Sapienza
Università di Roma

Ramsey-type Principles, Regressivity and Well-Orderings

Ingegneria dell'Informazione, Informatica e Statistica
PhD in Computer Science (XXXV cycle)

Leonardo Mainardi
ID number 1310167

Advisor
Prof. Lorenzo Carlucci

Academic Year 2023/2024

Thesis defended on 28 May 2024
in front of a Board of Examiners composed by:

Prof. Lamberto Ballan (chairman)
Prof. Giovanni Petri
Prof. Alessandro Raganato

Ramsey-type Principles, Regressivity and Well-Orderings
PhD thesis. Sapienza University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: leonardo.mainardi@uniroma1.it

Contents

Introduction

Ramsey's Theorem and Hindman's Theorem are among the principles that are studied the most in Reverse Mathematics, that is, the branch of Mathematics whose aim is, informally, to determine the minimal set of axioms needed to prove some theorem. This is likely due not only to historical reasons – one of the first independence results concerning Peano Arithmetic was indeed about a variation of Ramsey's Theorem, namely the Paris-Harrington Principle – but also derives from the fact that they are quite natural statements and require almost no coding in order to be formalized in Second Order Arithmetic. Moreover, their combinatorial nature put them in a privileged position, because of the close connection between Reverse Mathematics and the concept of "combinatorial core" (i.e., the part of Combinatorics needed to formalize and prove a principle), as pointed out by Dzhafarov and Mummert [DM] and previously by Hirschfeldt [Hir].
Also, both Ramsey's Theorem and Hindman's Theorem can be easily parameterized (for instance, by varying the number of colours): hence, several variations can be formulated and investigated using tools from Proof Theory and Computability Theory. Here, we study various versions of the two theorems by first relating them to different well-ordering preservation principles, i.e., to statements of the form *"if $\mathcal{X}$ is well-ordered, then $f(\mathcal{X})$ is well-ordered"*, where f is some (natural) operator transforming linear orders into linear orders. Many important subsystems of Second Order Arithmetic are known to be equivalent to such principles: therefore, after introducing (in Chapter 1) the needed concepts, in Chapter 2 – which is based on a joint work with Zdanowski [CMZ] – we extend a method proposed in [CZ1] for proving lower bounds on the logical strength of different Ramsey-theoretic principles using characterizations of formal systems (namely, ACA_0, ACA_0' and ACA_0^+) in terms of well-ordering preservation principles. We start by deriving the same implications obtained in [CZ1] concerning full Ramsey's Theorem and its restriction to colourings of triples, yet using a slightly different argument that allows us to extend the result to the stronger Ramsey's Theorem for relatively large sets (due to Pudlàk and Rödl and, independently, to Farmaki and Negrepontis), thus obtaining a direct combinatorial proof of a stronger well-ordering preservation principle (namely, the well-ordering preservation from $\mathcal{X}$ to $\varepsilon_{\mathcal{X}}$).
Moreover, we provide a direct implication from Hindman's Theorem (with the apartness property) for sums of three elements and two colours to the well-ordering preservation principle at the level of ACA_0.
Interestingly, all implications we obtain over RCA_0 also establish uniform computable reductions of well-ordering preservation principles to the corresponding Ramsey-type theorems.

Then, in Chapter 3, we focus on Ramsey's Theorem and Hindman's Theorem for infinitely many colours. We start by recalling that, when the Canonical Ramsey's Theorem by Erdős and Rado is applied to regressive functions, one obtains the Regressive Ramsey's Theorem (due to Kanamori and McAloon). Hence, we propose an analogous regressive variation of Taylor's "canonical" version of Hindman's Theorem: more precisely, we introduce the restriction of Taylor's Theorem to a subclass of the regressive functions, the λ-regressive functions, and prove some results about the Reverse Mathematics of this novel Regressive Hindman's Theorem and of natural restrictions of it.
In particular, we prove that imposing the so-called *apartness* condition on the solution set does not make the theorem stronger over RCA_0 (hence preserving a well-known property of the standard Hindman's Theorem), and that the first non-trivial restriction of the principle (with apartness) is equivalent to ACA_0. We furthermore point out that the well-ordering preservation principle for base-ω exponentiation is reducible to this same restriction by a uniform computable reduction, and observe that in this case the argument is more straightforward than the corresponding reduction to the bounded version of the standard Hindman's Theorem, due to the use of infinitely many colours.

Finally, in Chapter 4 (based on a joint work with Rathjen [CMR]), we recall how the outcome of ordinal analysis of a theory T carried out using well-orderings (in place of well-ordering preserving principles) is the least ordinal number whose well-foundedness can not be proved within T and that, if added as an axiom to T, makes the theory stronger. However, the strength of "augmented" theories of the form $\mathsf{RCA}_0 + \mathrm{WO}(\alpha)$, with α being at least the proof-theoretic ordinal of RCA_0 (i.e., ω^ω), has not been investigated much. Therefore, in this final chapter, we fill an apparent gap in the literature by showing that the proof-theoretic ordinal of the theory $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$ is δ^ω, for any ordinal δ satisfying $\omega \cdot \delta = \delta$ (e.g., ω^ω, ω^{ω^ω}, ε_0). Theories of the form $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$ are of interest in Proof Theory and Reverse Mathematics because of their connections to a number of well-investigated combinatorial principles related to various subsystems of Arithmetic: hence, our result also provides an ordinal analysis for all such theorems.

Chapter 1

Preliminary notions

All the results provided in the this thesis fall within the scope of Proof Theory and Reverse Mathematics. Hence, in Sec. 1.1, we give a brief introduction to these two topics, in order to provide an adequate context to the present dissertation and to explicitly formalize the main concepts discussed throughout the rest of the thesis. Furthermore, we will often refer to some notions related to Computability Theory, that are then formalized in Sec. 1.2.

