of finitely many elements of X, then Nash–Williams' result is that, if $(X, \leqslant)$ is a w.q.o., then $\mathrm{S}^F(X)$ is well-quasi-ordered by $\leqslant_S$. This clearly follows from Theorem 3.6 above, since any infinite set of $\leqslant_S$-incomparables in $\mathrm{S}^F(X)$ would be contained in $\mathrm{S}^F_\alpha(X)$ for some ordinal $\alpha$. However, it does not make sense to talk about $o(S^F(X), \leqslant_S)$, since for any extension $\leqslant_S^+$ of $\leqslant_S$ to a well-ordering of $\mathrm{S}^F(X)$ $(\mathrm{S}^\mathrm{F}(X), \leqslant_\mathrm{S}^+)$ is isomorphic to the class of all ordinals with the usual well-ordering. So we now proceed to give upper bounds for $o^F_\alpha(X)$ in terms of o(X), for each ordinal $\alpha$.

**Lemma 3.7** *If $\alpha = \beta + \gamma$, then $o^F_\alpha(X) \leqslant (o^F_{\beta+1}(X)) \otimes (o^F_\gamma(X))$.*

*Proof* Immediate by Theorem 2.7 and the observations in the proof of the relevant case of Theorem 3.6.

**Lemma 3.8** *If $\alpha$ is a limit ordinal or the successor of a limit ordinal, then*

$$o^F_\alpha(X) \leqslant \sup_{\substack{\gamma<\alpha \\ m\in\omega \\ x_1,\dots,x_m\in X}} \left[o^F_\gamma(X) \# (o^F_\gamma(L_{x_1\dots x_m} \times X) \otimes o(L_{x_1\dots x_m}))\right]$$

*(see proof of Theorem 3.6 for the definition of $L_{x_1\dots x_m}$).*

*Proof* Immediate by Theorems 2.6 and 2.7 and the proof of the relevant case of Theorem 3.6.

**Theorem 3.9** *If $\alpha = \omega^{1+\beta_n} + \dots + \omega^{1+\beta_0}$, where $0 \leqslant \beta_0 \leqslant \beta_1 \leqslant \dots \leqslant \beta_{n-1} \leqslant \beta_n$, then*

$$o^F_\alpha(X) \leqslant f^+\binom{o(X)}{2\beta_n+2} \otimes \dots \otimes f^+\binom{o(X)}{2\beta_1+2} \otimes f^+\binom{o(X)}{2\beta_0+1}, \quad \textit{and}$$

$$o^F_{\alpha+1}(X) \leqslant f^+\binom{o(X)}{2\beta_n+2} \otimes \dots \otimes f^+\binom{o(X)}{2\beta_1+2} \otimes f^+\binom{o(X)}{2\beta_0+2}.$$

*Proof* By transfinite induction on $\alpha$ with subsidiary transfinite induction on $o(X)$.

(i) $\alpha = \omega^{1+\beta_n} + \dots + \omega^{1+\beta_0}$, *where* $n > 0$: In this case, by Lemma 3.7,

$$\begin{aligned} o^F_\alpha(X) &\leqslant o^F_{\omega^{1+\beta_n}+1}(X) \otimes \dots \otimes o^F_{\omega^{1+\beta_1}+1}(X) \otimes o^F_{\omega^{1+\beta_0}}(X) \\ &\leqslant f^+\binom{o(X)}{2\beta_n+2} \otimes \dots \otimes f^+\binom{o(X)}{2\beta_1+2} \otimes f^+\binom{o(X)}{2\beta_0+1} \end{aligned}$$

by (main) induction hypothesis; and, again by Lemma 3.7,

$$o^F_{\alpha+1}(X) \leqslant o^F_{\omega^{1+\beta_{n+1}}}(X) \,\#\!\!\#\, \dots \,\#\!\!\#\, o^F_{\omega^{1+\beta_{1+1}}}(X) \,\#\!\!\#\, o^F_{\omega^{1+\beta_{0+1}}}(X)$$
$$\leqslant f^+\binom{o(X)}{2\beta_n+2} \,\#\!\!\#\, \dots \,\#\!\!\#\, f^+\binom{o(X)}{2\beta_1+2} \,\#\!\!\#\, f^+\binom{o(X)}{2\beta_0+2}$$

by (main) induction hypothesis.

(ii) $\alpha = \omega = \omega^{1+0}$ :$o^F_\omega(X) = o(S^F_\omega(X), \leqslant_S) = o(X^*, \leqslant^*)$
$$= f^+\binom{o(X)}{1} = f^+\binom{o(X)}{2.0+1};$$

$o^F_{\omega+1}(X) \leqslant f^+\binom{o(X)}{2}$ follows from this exactly as in the following case.
(iii) $\alpha = \omega^{1+\beta}$, where $\beta > 0$ and $o(X) > 0$ :

Suppose that $\gamma < \alpha$, $m \in \omega$ and $x_1, \dots, x_n \in X$. We first show that $o^F_\alpha(X) \leqslant f^+\binom{o(X)}{2+1}$.

By Lemma 3.8 it is sufficient to prove that

$$o^F_\gamma(X) \,\#\, (o^F_\gamma(L_{x_1\dots x_n} \times X) \,\#\!\!\#\, o(L_{x_1\dots x_n})) < f^+\binom{o(X)}{2\beta+1}.$$

Now since $\gamma < \alpha$ there is an $n \in \omega$ and $\beta_0 \leqslant \dots \leqslant \beta_n < \beta$ such that $\gamma \leqslant \omega^{1+\beta_n} + \dots + \omega^{1+\beta_0}$.

Hence, by (main) induction hypothesis,

$$o^F_\gamma(X) \leqslant f^+\binom{o(X)}{2\beta_n+2} \,\#\!\!\#\, \dots \,\#\!\!\#\, f^+\binom{o(X)}{2\beta_1+2} \,\#\!\!\#\, f^+\binom{o(X)}{2\beta_0+1}$$
$$< f^+\binom{o(X)}{2\beta+1} \text{ by Theorem 2.17 and Lemma 2.18.}$$

Now $L_{x_1\dots x_m} = \dot{\bigcup}_{i=0}^{m} \left[S^F_\alpha(L(x_1)) \times X \times \dots \times X \times S^F_\alpha(L(x_i))\right]$.

Hence

$$o(L_{x_1\dots x_m}) \leqslant \mathop{\#}_{i=0}^{m} \left[o^F_\alpha(L(x_1)) \,\#\!\!\#\, o(X) \,\#\!\!\#\, \dots \,\#\!\!\#\, o(X) \,\#\!\!\#\, o^F_\alpha(L(x_i))\right]$$
by Theorems 2.6 and 2.7.

But $o(X) < f^+\binom{o(X)}{2\beta+1}$ by Theorem 2.17a); and for each $j \leqslant m$

$$o_\alpha^F(L(x_j)) \leqslant f^+\binom{l(x_j)}{2\beta+1} \text{ by (subsidiary) induction hypothesis}$$
$$< f^+\binom{o(X)}{2\beta+1} \text{ by Theorem 2.17b).}$$

Hence, since by Lemma 2.18 $f^+\binom{o(X)}{2\beta+1}$ is closed under $\#$ and $\otimes$,

$$o(L_{x_1\ldots x_m}) < f^+\binom{o(X)}{2\beta+1} \text{ and } o(L_{x_1\ldots x_m} \times X) < f^+\binom{o(X)}{2\beta+1}.$$

But, by (main) induction hypothesis,

$$o_\gamma^F(L_{x_1\ldots x_m} \times X) \leqslant f^+\binom{o(L_{x_1\ldots x_m} \times X)}{2\beta_n+2} \otimes \cdots \otimes f^+\binom{o(L_{x_1\ldots x_m} \times X)}{2\beta_0+2}.$$

But, by the above and Theorem 2.17b),

$$f^+\binom{o(L_{x_1\ldots x_m} \times X)}{2\beta_k+2} < f^+\binom{o(X)}{2\beta+1} \text{ for each } k \leqslant n.$$

Hence, since $f^+\binom{o(X)}{2\beta+1}$ is closed under $\otimes$ by Lemma 2.18, $o_\gamma^F(L_{x_1\ldots x_m} \times X) < f^+\binom{o(X)}{2\beta+1}$.
Thus we have shown that
$o_\gamma^F(X) < f^+\binom{o(X)}{2\beta+1}$, $o(L_{x_1\ldots x_m}) < f^+\binom{o(X)}{2\beta+1}$ and $o_\gamma^F(L_{x_1\ldots x_m} \times X) < f^+\binom{o(X)}{2\beta+1}$.
Hence, since by Lemma 2.18 $f^+\binom{o(X)}{2\beta+1}$ is closed under $\#$ and $\otimes$,

$$o_\gamma^F(X) \# o_\gamma^F(L_{x_1\ldots x_m} \times X) \otimes o(L_{x_1\ldots x_m}) < f^+\binom{o(X)}{2\beta+1}, \text{ q.e.d.}$$

Now we show that $o_{\alpha+1}^F(X) \leqslant f^+\binom{o(X)}{2\beta+2}$.
By Lemma 3.8 it is sufficient to show that, for each $m \in \omega$ and all $x_1, \ldots, x_m \in X$,

$$o_\alpha^F(X) \# (o_\alpha^F(L_{x_1\ldots x_m} \times X) \otimes o(L_{x_1\ldots x_m})) < f^+\binom{o(X)}{2\beta+2}$$

where $L_{x_1\ldots x_m} = \dot{\bigcup}_{i=0}^{m} \left[S_{\alpha+1}^F(L(x_1)) \times X \times \ldots \times X \times S_{\alpha+1}^F(L(x_i))\right]$.