1.1 Reverse Mathematics and Proof Theory

1.1.1 Reverse Mathematics

Any mathematical proof is carried out within some theory, which is made up of a language, a set of axioms and a set of inference rules.[1] However, in most branches of Mathematics, this aspect is not particularly relevant, and the axiomatic system adopted is typically left implicit. Yet, by focusing on the axioms used – or, even better, *needed* – to prove a theorem, it is possible to derive interesting information about the theorem itself, especially concerning its "strength".
This is indeed the ultimate goal of Reverse Mathematics: given a principle P and an adequately weak base theory T – i.e., a theory strong enough to prove basic facts about Mathematics, yet not so strong to already prove P – the aim is to figure out what axioms can be added to T in order to obtain P. If all of such axioms are actually necessary, then the result can be "reversed" – hence the name Reverse Mathematics – thus proving the axioms starting from P itself and, therefore, showing that they are equivalent to P over the theory T. This is interesting because it makes it possible to compare the strength of theorems, even concerning completely different topics in Mathematics. For instance, if we manage to get the *reversal* of two distinct theorems over the same theory, then we can conclude that the two theorems are equivalent to each other: in this sense, we can state – for instance – that Gödel's Completeness Theorem (a well-known result in Logic) is equivalent to Jordan's Curve Theorem (a principle pertaining Topology), since both of them are equivalent to the axioms of WKL_0 when using RCA_0 as a base theory (as showed by Simpson in [Sim2]; see Sec. 1.1.3 below for further details about RCA_0 and WKL_0).

[1] Often, all the theorems derivable from the axioms are considered part of the theory as well.

Furthermore, if we show that a principle P_1 is equivalent to a theory T_1 and a principle P_2 is equivalent to a different theory T_2, with $T_2 \supset T_1$ (that is, all theorems of T_1 are also theorems of T_2, but not vice versa), then of course the two theorems are not equivalent to each other, with P_1 being implied by P_2.
This is the case, for instance, of the infinite version of Ramsey's Theorem for triples and two colours, which can be proved to be strictly stronger than the aforementioned Jordan's Curve Theorem (or, equivalently, to Gödel's Completeness Theorem), since the former is equivalent to ACA_0, a theory including (hence, properly stronger than) WKL_0.
The idea, then, is to classify theorems according to their "equivalence classes", which in turn can be ordered based on their strength – or, more precisely, based on the relation given by the implications between theorems belonging to different classes.[2]
This way, Reverse Mathematics also helps to highlight possible clustering phenomena around some theories: this is exactly what happened in the 1970s, when Harvey Friedman noticed [Fri2, Fri3] that a number of theorems of ordinary Mathematics resulted to be equivalent to five specific axiomatic systems[3], later named "The Big Five".
However, in the last few decades, many mathematical principles have been shown to be equivalent to none of these five systems, hence making the whole picture of the relations between theorems much more complex: nevertheless, the Big Five have maintained a certain importance in Reverse Mathematics and they are still widely used today. Many results in the present dissertation are related to some of these systems, so we formally define them in Sec. 1.1.3.

1.1.2 Second Order Arithmetic

Reverse Mathematics can be carried out over any axiomatic system, yet in most cases Second Order Arithmetic (or some of its subsystems) is used. According to [DM], this is because "*it is strong enough to formalize much of ordinary mathematics, but weak enough that it does not overshadow the theorems being studied in the way that set theory does*". Second Order Arithmetic is indeed a much stronger theory than its first-order counterpart Peano Arithmetic (PA), since it uses a language $\mathcal{L}_2$ with two sorts of variables – the *number variables* and the *set variables* – to both of which quantifiers can be applied. This, together with the introduction of the membership symbol $\in$ and an additional axiom for constructing sets, allows Second Order Arithmetic to be more expressive and more powerful than PA: for instance, it can formalize and prove results about real numbers, while PA can not. Hence, the importance of Second Order Arithmetic in Reverse Mathematics is due to the fact that the statement of many mathematical principles can not even be expressed in PA.
Formally, this theory can be formalized as in the next Definition (following [DM]).

[2] Notice, however, that this order can not be linear since, by Gödel Incompleteness Theorem, there are arithmetical theories T that can be consistently extended using (alternately) both a new axiom A and its negation $\neg A$: in that case, none of these two extended theories (i.e., $T \cup \{A\}$ and $T \cup \{\neg A\}$) implies the other.

[3] Actually, the weakest of these five subsystems – that is, the aformentioned system RCA_0 – is only used as a base theory, hence proper equivalences with mathematical principles are only obtained with the remaining four subsystems.

Definition 1.1.1. Second order arithmetic (Z_2) *is defined by the following axioms:*

1. $(\forall x, y, z)[(x + y) + z = x + (y + z)]$
2. $(\forall x, y)[x + y = y + x]$
3. $(\forall x, y, z)[(x \cdot y) \cdot z = x \cdot (y \cdot z)]$
4. $(\forall x, y)[x \cdot y = y \cdot x]$
5. $(\forall x, y, z)[x \cdot (y + z) = x \cdot y + x \cdot z]$
6. $(\forall x)[x + 0 = x \wedge x \cdot 0 = 0 \wedge x \cdot 1 = x]$
7. $(\forall x, y, z)[x < y \wedge y < z \rightarrow x < z]$
8. $(\forall x)[x \not< x]$
9. $(\forall x, y)[x < y = \vee x = y \vee y < x]$
10. $(\forall x, y, z)[x < y \rightarrow x + z < y + z]$
11. $(\forall x, y, z)[0 < z \wedge x < y \rightarrow x \cdot z < y \cdot z]$
12. $(\forall x, y)(\exists z)[x < y \rightarrow x + z = y]$
13. $(0 < 1) \wedge (\forall x)[x > 0 \rightarrow x \geq 1]$
14. $(\forall x)[(x = 0 \vee x > 0)]$
15. $[\varphi(0) \wedge (\forall x)[(\varphi(x) \rightarrow (\varphi(x + 1)]] \rightarrow (\forall x)\varphi(x)$*, for all* $\mathcal{L}_2$*-formulas* φ
16. $(\exists X)(\forall x)[x \in X \leftrightarrow \varphi(x)]$*, for all* $\mathcal{L}_2$*-formulas* φ *containing no set variables* X*.*

Axioms 1-14 form a first-order subtheory that is often called PA^-, for it corresponds to the axiomatization of PA without one axiom (namely, the axiom 15 for $\mathcal{L}_1$, that is the language of PA).
Axioms 15 and 16 (named, respectively, *induction axiom* and *comprehension axiom*) are actually axiom schemas, as they consist of infinitely many axioms (one for each $\mathcal{L}_2$-formula φ). As we will explain in the next Section, both of them can be restricted by limiting the scope of φ to some subclass Γ (possibly different for the two axioms) of $\mathcal{L}_2$-formulas: in that case, they are typically denoted by $\mathsf{I}\Gamma$ and Γ-CA, respectively. These subclasses are usually specified using arithmetical and analytical hierarchies, so it is worth to formalize these two notions.

Definition 1.1.2. *The* arithmetical hierarchy *consists of the following classes of formulas in* $\mathcal{L}_2$*.*

- *The classes* Δ^0_0*,* Σ^0_0 *and* Π^0_0 *are the same, and consist of all formulas with no unbounded quantifiers.*
- *For any* $n \in \omega$*,* Σ^0_{n+1} *is the class of formulas* $\exists x_1 \exists x_2 \ldots \exists x_k \, \psi(x_1, x_2, \ldots, x_k)$*, where* $k > 0$ *and* ψ *is a formula in* Π^0_n*.*
- *For any* $n \in \omega$*,* Π^0_{n+1} *is the class of formulas* $\forall x_1 \forall x_2 \ldots \forall x_k \, \psi(x_1, x_2, \ldots, x_k)$*, where* $k > 0$ *and* ψ *is a formula in* Σ^0_n*.*

When dealing with the first order language $\mathcal{L}_1$, the same classification can be adopted, but the superscript 0 is often omitted.

It is noteworthy that not all arithmetical formulas belong to the arithmetical hierarchy: however, any $\mathcal{L}_2$-formula (and any $\mathcal{L}_1$-formula as well) can be written in prenex normal form, therefore each arithmetical formula can actually be related to some class in Definition 1.1.2.

Definition 1.1.3. *The* analytical hierarchy *consists of the following classes of formulas in* $\mathcal{L}_2$.

- *The classes* Δ^1_0, Σ^1_0 *and* Π^1_0 *are the same, and consist of all formulas with no set quantifiers, called* arithmetical formulas.
- *For any* $n \in \omega$, Σ^1_{n+1} *is the class of formulas* $\exists X_1 \dots \exists X_k\, \psi(X_1, \dots, X_k)$, *where* $k > 0$ *and* ψ *is a formula in* Π^1_n.
- *For any* $n \in \omega$, Π^1_{n+1} *is the class of formulas* $\forall X_1 \dots \forall X_k\, \psi(X_1, \dots, X_k)$, *where* $k > 0$ *and* ψ *is a formula in* Σ^1_n.

Restrictions on induction and comprehension axioms based on arithmetical and analytical hierarchies are fundamental in this thesis and, in general, in Reverse Mathematics. However, in order to be effective (especially in terms of models), they require a specific semantics for Z_2, which is briefly explained in the next paragraph.

Semantics of Z_2. Second Order Arithmetic is commonly interpreted in two different ways. The actual "second order" interpretation is called *full semantics*, and adopts structures that only specify the first-order domain – that is, the domain of number variables – and the meaning of each symbol in $\mathcal{L}_2$, hence assuming set variables to range over any possible subset of the domain.
However, full semantics is quite useless in Reverse Mathematics, for it causes Second Order Arithmetic to become a *categorical* theory – i.e., a theory admitting only one model (up to isomorphism), and that has some undesirable consequences (for instance, Gödel's Completeness Theorem no longer holds).
Therefore, in Reverse Mathematics (and in this thesis as well), *Henkin semantics* is commonly used in place of full semantics, meaning that set variables can range over an arbitrary subset of the power set of the domain (and not necessarily over the whole power set, as in full semantics), thus turning Z_2 into a first-order theory[4]. In that case, however, several logical metatheorems (such as the Completeness Theorem and the Compactness Theorem) do hold, hence allowing models of Z_2 – or of any of its subsystems – to have non-standard elements in their domain. Then, as a consequence of adopting Henkin semantics, we have to distinguish the domain of our models from the set of standard natural numbers: so, in the rest of the thesis, we adopt the notation $\mathbf{N}$ to denote the domain of any (possibly non-standard) model of the theory being used, while we reserve ω for the set of standard natural numbers.

1.1.3 Subsystems of Second Order Arithmetic

As mentioned above, Second Order Arithmetic is much weaker than other well-known theories – like ZFC: nevertheless, it often turns out to be still too powerful to be

[4]Nevertheless, in this thesis we will occasionally refer to "second order statements", which might then sound as an abuse of terminology. However, recall that our theories use a second-order language: hence, by this expression, we mean any statement containing at least one set variable.

used as a base system in Reverse Mathematics, since the axioms of Z_2 already imply most ordinary mathematical theorems. The solution, then, is to weaken some of these axioms, hence obtaining subsystems of the full theory. Of course, there exist infinitely many subsystems of Z_2, each given by an arbitrary restriction of the original axioms: however, some choices turn out to be more meaningful than others, mainly because of the "Big Five phenomenon". So, the vast majority of results in Reverse Mathematics makes use of one out of these five prominent subsystems: RCA_0, WKL_0, ACA_0, ATR_0 and Π^1_1-CA_0. All of them are obtained by restricting the induction axiom and the comprehension axiom schema of Z_2 (WKL_0 and ATR_0 also contain further axioms). [Sim2] and [DM] are good references about this topic, while here we only give some details about RCA_0 and ACA_0, as they are the only subsystems of Z_2 among the Big Five that are used within this thesis. The former is the most commonly adopted as a base system, because several useful basic principles are already derivable in RCA_0, but it still leaves space for Reverse Mathematics analysis of most ordinary Mathematics theorems, for they often require stronger systems to be proved. RCA_0 is indeed a quite weak theory: its name stands for *Recursive Comprehension Axiom*, hence it formalizes only the recursive part – i.e., the computable part – of Arithmetic. This is achieved by limiting the induction axiom to Σ^0_1-formulas and the comprehension axiom schema as in the next definition.