Now $o_\alpha^F(X) \leqslant f^+\binom{o(X)}{2\beta+1} < f^+\binom{o(X)}{2\beta+2}$;

$$o(L_{x_1\ldots x_m}) \leqslant \#_{i=0}^{m}\left[o_{\alpha+1}^F(L(x_1)) \otimes o(X) \otimes \ldots \otimes o(X) \otimes o_{\alpha+1}^F(L(x_i))\right]$$
$$< f^+\binom{o(X)}{2\beta+2}, \text{ since } o(X) < f^+\binom{o(X)}{2\beta+2} \text{ and}$$

$o_{\alpha+1}^F(L(x_j)) \leqslant f^+\left(l(x_j)2\beta+2\right) < f^+\left(o(X)2\beta+2\right)$ for each $j \leqslant m$ by (subsidiary)induction hypothesis;

and

$$o_\alpha^F(L_{x_1\ldots x_m} \times X) \leqslant f^+\binom{o(L_{x_1\ldots x_m}) \otimes o(X)}{2\beta+1} < f^+\binom{o(X)}{2\beta+2}.$$

Hence $o_\alpha^F(X) \# (o_\alpha^F(L_{x_1\ldots x_m} \times X) \otimes o(L_{x_1\ldots x_m})) < f^+\binom{o(X)}{2\beta+2}$, q.e.d.

(iv) $o(x) = 0$ : Then for each ordinal $\alpha S_\alpha^F(X)$ contains at most the empty sequence, so $o_\alpha^F(X) \leqslant 1 < \omega = f^+\binom{0}{1}$ from which the assertion follows by Theorem 2.17.

At this point a digression on better-quasi-orderings (b.q.o.) seems appropriate. Nash–Williams [17] coined this notion (all well-orderings are b.q.o.'s and all b.q.o.'s are w.q.o.'s, but the reverse inclusions do not hold) and it has proved to be just the right notion to use in proofs of statements of the form 'if $A$ is a well-ordering, then $B$ is a w.q.o.', by proving 'if $A$ is a b.q.o., then $B$ is also a b.q.o.'. Examples are Nash–Williams' theorem that if $(X, \leqslant)$ is a b.q.o., then so is $(\mathrm{S}_\alpha(X), \leqslant_S)$ for each ordinal $\alpha$ [18] and Laver's theorem that if $(X, \leqslant)$ is a b.q.o. then so is the class of all (finite and infinite) trees with labels in $X$ under the relation $\leqslant_T$ (proved for $X$ a singleton in Nash–Williams [17]). Both these theorems are false if 'b.q.o.' is replaced by 'w.q.o.' (for a counterexample see [20]). We have not investigated the question whether the methods of this paper can be adapted to be applied to b.q.o.'s and to obtain such results and give bounds for the corresponding maximal order types.

## 4 Applications of the Results in Sect. 3

### *4.1 Monotonic Increasing Ordinal Functions*

Reference [12] have already given an interesting application of their result Theorem 2.9 to hierarchies of (e.g. recursive) functions, and [23] contains another application of de Jongh and Parikh's results and methods. We shall now apply the results of Sect. 3.1 to obtain upper bounds for the order types of the sets of ordinals generated by certain ordinal functions. These results have a bearing on questions about systems of notations for ordinals, since such systems of notations are in general sets of terms (as in Definition 4.1) corresponding to the set of ordinals generated by one or more ordinal functions. Thus the results below show that, if such a system of notations is to reach $f^{+}\binom{1}{\omega}$, the ordinal functions used cannot be both monotonic and increasing (see Definition 4.5). In fact, there are two main lines of work which have produced systems of notations for such large ordinals; the Bachmann approach [1] makes use of functions which are (essentially) monotonic but not increasing, and the Takeuti ordinal diagram approach [25] uses functions which are increasing but not monotonic.

We shall use our considerations on monotonic increasing ordinal functions to show that the upper bounds calculated in Sect. 3.1 for the maximal order types of the w.q.o.'s $(T(X_0, \ldots, X_n), \leqslant_T)$ are best possible.

**Definition 4.1** Let $X_0, \ldots, X_n$ be any sets. Define $\mathrm{Term}(X_0, \ldots, X_n)$, the set of terms generated by $X_0, \ldots, X_n$, inductively as follows:

If $x \in X_0$, then $x \in \mathrm{Term}(X_0, \ldots, X_n)$;

If $x \in X_i$ and $t_1, \ldots, t_i \in \mathrm{Term}(X_0, \ldots, X_n)$, then $f_x(t_1, \ldots, t_i) \in \mathrm{Term}(X_0, \ldots, X_n)$.

Thus $\mathrm{Term}(X_0, \ldots, X_n)$ is obtained by associating an i-ary function symbol with each element of $X_i$ and generating the set of terms which can be formed using all these function symbols (for $x \in X_0$, $f_x \equiv x$ is a 0-ary function symbol and therefore a constant).

If $(X_i, \leqslant_i)$ $(i = 0, \ldots, n)$ are q.o.'s, then define a relation $\leqslant_{\mathrm{Term}}$ on $\mathrm{Term}(X_0, \ldots, X_n)$ inductively as follows:

$$
\begin{aligned}
&t_j \leqslant_{\mathrm{Term}} f_x(t_1, \ldots, t_i) \text{ for all } j \leqslant i \text{ and all}\\
&\qquad\qquad\qquad\qquad f_x(t_1, \ldots, t_i) \in \mathrm{Term}(X_0, \ldots, X_n);\\
&x \leqslant_i y \quad \& \quad t_1 \leqslant_{\mathrm{Term}} t_1' \quad \& \quad \ldots \quad \& \quad t_i \leqslant_{\mathrm{Term}} t_i'\\
&\qquad \Rightarrow f_X(t_1, \ldots, t_i) \leqslant_{\mathrm{Term}} f_y(t_1', \ldots, t_i')\\
&\qquad\qquad\qquad\qquad \text{for all } i \leqslant n, \text{ all } x, y \in X_i \text{ and all}\\
&\qquad\qquad\qquad\qquad t_1, \ldots, t_i, t_1', \ldots, t_i{}' \in \mathrm{Term}(X_0, \ldots, X_n);\\
&x \leqslant_{\mathrm{Term}} y \quad \& \quad y \leqslant_{\mathrm{Term}} z \Rightarrow x \leqslant_{\mathrm{Term}} z.
\end{aligned}
$$

Higman [11] proved the following result: If $(X_i, \leqslant_i)$ is a w.q.o. for each $i \leqslant n$, then so is $(\mathrm{Term}(X_0, \ldots, X_n), \leqslant_{\mathrm{Term}})$. We show that $(\mathrm{Term}(X_0, \ldots, X_n), \leqslant_{\mathrm{Term}})$ is isomorphic to $(\mathrm{T}(X_0, \ldots, X_n), \leqslant_T)$ (this is well-known but we provide a proof for completeness' sake), so that Theorem 3.2 yields an upper bound (the best possible one, as we shall see below) for the maximal order type of $(\mathrm{Term}\,(X_0, \ldots, X_n), \leqslant_{\mathrm{Term}})$.

**Definition 4.2** We define a 1–1 map $F$: $\text{Term}(X_0, \ldots, X_n) \to T(X_0, \ldots, X_n)$ inductively as follows:

$F(x) = .x$ for each $x \in X_0$;

$F(f_x(t_1, \ldots, t_i)) = x(F(t_1), \ldots, F(t_i))$.

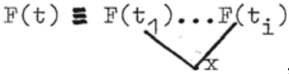
Thus $F$ maps $f_x(t_1, \ldots, t_i)$ into


**Lemma 4.3** *For any $s, t \in Term(X_0, \ldots, X_n)$,*
*$s \leqslant_{Term} t \iff F(s) \leqslant_T F(t)$ if the $X_i$ are pairwise disjoint.*

*Proof* of '⇒' by induction on the length of the proof that s $\leqslant_{\text{Term}} t$.