Definition 1.1.4. *The schema of* Δ^0_1-comprehension *consists of all axioms of the form*

$$\forall x\,(\varphi(x) \leftrightarrow \psi(x)) \to (\exists X)(\forall x)[x \in X \leftrightarrow \varphi(x)]$$

where $\varphi(x)$ is a Σ^0_1-formula with no free occurrences of X, and $\psi(x)$ is a Π^1_1-formula.

We denote Δ^0_1-comprehension by Δ^0_1-CA, thus extending the notation Γ-CA introduced in the previous Section (where Γ was required to be a class of formulas, which Δ^0_1 is not).

Definition 1.1.5. RCA_0 *is the formal system obtained by adding to* PA^- *the axioms* $\mathsf{I}\Sigma^0_1$ *and* Δ^0_1-CA*.*

The reason behind the restriction over the comprehension axiom is clear: it precludes the possibility of constructing sets that are not decidable (Δ^0_1-definable sets are indeed the class of sets for which we can computably decide whether an element belongs to the set or not). The choice of limiting induction to Σ^0_1-formulas, instead, is due both to historical reasons and to the necessity of reaching a compromise between not allowing induction at all (an extremely constructivist approach) and allowing full induction, with both approaches having valid philosophical justifications. Again, a detailed analysis can be found in [DM], where it is also highlighted how Σ^0_1-induction is necessary (actually, equivalent) to carry out even just finite recursion, and how including stronger induction principles would break relevant conservativity results of RCA_0 with respect to weaker, finitistic theories.

The other subsystem of Z_2 that will be largely mentioned in this thesis is ACA_0, whose name is the acronym for *Arithmetical Comprehension Axiom*: this theory is indeed obtained by adding the arithmetical comprehension axiom schema to RCA_0.

Definition 1.1.6. ACA_0 *is the formal system obtained by adding to* RCA_0 *the axioms* Σ^0_k-CA*, for all $k \in \omega$.*

Therefore, in ACA_0, we can derive the existence of any set defined by an arithmetical formula, i.e. a formula containing no set quantifiers. This, together with Σ^0_1-induction, also implies arithmetical induction (that is, induction over arithmetical formulas), which is then included in ACA_0.
Of course, such a system is much stronger than RCA_0: in fact, ACA_0 is a conservative extension of Peano Arithmetic, meaning that any theorem of PA is a theorem of ACA_0 and that any theorem of ACA_0 in the language of PA is already a theorem of PA (in other words, PA is the first-order part of ACA_0). This system is quite recurring in Reverse Mathematics, since it turns out to be equivalent to a large number of well-known mathematical principles, as is the case with several combinatorial theorems discussed in this thesis.

Then, many equivalent formulations of ACA_0 can be found in the literature. The next Theorem summarizes the principles used in the next chapters to characterize ACA_0. In the statement of the Theorem and throughout the rest of this section, we mention the notion of *Turing jump*, which will be formalized in Section 1.2: informally, the Turing jump of a set X – in symbols, X' – is the Halting Problem relativized to X, that is, the set of indexes i of algorithms (in some fixed enumeration of algorithms) that halt on input i when using an oracle for X.

Theorem 1.1.7. *Over* RCA_0*, the following are equivalent:*

1. ACA_0.
2. *For any injective function* $f : \mathbf{N} \to \mathbf{N}$ *there exists a set* $X \subseteq \mathbf{N}$ *such that* $\forall n(n \in X \leftrightarrow \exists m(f(m) = n))$*, i.e., the range of any injective function exists.*
3. $\forall X \exists Y(Y = X')$*, i.e., the Turing jump of any set exists.*

Proof. (1) is equivalent to (2) by Lemma III.1.3 in [Sim2], while (1) is equivalent to (3) by Corollary 5.6.3 in [DM]. ☐

We denote by RAN the principle (2) in Theorem 1.1.7. By extending (3) via an iteration of the Turing jump operator, instead, we can obtain theories that are strictly stronger than ACA_0. This idea is indeed applied to formalize two further subsystems of Z_2 that are widely used in Reverse Mathematics, namely ACA_0' and ACA_0^+. The former is based on the notion of n-th Turing jump, defined in $\mathcal{L}_2$ as follows: fixed $X \subseteq \mathbf{N}$, we set $X^{(0)} = X$ and $X^{(n+1)} = (X^{(n)})'$ for any n.

Definition 1.1.8. ACA_0' *is the formal system obtained by adding to* RCA_0 *the following axiom:* $\forall n \forall X \exists Y(Y = X^{(n)})$.

Notice that, in the previous definition, n is internally quantified, hence it can be non-standard. This aspect is crucial, for it is what makes ACA_0' strictly stronger than ACA_0: indeed, any non-standard model $\mathcal{M}$ of ACA_0 including only arithmetically definable sets does not contain $\emptyset^{(e)}$ for e non-standard, otherwise such a set would be Σ^0_n-definable for some standard $n < e$, and then Turing-reducible to $\emptyset^{(n)}$ by Post's Theorem, thus contradicting the fact (provable in ACA_0) that $\emptyset^{(m)} \not\geq_T \emptyset^{(m+1)}$ for any m, since $\emptyset^{(e)} \leq_T \emptyset^{(n)}$ by the argument above and $\emptyset^{(n+1)} \leq_T \emptyset^{(e)}$ due to $(n+1) < e$.

Then, by introducing another Turing jump-related concept, we can formalize in $\mathcal{L}_2$ the system ACA_0^+: to do this, we define the ω-jump of $X \subseteq \mathbf{N}$ as a set coding all the n-th Turing jumps of X; more formally, $X^{(\omega)} = \{\langle x, n\rangle \mid n \in \mathbf{N}, x \in X^{(n)}\}$, where $\langle \cdot, \cdot \rangle$ is any computable pairing function.

Definition 1.1.9. ACA_0^+ *is the formal system obtained by adding to* RCA_0 *the following axiom:* $\forall X \exists Y (Y = X^{(\omega)})$.

ACA_0^+ is a strengthening of ACA_0', and these two theories represent nice examples of subsystems of Z_2 not belonging to The Big Five but still quite recurring in Reverse Mathematics: they are indeed equivalent to several natural principles – or, in some cases, they are just the best known upper bound – as we will see in the next Chapters.

1.1.4 Well-Orderings and Ordinal Analysis

Although not strictly related to Reverse Mathematics, well-orderings represent a relevant topic in Proof Theory. This is because they are often used as an alternative method to measure the strength of theorems and theories, at least since Gentzen has shown that, over the weak first order theory PRA, the well-ordering[5] of ε_0 implies the consistency of PA and then, by Gödel's Incompleteness Theorem, it is not derivable in PA (in contrast to the well-ordering of any other ordinal below ε_0). Since then, plenty of theorems and theories have been related to some well-ordering. Hence, it is worth formalizing some relevant definitions and results about this topic.
First, let us recall the definitions of well-founded ordering and well-ordering.

Definition 1.1.10. *An ordering* $\mathcal{X} = (X, <_{\mathcal{X}})$ *is* well-founded *if there exist no infinite strictly* $<_{\mathcal{X}}$*-descending sequences of elements in* X*, i.e., there are no functions* $f : \omega \to X$ *such that* $f(n) >_{\mathcal{X}} f(n+1)$ *for all* $n \in \omega$*. In that case, we write* $\mathrm{WF}(\mathcal{X})$.

Alternative definitions of well-founded ordering can be found in the literature: here, we are only interested in the following version.

Definition 1.1.11. *An ordering* $\mathcal{X} = (X, <_{\mathcal{X}})$ *is* well-founded *if every non-empty subset of* X *contains a* $<_{\mathcal{X}}$*-minimal element.*

$$(\forall Y \subseteq X)[Y \neq \emptyset \to (\exists m \in Y)(\forall y \in Y)(m \leq_{\mathcal{X}} y)]$$

In that case, we write $\mathrm{ME}(\mathcal{X})$.

However, we can show that in RCA_0 these two versions are equivalent to each other.

Theorem 1.1.12. *Over* RCA_0*,* $\mathrm{WF}(\mathcal{X})$ *is equivalent to* $\mathrm{ME}(\mathcal{X})$ *for any ordering* $\mathcal{X} = (X, <_{\mathcal{X}})$.

Proof. First, suppose $\mathrm{WF}(\mathcal{X})$ holds but $\mathrm{ME}(\mathcal{X})$ does not. Let S be a non-empty subset of X with no $<_{\mathcal{X}}$-minimal element, and let $\{s_0, s_1, \ldots\}$ be any enumeration of S. We can computably construct a function $f : \omega \to X$ by setting $f(0) = s_0$

[5] The actual result uses the principle of *transfinite induction* in place of well-ordering, for the latter is a second order statement, hence outside the language of PRA. However, well-orderings can be somehow formalized in $\mathcal{L}_1$ as well: details are given below and in Chapter 4, where indeed we formalize and derive well-orderings within a first order theory.

and $f(n+1) = s_k$, where k is minimal such that $s_k <_{\mathcal{X}} f(n)$. The existence of k is guaranteed by the definition of S. Then, we have a contradiction with $\mathrm{WF}(\mathcal{X})$, since f is an infinite strictly $<_{\mathcal{X}}$-decreasing sequence of elements in X.

Now we prove that $\mathrm{ME}(\mathcal{X})$ implies $\mathrm{WF}(\mathcal{X})$. In the argument below, recall that $<$ might be different from $<_{\mathcal{X}}$. Then, assume $\mathrm{ME}(\mathcal{X})$ and, by way of contradiction, let $f : \omega \to X$ be an infinite strictly $<_{\mathcal{X}}$-descending sequence in $\mathcal{X}$. Notice that the range of f contains no $<_{\mathcal{X}}$-minimal element, but the existence of such a set is not provable in RCA_0. Therefore, we need to adopt a different approach in order to reach a contradiction. In particular, we recursively construct a set $S = \{s_0 < s_1 < \cdots\}$ by defining $s_0 = f(0)$ and, if $s_i = f(n)$ for some n, then $s_{i+1} = f(n+k)$, where $k \in (0, s_i + 1]$ is minimal such that $f(n+k) > f(n+j)$ for all $j < k$. Clearly each s_{i+1} exists, otherwise $\{f(n+1), f(n+2), \ldots, f(n+s_i+1)\}$ would only contains elements less than s_i, which is impossible since it contains (s_i+1) many elements. Then, S is infinite, so for any m such that $f(m) \in S$, there exists $n > m$ such that $f(n) \in S$. Also, $f(n) <_{\mathcal{X}} f(m)$, since we assumed f to be strictly $<_{\mathcal{X}}$-decreasing. Therefore, S does not contain a $<_{\mathcal{X}}$-minimal element, contra our hypothesis. □

We can then formalize the concept of well-ordering.