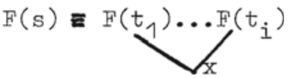
*Case 1*: s $\equiv t_j$ for some $j \leqslant i$ and $t \equiv f_x(t_1, \ldots, t_i)$.

Then

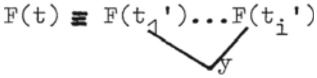



, and hence $F(s) \equiv F(t_j) \leqslant_T F(t)$.

*Case 2*: s $\equiv f_x(t_1, \ldots, t_i)$, $t \equiv f_y(t'_1, \ldots, t'_i)$, where $x \leq_i y$ and $t_j \leqslant_{\text{Term}} t'_j$ for all $j \leqslant i$ and, for each $j \leqslant i$, the proof of $t_j \leqslant_{Term} t'_j$ is shorter than that of s $\leqslant_{Term}$ t. Then, by induction hypothesis, $F(t_j) \leqslant_{Term} F(t'_j)$ for all $j \leqslant i$. Now




and



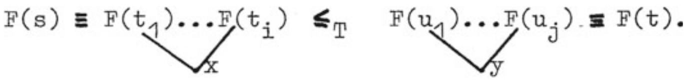

.

Hence $F(s) \leqslant_T F(t)$.

*Case 3*: There is some $u$ such that the proofs of s $\leqslant_{\text{Term}} u$ and $u \leqslant_{\text{Term}} t$ are shorter than that of s $\leqslant_{\text{Term}} t$. Then $F(s) \leqslant_T F(u)$ and $F(u) \leqslant_T F(t)$ by induction hypothesis; hence $F(s) \leqslant_T F(t)$.

Proof of '⇐ ' by induction on the total number of vertices in s and $t$:

Suppose s $\equiv f_x(t_1, \ldots, t_i)$, $t \equiv f_y(u_1, \ldots, u_j)$, where $i, j$ may be 0 (if $x$ resp. $y \in X_0$). Then




Let $g$ be the homeomorphic mapping from $F(s)$ to $F(t)$.

*Case 1* : $g$ maps the root of $F(s)$ into the root of $F(t)$.

Then $i = j, x, y \in X_i$ and $x \leqslant_i y$, and $F(t_k) \leqslant_T F(u_k)$ for all $k \leqslant i$. Hence, by induction hypothesis, $t_k \leqslant_{\text{Term}} u_k$ for each $k \leqslant i$. Hence s $\leqslant_{\text{Term}} t$.

*Case 2*: $g$ maps the root of $F(s)$ into a vertex $V$ of $F(t)$ which is not the root. Then, if $\tau$ is the subtree of $F(t)$ consisting of $V$ and all its successors, $F(s) \leqslant_T \tau$. But $\tau = F(t_0)$, where $t_0$ is a subterm of $t$. Hence, by induction hypothesis, s $\leqslant_{\text{Term}} t_0$. Hence s $\leqslant_{\text{Term}} t$.

**Corollary 4.4**

$$o(\mathrm{Term}(X_0, \ldots, X_n), \leqslant_{\mathrm{Term}}) \leqslant f^+ \begin{pmatrix} -1 + o(X_0) & \cdots & o(X_n) \\ 0 & \cdots & n \end{pmatrix}.$$

*Proof* Follows immediately from Lemma 4.3 and Theorem 3.2.

**Definition 4.5** Let $\phi$ be an $n$-ary function on the ordinals, $\phi : \mathrm{On}^n \to \mathrm{On}$, where On denotes the class of all ordinals. $\phi$ is *monotonic* iff for all ordinals $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ $\alpha_1 \leqslant \beta_1, \ldots, \alpha_n \leqslant \beta_n \Rightarrow \phi\alpha_1 \cdots \alpha_n \leqslant \phi\beta_1 \cdots \beta_n$. $\phi$ is *increasing* iff for all ordinals $\alpha_1, \ldots, \alpha_n$ and each $i \leqslant n, \alpha_i \leqslant \phi\alpha_1 \cdots \alpha_n$.

If $\Phi$ is a set of ordinal functions, CL($\Phi$), the *closure set of* $\Phi$, is the smallest set of ordinals containing 0 and closed under all elements of $\Phi$. cl($\Phi$) the *closure ordinal of* $\Phi$, is the order type of CL($\Phi$) under the usual well-ordering on the ordinals.

If $\phi, \psi$ are any two $n$-ary ordinal functions, define $\phi \leqslant_{fn} \psi \iff \phi\alpha_1 \cdots \alpha_n \leqslant \psi\alpha_1 \cdots \alpha_n$ for all ordinals $\alpha_1, \ldots, \alpha_n$.

**Theorem 4.6** *Let* $\Phi = X_0 \cup \cdots \cup X_n$ *be a set of ordinal functions, where for each* $i$ $X_i$ *contains only* $i$*-ary functions and* $(X_i, \leqslant_{fn})$ *is a w.q.o. (Then* $X_0$ *is a set of ordinals.) If we define* $\leqslant_i \equiv \leqslant_{fn} \restriction X_i$ *for each* $i$ *and* $X'_0 = \{0\} \dot{\bigcup} X_0$, $\leqslant'_0 = \leqslant_0 \cup \{0\} \times X'_0$ *(i.e. 0 becomes the least element of* $X'_0$ *under the w.q.o.* $\leqslant'_0$*), then*

$$\mathrm{cl}(\Phi) \leqslant o(\mathrm{Term}(X'_0, X_1, \ldots, X_n), \leqslant_{Term}).$$

*Proof* We define a map $G$: Term$(X'_0, X_1, \ldots, X_n) \to$ CL($\Phi$) in the obvious way:
$G(0) = 0$;

$G(x) = x$ for each $x \in X_0$;

$$G(f_x(t_1, \ldots, t_i)) = f_x(G(t_1), \ldots, G(t_i)) \text{ for all } x \in X_i, \quad t_1, \ldots, t_i \in \mathrm{Term}(X'_0, X_1, \ldots, X_n).$$

It is easy to see (by transfinite induction on $\alpha$) that for every $\alpha \in$ CL($\Phi$) there is a $t \in$ Term$(X'_0, X_1, \ldots, X_n)$ such that $\alpha = G(t)$ — i.e. $G$ is surjective; and also that $G$ is order-preserving, i.e. s $\leqslant_{\mathrm{Term}}$ $t \Rightarrow G(s) \leqslant G(t)$; however, $G$ is in general not 1–1 – there are $\leqslant_{\mathrm{Term}}$ -incomparable s and $t$ such that $G(s) = G(t)$ and, in general, s $<_{\mathrm{Term}}$ $t$ such that $G(s) = G(t)$. So we adjust $G$ as follows:

Let Term$(X'_0, X_1, \ldots, X_n)$ be well-ordered in some arbitrary way compatible with $\leqslant_{\mathrm{Term}}$ (e.g. see Sect. 2.1.4). If $\beta$ is the order type of this well-ordering, let

$\| \ \|$: Term$(X'_0, X_1, \ldots, X_n) \to \beta$ associate with each term t its position in the well-ordering.

Now we can define $G'$ : Term$(X'_0, X_1, \ldots, X_n) \to \beta. \sup\{\alpha | \alpha \in CL(\Phi)\}$ as follows:

$$G'(t) = \beta.G(t) + \| t \| .$$

Now $G'$ is 1–1 strictly order-preserving and, since $G$ was surjective, the image of $G'$

has order type $\geqslant$ cl($\Phi$). Hence, if we define a relation $\leqslant^+$ on Term$(X_0', X_1, \ldots, X_n)$ by:

$$s \leqslant^+ t \Leftrightarrow G'(s) \leqslant G'(t),$$

then $\leqslant^+$ is a well-ordering of order type at least cl($\Phi$) which extends $\leqslant_{\text{Term}}$. Hence

$$\text{cl}(\Phi) \leqslant \text{o}(\text{Term}(X_0', X_1, \ldots, X_n), \leqslant_{\text{Term}}).$$

**Corollary 4.7** *If $\Phi = X_0 \cup \ldots \cup X_n$ is a set of monotonic increasing ordinal functions, where for each i $X_i$ contains only i-ary functions and $(X_i, \leqslant_{fn})$ is a w.q.o., then*

$$cl(\Phi) \leqslant f^+ \begin{pmatrix} o(X_0, \leqslant_{fn}) & \ldots & o(X_n, \leqslant_{fn}) \\ 0 & \ldots & n \end{pmatrix}.$$

*Proof* Follows from Corollary 4.4 and Theorem 4.6, since $-1 + o(X_0') = -1 + (1 + o(X_0)) = o(X_0)$.