Definition 1.1.13. *A well-founded linear ordering $\mathcal{X}$ is called a* well-ordering*, in symbols* $\mathrm{WO}(\mathcal{X})$*.*

It is trivial to figure out linear orderings that are also well-founded, as is the case with the usual ordering on the standard natural numbers. Similarly, we can easily come up with examples of non well-founded linear orderings, e.g. the usual $<$-relation over the set of negative integers. As with any other mathematical principle, we can think of well-foundedness of some ordering $\mathcal{X}$ just as a theorem of a certain theory: therefore, in Reverse Mathematics, it is natural to wonder whether $\mathrm{WF}(\mathcal{X})$ can be proved within a specific formal system. This analysis is usually carried out over infinite linear orderings which – being also well-founded – are then isomorphic to transfinite ordinal numbers. However, the well-foundedness of a linear ordering $\mathcal{X}$ isomorphic to some transfinite ordinal α – hence, shortly, $\mathrm{WF}(\alpha)$ – always implies at least $\mathrm{WF}(\alpha+1)$: for this reason, the actual question usually pondered in Reverse Mathematics is rather about the least transfinite ordinal α whose well-foundedness can *not* be proved within a certain theory. There are indeed a large number of such results in the literature, one example being the independence of $\mathrm{WO}(\varepsilon_0)$ from ACA_0, straightforwardly derivable from the aforementioned Gentzen's proof of consistency of PA plus Gödel's Incompleteness Theorem and the Π^1_1-conservativity of ACA_0 over PA. In that case, it appears to be just a corollary of some major results: however, starting from Gentzen's result, this kind of analysis – later called *ordinal analysis* – became an interesting line of research on its own.

It is important to note that both Definition 1.1.10 and Definition 1.1.13 are second order statements, namely Π^1_1-statements: hence, the outcome of the ordinal analysis of a theory T – i.e., the least ordinal α such that T does not prove $\mathrm{WO}(\alpha)$ – is called the Π^1_1-*ordinal* of T. We formalize this concept in the next Definition.

Definition 1.1.14. *The* Π^1_1-ordinal *of a theory* T *– in symbols,* $ord(T)$ *– is the least ordinal* α *isomorphic to some primitive recursive linear ordering* $\mathcal{X}$ *whose well-foundedness can not be proved in* T*.*

Many alternative definitions concerning the ordinal of a theory can be found in the literature, even using just first order statements: however, the definition given above is probably the most adopted in Proof Theory, especially when dealing with Second Order Arithmetic.
An arguable aspect about the Π^1_1-ordinal of a theory T is that it does not catch all the nuances of the strength of T. For instance, the axiomatic systems PRA, RCA_0 and WKL_0 share all the same Π^1_1-ordinal – i.e., ω^ω – but they are not equivalent to each other, with WKL_0 being stronger than RCA_0 and RCA_0, in turn, being stronger than PRA.[6] Nevertheless, Π^1_1-ordinal analysis is still a useful and widely adopted method for measuring the strength of theories.

An alternative approach to ordinal analysis is instead carried out by using the preservation of well-orderings, formalized in the following definition.

Definition 1.1.15. *Let* f *be a function from linear orderings to linear orderings. Then, by* well-ordering preservation principle *– or just by* well-ordering principle *– we mean the following statement:*

$$\forall \mathcal{X}(\mathrm{WO}(\mathcal{X}) \to \mathrm{WO}(f(\mathcal{X}))).$$

If the previous statement holds, we write $\mathsf{WOP}(\mathcal{X} \mapsto f(\mathcal{X}))$.

Results concerning this kind of ordinal analysis typically show the equivalence – over a base theory – between a subsystem S of Z_2 and a certain well-ordering principle, i.e. between S and $\mathsf{WOP}(\mathcal{X} \mapsto f(\mathcal{X}))$ for some specific f. Such an equivalence can hardly be achieved when using Π^1_1-ordinals, since most of the relevant subsystems of Z_2 are extensions of RCA_0 obtained by adding axioms whose complexity is higher than Π^1_1. These axioms are often Π^1_2-statements (as in the case of ACA_0', ACA_0^+ and ATR_0)[7], which is indeed the complexity of the well-ordering principles.
Some notable results in this topic are reported below, where – adopting the same notation defined in [MM], also formalized in Sec. 2.2 below – $\boldsymbol{\omega}$, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varphi}$ are computable operators from linear orderings to linear orderings behaving on well-orderings like the usual ω, ε and φ functions do on ordinals.

---

[6]Here, the symbol ω does not stand for the set of standard natural numbers; rather, it represents the least transfinite ordinal: we use the same symbol since the latter is isomorphic to the usual $<$-ordering on the standard natural numbers. However, none of the mentioned theories can define the notion of ordinal number: hence, by a (quite common) abuse of notation, when we refer to some ordinal α within some subsystem T of Z_2 (for instance, when we assert that RCA_0 proves $\mathrm{WO}(\omega^2)$), we actually mean some T-definable ordering on the elements in the (possibly non-standard) domain of T that, when "observed" from the metatheory, corresponds to α. For this reason, when dealing with transfinite ordinals, we define $\boldsymbol{\omega}$, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varphi}$ operators to obtain orderings in RCA_0 that, in the standard model, are isomorphic to ordinals larger than ω.

[7]The fact that ACA_0' and ACA_0^+ are characterized by Π^1_2-axioms is clear from Definition 1.1.8, Definition 1.1.9 and the fact that the concept of Turing jump is arithmetically definable. For ATR_0, refer to Definition V.2.2 and Definition V.2.4 in [Sim2]. Finally, notice that even ACA_0 is clearly characterized by a Π^1_2-axiom when considering its equivalent formulations given in Theorem 1.1.7.

Theorem 1.1.16 (Girard [Gir], Hirst [Hir2]). *Over* RCA_0*,* ACA_0 *is equivalent to* $\mathsf{WOP}(\mathcal{X} \mapsto \boldsymbol{\omega}^{\mathcal{X}})$.