This result was proved for $X_0, \ldots, X_{n-1}$ empty ($n \geqslant 3$) and $X_n$ a singleton in [22], and can be deduced from the proof in [22] if the $X_i$ are finite or well-ordered by $\leqslant_{fn}$ and $\binom{1}{3} \leqq \begin{pmatrix} o(X_0) \ldots o(X_n) \\ 0 \ldots n \end{pmatrix}$. As far as I can see, however, the corresponding results about trees and terms (Corollary 4.4 and Theorem 2.2) cannot be inferred easily from the above result, since, for example, if $f_x(t), f_x(u) \in \text{Term}(X_0, \ldots, X_n)$ and $\leqslant^+$ is an extension of $\leqslant_{\text{Term}}$ to a well-ordering of Term $(X_0, \ldots, Xn)$, then

$$t \leqslant_{\text{Term}} u \to f_x(t) \leqslant^+ f_x(u) \text{ is true but}$$
$$t \leqslant^+ u \to f_x(t) \leqslant^+ f_x(u) \text{ is not in general } (t \text{ and } u \text{ may be } \leqslant_{\text{Term}} \text{-incomparable}),$$

and therefore the well-ordering induced on Term$(X_0, \ldots, X_n)$ by the functions in $X_0 \bigcup \ldots \bigcup X_n$ is not in general an extension of $\leqslant_{\text{Term}}$.

Corollaries 4.4 and 4.7 have already been proved by de Jongh (as yet unpublished) for $\begin{pmatrix} o(X_0) & \ldots & o(X_n) \\ 0 & \ldots & n \end{pmatrix} \leqq \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and in [10] for $\begin{pmatrix} o(X_0) & \ldots & o(X_n) \\ 0 & \ldots & n \end{pmatrix} \leqq \begin{pmatrix} 2 \\ 2 \end{pmatrix}$.

In the next two theorems we show that the bounds of Corollary 4.7 are best possible provided $\begin{pmatrix} o(X_0) & \ldots & o(X_n) \\ 0 & \ldots & n \end{pmatrix} \leqq \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

**Theorem 4.8** *If $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \geqq \underline{\alpha} = \begin{pmatrix} \alpha_0 & \ldots & \alpha_n \\ 0 & \ldots & n \end{pmatrix}$, then there are sets $X_0^{\underline{\alpha}}, \ldots, X_n^{\underline{\alpha}}$ such that*

*(i) each $X_i^{\underline{\alpha}}$ is a set of monotonic increasing i-ary functions well-ordered by $\leqslant_{fn}$ with order type $\alpha_i$;*

*(ii) $CL(\Phi^{\underline{\alpha}}) \supseteq f \begin{pmatrix} \alpha_0 & \ldots & \alpha_n \\ 0 & \ldots & n \end{pmatrix}$, where $\Phi^{\underline{\alpha}} = X_0^{\underline{\alpha}} \bigcup \ldots \bigcup X_n^{\underline{\alpha}}$.*

*Proof* We assume w.l.o.g. that $\alpha_n \neq 0$ and $n \geqslant 2$. We define the $X_i^{\underline{\alpha}}$ as follows:

$$X_0^{\underline{\alpha}} = \{f\begin{pmatrix} \beta\, \alpha_1 & \ldots & \alpha_n \\ 0\, 1 & \ldots & n \end{pmatrix} | \beta < \alpha_0\};$$

$$X_i^{\underline{\alpha}} = \{\lambda\beta_0 \ldots \beta_{i-1} f\begin{pmatrix} \beta_0 & \ldots & \beta_{i-1} & \beta & \alpha_{i+1} & \ldots & \alpha_n \\ 0 & \ldots & i-1 & i & i+1 & \ldots & n \end{pmatrix} | \beta < \alpha_i\}$$

$$\text{for } 0 < i < n;$$

$$X_n^{\underline{\alpha}} = \{\lambda\beta_0 \ldots \beta_{n-1} f\begin{pmatrix} \beta_0 & \ldots & \beta_{n-1} & \beta \\ 0 & \ldots & n-1 & n \end{pmatrix} | 0 < \beta < \alpha_n\} \bigcup \{f_n\},$$

where $f_n$ is an $n$-ary ordinal function defined as follows:

$$f_n\beta_0 \ldots \beta_{n-1} = \begin{cases} \omega^{\beta_1} + \beta_0 \text{ if } \beta_2 = \ldots = \beta_{n-1} = 0, \\ f\begin{pmatrix} \beta_0 & \ldots & \beta n-1 & 0 \\ 0 & \ldots & n-1 & n \end{pmatrix} \text{ otherwise.} \end{cases}$$

We define $\Phi^{\underline{\alpha}} = X_0^{\underline{\alpha}} \bigcup \ldots \bigcup X_n^{\underline{\alpha}}$.

The $X_i^{\underline{\alpha}}$ satisfy (i) by Theorem 2.16. In order to prove (ii), we shall prove the following assertion by transfinite induction on $\gamma$ with subsidiary transfinite induction on $\underline{\alpha}$ with respect to $\prec$:

$$\text{If } \gamma < f\begin{pmatrix} \alpha_0 & \ldots & \alpha_n \\ 0 & \ldots & n \end{pmatrix} \text{ then } \gamma \in \text{CL}(\Phi^{\underline{\alpha}}).$$

Now, since $\gamma < f\underline{\alpha}$, by 2.16 d),e) either $\gamma < f\begin{pmatrix} 0 \ldots 0 \\ 0 \ldots n \end{pmatrix}$ or there is a $j \leqslant n$ and a $\beta_j < \alpha_j$ such that

$$f\begin{pmatrix} \beta_j & \alpha_{j+1} & \ldots & \alpha_n \\ j & j+1 & \ldots & n \end{pmatrix} \leqslant \gamma < f\begin{pmatrix} \beta_j + 1 & \alpha_{j+1} & \ldots & \alpha_n \\ j & j+1 & \ldots & n \end{pmatrix}.$$

*Case 1* : $\gamma$ is not an $\varepsilon$-number, i.e. $\gamma \neq \omega^\gamma$. Then there are $\gamma_0, \gamma_1 < \gamma$ such that $\gamma = \omega^{\gamma_1} + \gamma_0 = f_n\gamma_0\gamma_1 0 \ldots 0$. But $0 \in \text{CL}(\Phi^{\underline{\alpha}})$ by definition of $\text{CL}(\Phi^{\underline{\alpha}})$ and $\gamma_0, \gamma_1 \in \text{CL}(\Phi^{\underline{\alpha}})$ by (main) induction hypothesis. Hence $\gamma \in \text{CL}(\Phi^{\underline{\alpha}})$. Note that if $\gamma < f\begin{pmatrix} 0 \ldots 0 \\ 0 \ldots n \end{pmatrix} = \omega$ then Case 1 holds.

*Case 2* : There is a $j > 0$ and a $\beta_j < \alpha_j$ such that

$$f\begin{pmatrix} \beta_j & \alpha_{j+1} & \ldots & \alpha_n \\ j & j+1 & \ldots & n \end{pmatrix} \leqslant \gamma < f\begin{pmatrix} \beta_j & \alpha_{j+1} & \ldots & \alpha_n \\ j & j+1 & \ldots & n \end{pmatrix}.$$

Then $f\begin{pmatrix} \beta_j & \alpha_{j+1} & \ldots & \alpha_n \\ j & j+1 & \ldots & n \end{pmatrix}$ is the least solution $\eta$ of the equation

$$f\begin{pmatrix} \eta & \beta_j & \alpha_{j+1} & \dots & \alpha_n \\ j-1 & j & j+1 & \dots & n \end{pmatrix} = \eta.$$

Hence, by 2.16 d),e) there is a $\gamma_0$ such that

$$\gamma_0 < \gamma \leqslant f\begin{pmatrix} \gamma_0 & \beta_j & \alpha_{j+1} & \dots & \alpha_n \\ j-1 & j & j+1 & \dots & n \end{pmatrix}.$$

We shall write $\underline{\beta} = \begin{pmatrix} \gamma_0 & \beta_j & \alpha_{j+1} & \dots & \alpha_n \\ j-1 & j & j+1 & \dots & n \end{pmatrix}$.

Now, since $n \geq 2$ and $\alpha_n \neq 0$, $\underline{\beta} \prec \binom{1}{2}$ holds only if $j = n = 2$ and $\beta_j = 0$. But then $\gamma_0 < \gamma \leqslant \omega^{\omega^{\gamma_0}}$; hence case 1 above holds. Thus we may assume that $\binom{1}{2} \leqq \underline{\beta}$.