Theorem 1.1.17 (Marcone-Montalbán [MM]). *Over* RCA_0*,* ACA_0' *is equivalent to* $\forall n \mathsf{WOP}(\mathcal{X} \mapsto \boldsymbol{\omega}^{\langle n, \mathcal{X} \rangle})$*, where* $\boldsymbol{\omega}^{\langle n, \mathcal{X} \rangle}$ *denotes the ordering obtained by applying* n *times the operator* $\boldsymbol{\omega}$ *to* $\mathcal{X}$.

Theorem 1.1.18 (Marcone-Montalbán [MM], Afshari-Rathjen [AR]). *Over* RCA_0*,* ACA_0^+ *is equivalent to* $\mathsf{WOP}(\mathcal{X} \mapsto \boldsymbol{\varepsilon}_{\mathcal{X}})$.

Theorem 1.1.19 (Marcone-Montalbán [MM]). *Over* RCA_0*,* $\Pi^0_{\omega^\alpha} - \mathsf{CA}$ *is equivalent to* $\mathsf{WOP}(\mathcal{X} \mapsto \boldsymbol{\varphi}(\alpha, \mathcal{X}))$.

Theorem 1.1.20 (Friedman-Montalbán-Weiermann [FMW]; Rathjen-Weiermann [RW]; Marcone-Montalbán [MM]). *Over* RCA_0*,* ATR_0 *is equivalent to* $\mathsf{WOP}(\mathcal{X} \mapsto \boldsymbol{\varphi}(\mathcal{X}, 0))$.

In Chapter 2, we carry out such an analysis for theories including different versions of Ramsey's Theorem. The outcome of this analysis could also be obtained by noting that all of such theories are equivalent to ACA_0, ACA_0' or ACA_0^+, thus it would suffice to apply the aforementioned results to derive the well-ordering principle associated with each of these theories. However, we propose novel combinatorial proofs relating these versions of Ramsey's Theorem – and a restriction of Hindman's Theorem as well – to the respective well-ordering principles; besides showing direct implications, our proofs also witness different kinds of computable reductions (defined in the next Section) between such principles, hence giving a deeper insight into the actual relations between Ramsey's Theorem and well-ordering principles.

1.2 Notions of reducibility

1.2.1 Basic definitions in Computability theory

Despite being mainly proof-theoretic related, our results will sometimes refer to computability concepts, which are then worth a formal definition.
We start with the fundamental notion of halting set.

Definition 1.2.1 (Halting set). *By* halting set relative to a set X *we mean the set*

$$X' = \{x \in \omega \mid \Phi^X_x(x) \text{ is defined}\}$$

where $(\Phi^X_i)_{i \in \omega}$ *is any fixed enumeration of* X*-computable partial functions.*
The set $\emptyset'$ *is usually called just* halting set *and is often denoted by* K.

The existence of the enumeration $(\Phi^X_i)_{i \in \omega}$ can be easily proved, and it can be formalized even in weak theories like RCA_0.

Starting from the previous Definition, we can then formalize the concept of Turing jump.

Definition 1.2.2 (n-th Turing jump). *For any* $n \in \omega$*, we denote by* $X^{(n)}$ *the* n-th Turing jump *of a set* X*, that is recursively defined as follows:*

- $X^{(0)} = X$
- $X^{(m+1)} = (X^{(m)})'$, *for any* $m < n$.

Notice that $X^{(1)} = X'$ *and it is usually called just* Turing jump of X.

Definition 1.2.3 (ω-Turing jump). *The* ω-Turing jump *of a set* X, *or just the* ω-jump of X, *is the set*

$$X^{(\omega)} = \{\langle x, n\rangle \mid n \in \omega,\ x \in X^{(n)}\}$$

where $\langle \cdot, \cdot \rangle$ *is any fixed, computable paring function.*

As we already noticed in Section 1.1, these notions are definable in $\mathcal{L}_2$ and are useful to define some relevant subsystems of Z_2, namely ACA_0' and ACA_0^+. We also showed how the former theory does not imply the latter, by a simple argument based on Post's Theorem, which we now formalize for the sake of completeness. Recall that a set X is *recursively enumerable in* Y if X is the domain of a partial Y-computable function, while it is *computable* (also, *recursive* or *decidable*) *in* Y if both X and its complement are recursively enumerable in Y.

Theorem 1.2.4 (Post's Theorem). *Let* $n \in \omega$.

- *A set is* Σ^0_{n+1}*-definable if and only if it is recursively enumerable in* $\emptyset^{(n)}$.
- *A set is* Π^0_{n+1}*-definable if and only if its complement is recursively enumerable in* $\emptyset^{(n)}$.

A simple corollary of Post's Theorem is that any Δ^0_{n+1}-definable set X is computable in $\emptyset^{(n)}$, meaning that, for any x, we can computably check whether $x \in X$ or not, using an oracle for $\emptyset^{(n)}$. This idea anticipates the notion of reducibility, that is discussed in the next sections for its relevance within the results presented in this dissertation.

1.2.2 Overview of reductions

The basic idea behind the concept of reducibility is to investigate whether a specific problem can be solved by using the solution to another problem. Such a question is of interest, for a positive answer would make it possible to link a novel problem to an existing one – which we might already know how to solve: in that case, a reduction would make it easier to obtain a solution to the original problem.
In this sense, we are mainly interested in computable reductions, i.e. reductions that can be carried out by a Turing machine, or by any other equivalent computational model allowed by the Church-Turing thesis. A first example of such a kind of reducibility is given by *Turing reducibility.*

Definition 1.2.5 (Turing reducibility). *A set* X *is* Turing-reducible *to a set* Y, *in symbols* $X \leq_{\mathrm{T}} Y$, *if* X *can be decided by a Turing machine with an oracle for* Y.