Now $\gamma_0 \in \mathrm{CL}(\Phi^{\underline{\alpha}})$ by (main) induction hypothesis. Hence if $\gamma = f\underline{\beta}$ then $\gamma \in \mathrm{CL}(\Phi^{\underline{\alpha}})$ thus we may assume that $\gamma_0 < \gamma < f\underline{\beta}$. But then, by (subsidiary) induction hypothesis, $\gamma \in \mathrm{CL}(\Phi^{\underline{\beta}})$.

Now $X_i^{\underline{\beta}} \subseteq X_i^{\underline{\alpha}}$ for all $i \neq j-1$.

Moreover, by (subsidiary) induction hypothesis, $\gamma_0 \subseteq \mathrm{CL}(\Phi^{\underline{\alpha}})$. Hence if $\phi \in X_{j-1}^{\underline{\beta}}$ then there is a $\delta_{j-1} < \gamma_0$ such that

$$\phi = \lambda\eta_0 \dots \eta_{j-2} f\begin{pmatrix} \eta_0 & \dots & \eta_{j-2} & \delta_{j-1} & \beta_j & \alpha_{j+1} & \dots & \alpha_n \\ 0 & \dots & j-2 & j-1 & j & j+1 & \dots & n \end{pmatrix},$$

and if we define

$$\psi = \lambda\eta_0 \dots \eta_{j-1} f\begin{pmatrix} \eta_0 & \dots & \eta_{j-1} & \beta_j & \alpha_{j+1} & \dots & \alpha_n \\ 0 & \dots & j-1 & j & j+1 & \dots & n \end{pmatrix}$$

then $\psi \in X_j^{\underline{\alpha}}$ and

$$\phi\eta_0 \dots \eta_{j-2} = \psi\eta_0 \dots \eta_{j-2}\delta_{j-1} \text{ for all } \eta_0, \dots \eta_{j-2}.$$

Hence, since $\delta_{j-1} \in \mathrm{CL}(\Phi^{\underline{\alpha}})$, $\mathrm{CL}(\Phi^{\underline{\beta}}) \subseteq \mathrm{CL}(\Phi^{\underline{\alpha}})$ and hence $\gamma \in \mathrm{CL}(\Phi^{\underline{\alpha}})$.

*Case 3*: There is a $\beta_0 < \alpha_0$ such that

$$f\begin{pmatrix} \beta_0 & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix} \leqslant \gamma < f\begin{pmatrix} \beta_0 + 1 & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix}.$$

If equality holds then clearly $\gamma \in \mathrm{CL}(\Phi^{\underline{\alpha}})$; hence we may assume that $f\begin{pmatrix} \beta_0 & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix} < \gamma < f\begin{pmatrix} \beta_0 + 1 & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix}$.

Let $j \leqslant n$ be the least nonzero integer such that $\alpha_j > 0$. Then $\gamma$ lies between two consecutive common solutions of all the equations

$$f\begin{pmatrix} \eta & \beta_j & \alpha_{j+1} & \dots & \alpha_n \\ j-1 & j & j+1 & \dots & n \end{pmatrix} = \eta \qquad (\beta_j < \alpha_j).$$

Hence, by 2.16 d),e), there is a $\beta_j < \alpha_j$ and a $\gamma_0$ such that

$$\gamma_0 < \gamma < f\begin{pmatrix} \gamma_0 & \beta_j & \alpha_{j+1} & \dots & \alpha_n \\ j-1 & j & j+1 & \dots & n \end{pmatrix}.$$

The proof now proceeds as in Case 2.

**Theorem 4.9** *If* $\binom{1}{2} \leqq \underline{\alpha} = \begin{pmatrix} \alpha_0 & \dots & \alpha_n \\ 0 & \dots & n \end{pmatrix}$*, then there are sets* $Y_0^{\underline{\alpha}}, \dots, Y_n^{\underline{\alpha}}$ *such that*

*(i)* *each* $Y_i^{\underline{\alpha}}$ *is a set of monotonic increasing i-ary functions well-ordered by* $\leqslant_{fn}$ *with order type* $\alpha_i$*;*

*(ii)* $\mathrm{CL}(\Phi^{\underline{\alpha}}) \supseteq f^+\begin{pmatrix} \alpha_0 & \dots & \alpha_n \\ 0 & \dots & n \end{pmatrix}$ *where* $\Phi^{\underline{\alpha}} = X_0^{\underline{\alpha}} \cup \dots \cup X_n^{\underline{\alpha}}$.

*Proof* The assertion follows from Theorem 4.8 if $f^+\underline{\alpha} = f\underline{\alpha}$. Hence we may assume that this is not the case, i.e. that there is an $\alpha \leqslant \alpha_0$ and an $n < \omega$ such that $\alpha_0 = \alpha + n$,

$$f\begin{pmatrix} \alpha & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix} = \alpha \text{ or } f\begin{pmatrix} \alpha & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix} = \alpha_j \text{ for some } j > 0$$

and

$$f^+\begin{pmatrix} \alpha_0 & \dots & \alpha_n \\ 0 & \dots & n \end{pmatrix} = f\begin{pmatrix} \alpha_0 + 1 & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix}$$

We shall assume that $f\begin{pmatrix} \alpha & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix} = \alpha$ and indicate at the end of this proof how to adapt the proof if

$$f\begin{pmatrix} \alpha & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix} = \alpha_j \text{ for some } j > 0.$$

We again assume w.l.o.g. that $\alpha_n \neq 0$ and $n \geqslant 2$, and define the $Y_i^{\underline{\alpha}}$ as follows:

$$Y_0^{\underline{\alpha}} = \alpha \cup \left\{ f\begin{pmatrix} \alpha + m & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix} | 0 < m \leqslant n \right\};$$

$Y_i^{\underline{\alpha}} = X_i^{\underline{\alpha}}$ (see proof of Theorem 4.8) for $0 < i < n$; and
$Y_n^{\underline{\alpha}} \{\phi^+ | \phi \in X_n^{\underline{\alpha}}\}$, where $\phi^+$ is defined by $\phi^+\beta_0 \dots \beta_{n-1} = \max\{\alpha, \phi\beta_0 \dots \beta_{n-1}\}$.

The $Y_i^{\underline{\alpha}}$ satisfy (i) by Theorem 2.16. In order to prove (ii), we shall prove the following assertion by transfinite induction on $\gamma$ with subsidiary transfinite induction on $\underline{\alpha}$ with respect to $\prec$:

If $\gamma < f^+\underline{\alpha}$ then $\gamma \in \mathrm{CL}(\Phi^{\underline{\alpha}})$.

*Case 1* : $\gamma < \alpha = f\begin{pmatrix} \alpha & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix}$. Then $\gamma \in Y_0^{\underline{\alpha}} \subseteq \mathrm{CL}(\Phi^{\underline{\alpha}})$.

*Case 2*: $\gamma \geqslant \alpha$ and $\gamma$ is not an $\varepsilon$-number. Then, since $\alpha$ is an $\varepsilon$-number, $\gamma > \alpha$ and there are $\gamma_0, \gamma_1 < \gamma$ such that $\alpha \leqslant \gamma_0, \gamma_1$ and $\gamma = \omega^{\gamma_1} + \gamma_0 = f_n^+ \gamma_0 \gamma_1 0 \dots 0$. But $0 \in \mathrm{CL}(\Phi^{\underline{\alpha}})$ by definition of $\mathrm{CL}(\Phi^{\underline{\alpha}})$ and $\gamma_0, \gamma_1 \in \mathrm{CL}(\Phi^{\underline{\alpha}})$ by induction hypothesis. Hence $\gamma \in \mathrm{CL}(\Phi^{\underline{\alpha}})$.

*Case 3*: Cases 1 and 2 do not hold. Then $f\begin{pmatrix} \alpha & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix} \leqslant \gamma < f\begin{pmatrix} \alpha+n+1 & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix}$; hence there is some $m \leqslant n$ such that $f\begin{pmatrix} \alpha+m & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix} \leqslant \gamma$ $< f\begin{pmatrix} \alpha+m+1 & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix}$.

Now if strict inequality holds then the assertion follows as in the proof of Case 3 of Theorem 4.8. Hence we may assume that $\gamma = f\begin{pmatrix} \alpha+m & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix}$.

But then if $m > 0 \quad \gamma \in Y_0^{\underline{\alpha}} \in \mathrm{CL}(\Phi^{\underline{\alpha}})$,
and if $m = 0$ $\gamma = f_n^+ 0 \dots 0 \in \mathrm{CL}(\Phi^{\underline{\alpha}})$, q.e.d.

We now indicate how to adapt the above proof if $f\begin{pmatrix} \alpha & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix} = \alpha_j$ for some $j > 0$.

Then, by Theorem 2.16b), $\alpha = 0, \alpha_0 = n$ and $\alpha_1, \dots, \alpha_{j-1} = 0$. We then define

$$Y_0^{\underline{\alpha}} = \left\{ f\begin{pmatrix} m & \alpha_1 & \dots & \alpha_n \\ 0 & 1 & \dots & n \end{pmatrix} | 0 < m \leqslant n \right\},$$

$Y_1^{\underline{\alpha}}, \dots, Y_{j-1}^{\underline{\alpha}}$ to be empty, $Y_j^{\underline{\alpha}} = \{\phi_\beta | \beta < \alpha_j\}$, where

$$\phi_\beta \beta_0 \dots \beta_{j-1} = \begin{cases} \beta \text{ if } \beta_0 = \dots = \beta_{j-1} = 0, \\ \max\left\{\alpha_j, f\begin{pmatrix} \beta_0 \dots & \beta_{j-1} & \beta & \alpha_{j+1} & \dots & \alpha_n \\ 0 \dots & j-1 & j & j+1 & \dots & n \end{pmatrix}\right\} & \text{otherwise,} \end{cases}$$

$Y_i^{\underline{\alpha}} = X_i^{\underline{\alpha}}$ for $j < i \leqslant n$.

(If $j = n Y_j^{\underline{\alpha}}$ must be adjusted a little so as to contain $f_n$. Then the proof goes through as before, except that now $\alpha_j \subseteq CL(\Phi^{\underline{\alpha}})$ by virtue of the definition of $Y_j^{\underline{\alpha}}$ and not of $Y_0^{\underline{\alpha}}$; and $\alpha_j \in \mathrm{CL}(\Phi^{\underline{\alpha}})$ by virtue of the definition of $Y_j^{\underline{\alpha}}$, and not of $Y_n^{\underline{\alpha}}$.

## *4.2 X-Monotonic Increasing Ordinal Functions*

In this section we shall use the results of Sect. 3.2 to give bounds for the order types of the sets of ordinals generated by one binary function which is increasing but not quite monotonic.

**Definition 4.10** Let $X$ be any set. We define Term$_2(X)$, the set of terms generated from $X$ by a binary function symbol, inductively by:

If $x \in X$ than $x \in$ Term$_2(X)$; if $s, t \in$ Term$_2(X)$ then (s, t) $\in$ Term$_2$(X).

Thus Term$_2(X)$ is isomorphic to Term$(X, \phi, \{0\})$.

Note that each element of Term$_2(X)$ is either an element of $X$ or has the form $((\ldots(t_0, t_1)\ldots, t_{n-1}), t_n)$, where $t_0 \in X$ and $t_1, \ldots, t_n \in$ Term$_2(X)$.

With each $t \in$ Term$_2(X)$ we associate an element $h(t)$ of $X$ as follows:

If $t \in X$ then $h(t) = t$; if $t = (u, v)$ then $h(t) = h(u)$.

Thus, if $t = ((\ldots(t_0, t_1)\ldots, t_{n-1}), t_n)$, where $t_0 \in X$, then $h(t) = t_0$.

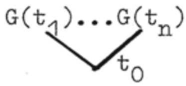
We now define a 1–1 map $G$: Term$_2(X) \to T\binom{X}{\omega}$

(T$\binom{X}{\omega}$ is the set of all finite structured trees with labels in $X$) as follows:

$$G(x) = .x \text{ for } x \in X;$$

$$G(((\ldots(t_0, t_1)\ldots, t_{n-1}), t_n)) = t_0(G(t_1), \ldots, G(t_n)) \text{for } t_0 \in X$$

Thus $G$ maps $((\ldots(t_0, t_1)\ldots, t_{n-1}), t_n)$ into




.

If $(X, \leqslant)$ is a q.o., we define a q.o. $\leqslant^2$ on Term$_2(X)$ inductively by:

(a) $x \leqslant y \Rightarrow x \leqslant^2 y$ for all $x, y \in X$;
(b) s $\leqslant^2 (s, t)$ and $t \leqslant^2 (s, t)$ for all $s, t \in$ Term$_2(X)$;
(c) $s_0 \leqslant t_0$ & $s_1 \leqslant^2 t_{i_1}$ & $\ldots$ & $s_n \leqslant^2 t_{i_n}$ & $1 \leqslant i_1 < \ldots < i_n \leqslant m$

$$\Rightarrow (\ldots(s_0, s_1)\ldots, s_n) \leqslant^2 (\ldots(t_0, t_1)\ldots, t_m)$$

for all $s_1, \ldots, s_n, t_1, \ldots, t_m \in$ Term$_2(X)$ and all $s_0, t_0 \in X$

(d) s $\leqslant^2 t$ & $t \leqslant^2 u \Rightarrow s \leqslant^2 u$ for all $s, t, u \in$ Term$_2(X)$.

Clause (c) of the definition of $\leqslant^2$ is a little unwieldly; note that if we replaced it by the simpler clause

(c′) $s_1 \leqslant^{2+} s_2$ & $t_1 \leqslant^{2+} t_2$ & $h(s_1) \leqslant h(s_2)$

$\Rightarrow (s_1, t_1) \leqslant^2 (s_2, t_2)$ for all $s_1, s_2, t_1, t_2 \in$ Term$_2(X)$,

then we should obtain an extension $\leqslant^{2+}$ of $\leqslant^2$. Since, by Theorem 4.11 below, $\leqslant^2$ is a w.q.o., $\leqslant^{2+}$ is also a w.q.o. and $o$(Term$_2(X), \leqslant^{2+}) \leqslant o$(Term$_2(X), \leqslant^2$). Thus Coro1lary 4.12 below also applies to (Term$_2(X), \leqslant^{2+}$).

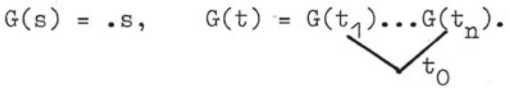
**Theorem 4.11** *$s \leqslant^2 t \Leftrightarrow G(s) \leqslant_T G(t)$ for all $s, t \in$ Term$_2(X)$.*

*Proof* We shall omit the proof of the implication from left to right since we do not need it in the following. We shall prove $G(s) \leqslant_T G(t) \Rightarrow s \leqslant^2 t$ by induction on the sum of the lengths of s and $t$.

*Case 1*: $s, t \in X$. Then $G(s) \leqslant_T G(t) \Leftrightarrow s \leqslant t \Leftrightarrow s \leqslant^2 t$.

*Case 2*: Just one of $s, t$ is an element of $X$. Then since $G(s) \leqslant_T G(t)$ we must have s $\in X$ and

$$t = (\ldots(t_0, t_1)\ldots, t_n), \text{ where } t_0 \in X;$$

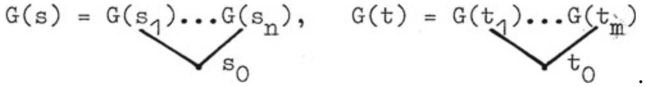



Thus s $\leqslant r$, where $r$ is the label of some node in $G(t)$.

Hence $r$ occurs somewhere in $t$. But then s $\leqslant^2 t$ by finitely many applications of clauses (b) and (d) of the definition of $\leqslant^2$.

*Case 3*: $s, t \notin X$. Then we may suppose that s $= (\ldots(s_0, s_1)\ldots, s_n)$ and $t = (\ldots(t_0, t_1)\ldots, t_m)$ for some $m, n > 0$, where $s_0, t_0 \in X$.

Then




Let $f$ be the homeomorphic mapping of $G(s)$ into $G(t)$.

If $f$ does not map the root of $G(s)$ into the root of $G(t)$ then $G(s) \leqslant_T G(t_i)$ for some $i \in \{1, \ldots, m\}$. But then s $\leqslant^2 t_i$ by induction hypothesis, and hence s $\leqslant^2 t$ by clauses (b) and (d) of the definition of $\leqslant^2$.

If on the other hand $f$ maps the root of $G(s)$ into the root of $G(t)$ then $s_0 \leqslant t_0$ and $f$ must map the roots of $G(s_1), \ldots, G(s_n)$ into some vertices of $G(t_{i_1}), \ldots, G(t_{i_n})$, where $1 \leqslant i_1 < i_2 < \ldots < i_n \leqslant m$. Thus $G(s_j) \leqslant_T G(t_{i_j})$ for each $j = 1, \ldots, n$. Hence $s_j \leqslant^2 t_{i_j}$ for each $j$ by induction hypothesis; hence s $\leqslant^2 t$ by clause (c) of the definition of $\leqslant^2$.

**Corollary 4.12** *If* $(X, \leqslant)$ *is a w.q.o., then so is* $(Term_2(X), \leqslant^2)$ *and* $o(Term_2(X), \leqslant^2) \leqslant f\binom{o(X)}{\omega}$.

*Proof* Follows from Theorem 3.4, since $(\text{Term}_2(X), \leqslant^2)$ is isomorphic to $(T\binom{X}{\omega}, \leqslant_T)$ by Theorem 4.11. Actually, it suffices if $(\text{Term}_2(X), \leqslant^2)$ is isomorphic to an extension of $\leqslant_T$ on $T\binom{X}{\omega}$; i.e. the implication from left to right in Theorem 4.11 suffices.

**Definition 4.13** Let $X$ be a set of ordinals. A binary ordinal function $\phi$ is said to be *X-monotonic* if and only if $\phi(\cdots\phi(\alpha_0, \alpha_1), \cdots, \alpha_n) \leqslant \phi(\cdots\phi(\beta_0, \beta_1), \cdots, \beta_m)$ whenever $m, n > 0$, $\alpha_0, \beta_0 \in X$, $\alpha_1, \cdots, \alpha_n, \beta_1, \cdots, \beta_m$ are any ordinals and there are $i_1, \cdots, i_n$ such that $1 \leqslant i_1 < \cdots i_n \leqslant m$ and $\alpha_j \leqslant \beta_{i_j}$ for each $j = 1, \cdots, n$.

**Corollary 4.14** *If $X$ is a set of ordinals and $\phi$ an $X$-monotonic increasing binary ordinal function, then the closure of $X \cup \{0\}$ under $\phi$ has order type at most $f\binom{||X||}{\omega}$, where $||X||$ denotes the order type of $X$.*

*Proof* Analogous to the proof of Theorem 2.6. Denote the closure of $X \cup \{0\}$ under $\phi$ by $CL_\phi(X)$ and the order type of $CL_\phi(X)$ by $cl_\phi(X)$, and define

$G : \mathrm{Term}_2(X) \to CL_\phi(X)$ by :

$G(0) = 0;$

$G(x) = x$ for all $x \in X$;

$G((s, t)) = \phi(G(s), G(t))$ for all $s, t \in \mathrm{Term}_2(X)$.

Then s $\leqslant^2 t \Rightarrow G(s) \leqslant G(t)$ for all $s, t \in \mathrm{Term}_2(X)$ and, as in the proof of 2.6, we may define a 1–1 strictly order-preserving map $G'$ from $\mathrm{Term}_2(X)$ into some ordinal, which induces a well-ordering on $\mathrm{Term}_2(X)$ which extends $\leqslant^2$ and has order type at least $cl_\phi(X)$. But, by Corollary 4.12, this well-ordering has order type at most $f\binom{||X||}{\omega}$.

*Remark* We shall not go into the question whether the bounds of Corollary 4.14 are best possible, because we do not know of any X-monotonic functions which are not monotonic and occur in mathematical practice; i.e. we do not know whether Corollary 4.14 has any useful applications. However, Corollary 4.14 does show that the monotonicity condition of Corollary 4.7 can be weakened (at least if $n = 2$ and $X_n$ is a singleton) at the cost of an increase in the bounds of Corollary 4.7. Thus there is some hope of displaying upper bounds for the closure ordinals of the sort of (sets of) functions which are used in the notation systems of [25] and [1, 4, 6, 19, 26, 43] etc., in which either monotonicity or increasingness fails to hold.

## *4.3 Possible Connections with Proof Theory*

To date, systems of notations for countable ordinals have been studied primarily in connection with proof theory, in order to find the ‘proof-theoretic ordinals (or, better, representations of ordinals)’ of interesting subsystems of analysis, In order not to have to digress here, we refer the reader to Sect. 6 of [5], which defines the notion of the ‘proof-theoretic ordinal’ of a subsystem of analysis and summarizes all results available to date on the proof-theoretic ordinals of particular systems.

Now Theorems 3.2 and 4.9 of this paper show that each ordinal $f^+\begin{pmatrix}\alpha_0 & \cdots & \alpha_n \\ 0 & \cdots & n\end{pmatrix}$ ($n < \omega$; $\alpha_0, \cdots, \alpha_n$ any ordinals) can be characterised in a natural way as the maximal order type of the w.p.o. $(T(1+\alpha_0, \alpha_1, \cdots, \alpha_n), \leqslant_T)$, where the order relation associated with each $\alpha_i$ is the usual well-ordering on the ordinals restricted to $\alpha_i$. Alternatively, by Corollary 4.7 and Theorem 4.9, $f^+\begin{pmatrix}\alpha_0 & \cdots & \alpha_n \\ 0 & \cdots & n\end{pmatrix}$ is the greatest ordinal which can be obtained as the closure ordinal of a set $\Phi = X_0 \cup \cdots \cup X_n$, where for each $i$ $X_i$ contains only monotonic increasing i-ary functions and $(X_i, \leqslant_{fn})$ is a w.q.o. with maximal order type $\alpha_i$. It is therefore natural to ask which of the ordinals $f^+\begin{pmatrix}\alpha_0 & \cdots & \alpha_n \\ 0 & \cdots & n\end{pmatrix}$ turn up as the proof-theoretic ordinal of some subsystem of analysis and, for those which do, whether the above characterisations of the given ordinal can give new insights into the proof theory of the corresponding subsystem of analysis. The first question is easy to answer; the second is an unbroached[2] open problem.

As we mentioned at the end of Sect. 2, for countable or finite $\alpha_0, \cdots, \alpha_n$, $\beta_1, \cdots, \beta_n$ such that $\binom{1}{2} \leqq \begin{pmatrix}\alpha_0 & \alpha_1 \cdots \alpha_n \\ 0 & \beta_1 \cdots \beta_n\end{pmatrix}$, $f\begin{pmatrix}\alpha_0 & \alpha_1 \cdots \alpha_n \\ 0 & \beta_1 \cdots \beta_n\end{pmatrix}$ is approximately $\phi_{\Omega^{\beta_n}\cdot\alpha_n+\cdots+\Omega^{\beta_1}\cdot\alpha_1}(\alpha_0)$. Moreover, in [3, 4] it is shown that, for 'nice' $\alpha$, $\phi_\alpha(\beta) = \Theta_{\alpha(-1+\beta)}$, where $\Theta$ is the function of the Feferman-Aczel notation which is used to express most of the ordinals mentioned in Sect. 6 of [5]. Thus the reader may check that all these ordinals except for $\mathcal{E}_0$, $\mathcal{E}_{\mathcal{E}_0}$ and $\Gamma_0$ are fixed points in the Klammer-symbol notation—that is to say, they cannot be expressed in the form $fA$, where $A$ contains only smaller ordinals. But then they also cannot be expressed in the form $f^+A$, where $A$ contains only smaller ordinals. So the only ordinals $f^+A$ which turn up non-trivially as the proof-theoretic ordinal of some subsystem of analysis are:

(i) $\varepsilon_0 = f^+\binom{1}{2}$, the proof-theoretic ordinal of arithmetic;
(ii) $\varepsilon_0 = f^+\begin{pmatrix}\varepsilon_0 & 1 \\ 0 & 2\end{pmatrix}$, the proof-theoretic ordinal elementary analysis;
(iii) $\Gamma_0 = f^+\binom{2}{2}$, the proof-theoretic ordinal of predicative analysis.

So the second of the two questions raised above reduces to the following questions:

(a) Do the characterisations of $\varepsilon_0 = f^+\binom{1}{2}$ given in this paper give any new insights into the proof theory of arithmetic?
(b) Do the characterisations of $\varepsilon_{\mathcal{E}_0} = f^+\begin{pmatrix}\mathcal{E}_0 & 1 \\ 0 & 2\end{pmatrix}$ given in this paper give any new insights into the proof theory of elementary analysis?
(c) Do the characterisations of $\Gamma_0 = f^+\binom{2}{2}$ given in this paper give any new insights into the proof theory of predicative analysis?

**Acknowledgements** I am grateful to Dick de Jongh for inviting me to Amsterdam in 1973 and introducing me to his work, and to Professor Helmut Pfeiffer for inviting me to Hannover in 1977 and putting to me a problem which led me to the results in this paper; to Professor Gert H.Müller for his active interest in my personal and mathematical welfare since 1974; and, last but by no means least, to my husband and children for their tolerance and support.

[2]Except for the very rough first attempt in Chap. II, Sect. 4 of [21].

## References


1. Bachmann, H. (1950). Die Normalfunktionen und das Problem der ausgezeichneten Folgen von Ordnungszahlen. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, *95*, 5–37.
2. Bachmann, H. (1967). *Transfinite Zahlen*. Springer, VII + 228 pp.
3. Bridge, J. (1972). *Some problems in mathematical logic*. Systems of ordinal functions and ordinal notations: Thesis, Oxford University.
4. Bridge, J. (1975). A simplification of the Bachmann method for generating large countable ordinals. *Journal Symbolic Logic*, *40*, 171–185.
5. Bridge, J. (1978). A summary of the literature concerning ordinal notations. To appear.
6. Buchholz, W. (1975). Normalfunktionen und konstruktive Systeme von Ordinalzahlen. Proof Theory Symposion, Kiel 1974 (Springer Lecture Notes 500), 4–25.
7. Carruth, P. (1942). Arithmetic of ordinals with application to the theory of ordered abelian groups. *Bulletin of the American Mathematical Society*, *48*, 262–271.
8. Feferman, S. (1968). Systems of predicative analysis, II: Representations of ordinals. *Journal Symbolic Logic*, *33*, 193–220.
9. Halmos, P. (1960). *Naive set theory*. Van Nostrand, vii + 104 pp.
10. Heinzmann, G. (1976). *Eine Charakterisierung der Ordinalzahl* $\Gamma_0$. Diplomarbeit: Heidelberg University.
11. Higman, G. (1952). Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society*, *3*(2), 326–336.
12. Jongh, de D. H. J., & Parikh, R. (1977). Well-partial orderings and hierarchies. *Indagationes Mathematicae*, *39*, 195–207.
13. Kruskal, J. B. (1960). Well-quasi-ordering, the tree theorem, and Vaszonyi's conjecture. *Transactions of the American Mathematical Society*, *95*, 210–225.
14. Kruskal, J. B. (1972). The theory of well-quasi-ordering: a frequently discovered concept. *Journal Combinatorial Theory (A)*, *13*, 297–305.
15. Nash-Williams, C.St.J.A. (1963). On well-quasi-ordering finite trees. *Mathematical Proceedings of the Cambridge Philosophical Society*, *59*, 833–835.
16. Nash-Williams, C.St.J.A. (1965). On well-quasi-ordering transfinite sequences. *Mathematical Proceedings of the Cambridge Philosophical Society*, *61*, 33–39.
17. Nash-Williams, C.St.J.A. (1965a). On well-quasi-ordering infinite trees. *Mathematical Proceedings of the Cambridge Philosophical Society*, *61*, 697–720.
18. Nash-Williams, C.St.J.A. (1968). On better-quasi-ordering transfinite sequences. *Mathematical Proceedings of the Cambridge Philosophical Society*, *64*, 273–290.
19. Pfeiffer, H. (1970). Ein Bezeichnungssystem fur Ordinalzahlen. *Archiv f. math. Logik u. Grundlagenforschung*, *13*, 74–90.
20. Rado, R. (1954). Partial well-ordering of sets of vectors. *Mathematika*, *1*, 89–95.
21. Schmidt, D. (1972). Thesis, Oxford University.
22. Schmidt, D. (1975). Bounds for the closure ordinals of replete monotonic increasing functions. *Journal Symbolic Logic*, *40*, 305–316.
23. Schmidt, D. (1978). Associative ordinal functions, well partial orderings and a problem of Skolem. Zeitschr. f. math. *Logik u. Grundl. d. Math.*, *24*, 297–302.
24. Schütte, K. (1954). Kennzeichnung von Ordnungszahlen durch rekursiv erklärte Funktionen. *Math. Annalen*, *127*, 15–32.
25. Takeuti, G. (1957). Ordinal diagrams. *Journal of the Mathematical Society of Japan*, *9*, 386–394.
26. Takeuti, G. (1960). Ordinal diagrams, II. *Journal of the Mathematical Society of Japan*, *12*, 385–391.
27. Wolk, E. S. (1967). Partially well ordered sets and partial ordinals. *Fundamenta Mathematicae*, *60*, 175–186.
28. Ackermann, W. (1950). Konstruktiver Aufbau eines Abschnitts der zweiten Cantorschen Zahlenklasse. *Math. Zeit.*, *53*, 403–413.

29. Buchholz, W. (1977). Eine Erweiterung der Schnitt eliminations methode. Habilitations-schrift. München.
30. Buchholz, W., & Pohlers, W. (1978). Provable well orderings of formal theories for transfinitely iterated inductive definitions. *Journal of Symbolic Logic*, *43*, 118–125.
31. Feferman, S. (1964). Systems of predicative analysis I. *Journal of Symbolic Logic*, *29*, 1–30.
32. Feferman, S. (1968a). Systems of predicative analysis II: representations of ordinals. *Journal of Symbolic Logic*, *33*, 193–220.
33. Feferman, S. (1968b). Lectures on proof theory. *Proceedings of the Leeds Summer School, Springer Lecture Notes*, *70*, 1–108.
34. Feferman, S. (1968c). Autonomous transfinite progressions and the extent of predicative mathematics. *Logic, Methodology and Philosophy of Science III*, N. Holland, 121–135.
35. Feferman, S. (1970b). Formal theories for transfinita iterations of generalized inductive definitions and some subsystems of analysis. See [39], 303–326.
36. Gentzen, G. (1938). Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie. *Forsch. zur Logik. und zur Grund. der exakten Wissenschaften New Series. Leipzig*, *4*, 19–44.
37. Gentzen, G. (1943). Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie. *Mathematische Annalen*, *119*, 140–161.
38. Gerber, H. (1967). An extension of Schütte's Klarnmersymbols. *Mathematische Annalen*, *174*, 203–216.
39. Gerber, H. (1970). Brouwer's bar theorem and a system of ordinal notations. *Intuitionism and proof theory*. A. Kino, J. Myhill & R. E. Vesley. N. Holland, 1970, 327–338.
40. Gödel, K. (1958). Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica.* l2, 280–287.
41. Howard, W. (1970). Assignment of ordinals to terms for primitive recursive functionals of finite type. See [39], 443–458.
42. Howard, W. (1972). A system of abstract constructive ordinals. *Journal of Symbolic Logic*, *37*, 355–374.
43. Isles, D. (1970). Regular ordinals and normal forms. See [39], 339–362.
44. Kino, A. (1958). A consistency proof of a formal theory of Ackermann's ordinal numbers. *Journal of the Mathematical Society of Japan*, *10*, 287–303.
45. Kino, A. (1968). On provably recursive functions and ordinal recursive functions. *Journal of the Mathematical Society of Japan*, *20*, 456–476.
46. Kino, A. (1970). Formalization of the theory of ordinal diagrams of infinite order. See [39], 363–376.
47. Kreisel, G. (1963). Generalized inductive definitions. Stanford report.
48. Kreisel, G. (1965). Mathematical logic. Lectures on Modern Mathematics. Ed. Saaty, Wiley, 95–195.
49. Martin-Löf, P. (1971). Hauptsatz for the intuitionistic theory of iterated inductive definitions. *Proceedings of the 2nd. Scand. Logic Symposium*. N. Holland, pp. 179–216.
50. Pohlers, W. (1973). Eine Grenze für die Herleitbarkeit der transfiniten Induktion in einem schwachen $\Pi^l_l$ Fragment der Analysis. Dissertation. München.
51. Pohlers, W. (1975). An upper bound for the provability of transfinite induction in systems with N-times iterated inductive definitions. Proof theory symposium. Kiel 1974. *Springer Lecture Notes*, *500*, 271–289.
52. Pohlers, W. (1977). *Beweistheorie der iterierten Induktive Definitionen*. München: Habilitationsschrift.
53. Pohlers, W. (1978). Ordinals connected with formal theories for transfinitely iterated inductive definitions. *Journal of Symbolic Logic*, *43*, 161–182.
54. Schütte, K. (1964). Eine Grenze für die Beweisbarkeit der transfiniten Induktion in der verzweigten Typenlogik. *Archiv für mathematische Logik und Grundlagenforschung.*, *1*, 45–60.
55. Schütte, K. (1965). Predicative well orderings. Formal systems and recursive functions. N. Holland, pp. 280–303.

56. Schütte, K. (1968). Ein konstruktives System von Ordinalzahlen, I, II. *Archiv für mathematische Logik und Grundlagenforschung*, *11*, 126–137 and (12 (1969) 12–17)
57. Schütte, K. (1977). *Proof theory*. Springer.
58. Tait, W. (1965). Infinitely long terms of transfinite type. Formal systems and recursive functions. N. Holland, pp. 76–185.
59. Tait, W. (1968). Normal derivability in classical logic. The syntax and semantics of infinitary languages. *Springer Lecture Notes*, *72*, 204–256.
60. Takeuti, G. (1961). On the inductive definition with quantifiers of second order. *Journal of the Mathematical Society of Japan*, *13*, 333–341.
61. Takeuti, G. (1967). Consistency proofs of subsystems of classical analysis. *Annals of Mathematics*, *86*, 299–348.
62. Takeuti, G. (1975). *Proof theory*. N. Holland.
63. Zucker, J. (1971). Proof theoretic studies of systems of iterated inductive definitions and subsystems of analysis. Dissertation. Stanford.