If two sets X, Y Turing-reduce each other, i.e. $X \leq_{\mathrm{T}} Y$ and $Y \leq_{\mathrm{T}} X$, then X and Y are said to be *Turing equivalent*, in symbols $X \equiv_{\mathrm{T}} Y$. Since $\equiv_{\mathrm{T}}$ is an equivalence

relation, we can define the notion of *Turing degree of* X as the equivalence class of X under $\equiv_{\mathrm{T}}$. For instance, any decidable set belongs to $[\emptyset]_{\equiv_{\mathrm{T}}}$. However, these equivalence classes do not form a chain, due to $\leq_{\mathrm{T}}$ not being a linear order; therefore, there exist sets $\leq_{\mathrm{T}}$-incomparable.
One notable result concerning $\leq_{\mathrm{T}}$ is the following.

Proposition 1.2.6. *For any* $n \in \omega$, $\emptyset^{(n)} \lneq_T \emptyset^{(n+1)}$.

Proof. It is derivable from the prominent paper of Turing [Tur]. □

1.2.3 Computable reductions

Turing reductions are defined on sets, that can be considered as collections of instances (with a positive answer) to decision problems, i.e. to problems posing questions with boolean answers: for instance, whether a number is prime, a graph is planar, or a map can be coloured using 3 colours and avoiding adjacent regions to have the same colour.
However, many mathematical problems have a different form, that can be represented using a Π^1_2-formula (hence they are typically called Π^1_2-*problems*):

$$\forall X\,(\varphi(X) \to \exists Y\,\psi(X,Y))$$

where φ and ψ are arithmetical formulas. We say that X is an *instance* of a Π^1_2-problem P if $\varphi(X)$ holds, and that Y is a *solution* for X to P if $\psi(X,Y)$ holds. We are able to solve a Π^1_2-problem P if we can find a solution for any instance of P.
Notice that the well-ordering principles discussed in Section 1.1 can be restated as Π^1_2-principles. This can be done by considering their contrapositive form:

$$\forall \mathcal{X}(\neg \mathrm{WO}(f(\mathcal{X})) \to \neg \mathrm{WO}(\mathcal{X})) \tag{1.1}$$

hence redefining such principles in terms of instances (that is, infinite descending sequences of elements in $f(\mathcal{X})$) and solutions (infinite descending sequences of elements in $\mathcal{X}$). Actually, this formulation does not fully comply with the definition of Π^1_2-problem given above, since $\neg\mathrm{WO}(f(\mathcal{X}))$ is not arithmetical. Then, a more pedantic reformulation of a well-ordering principle as a Π^1_2-problem would rather be:

$$\forall \mathcal{X}(\forall s\!:\!\mathbf{N}\to f(\mathcal{X}))\big[\forall x(s(x) >_{f(\mathcal{X})} s(x{+}1)) \to (\exists t\!:\!\mathbf{N}\to\mathcal{X})\big(\forall x(t(x) >_{\mathcal{X}} t(x{+}1))\big)\big].$$

In that case, an instance of the problem would be a Turing join between $\mathcal{X}$ and s. By adopting this formulation, we might even introduce a further constraint on the solutions, namely we might require t to only contain $\mathcal{X}$-terms appearing in s: in fact, all our solutions will satisfy this additional property. This can be done because ψ in the definition of Π^1_2-*problem* takes X as a parameter, meaning that the condition defining the solutions can use information about the instance. However, for the sake of readability, throughout the rest of the thesis we will omit all these details while dealing with well-ordering principles, and we will just adopt the usual formulation given in (1.1).

At this point, one can wonder whether a notion of reducibility can be formulated for this class of problems. We positively answer this question with the following

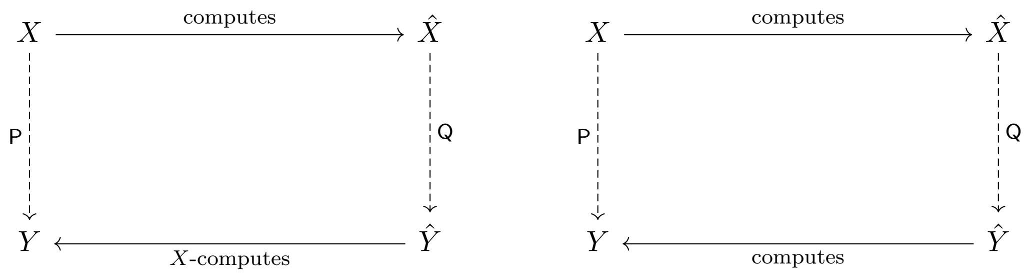

Figure 1.1. Diagram of computable reduction (left) and strong computable reduction (right) of a principle P to a principle Q.

definition, where $\oplus$ denotes the Turing join operation between sets, i.e., $X \oplus Y = \{2x \mid x \in X\} \cup \{2y+1 \mid y \in Y\}$.

Definition 1.2.7 (Computable reduction, $\leq_c$)**.** *Let* P, Q *be* Π^1_2*-principles. Then* Q *is* computably reducible *to* P *(denoted* $\mathsf{Q} \leq_c \mathsf{P}$*) if for every instance* X *of* Q *there exists an* X*-computable instance* $\widehat{X}$ *of* P *such that, if* $\widehat{Y}$ *is a solution to* P *for* $\widehat{X}$*, then there exists an* $X \oplus \widehat{Y}$*-computable solution to* Q *for* X*.*

Notice that the solution set $\hat{Y}$ in Definition 1.2.7 must be obtained by a single application of the principle P. Such a constraint is not required when proving the implication from Q to P over RCA_0 though: therefore, an implication over RCA_0 does not necessarily witnesses a computable reduction. For instance, in Proposition 3.2.10, we will prove that Hindman's Theorem is derivable over RCA_0 from our novel regressive version of the same theorem, but the argument does not establish a computable reduction, since it requires a further application of the latter principle (actually, of the weakest principle RT^1) to obtain a solution to HT. However, in that case, it is not guaranteed that a computable reduction (possibly based on a different argument) does not exist at all. Still, there are principles equivalent over RCA_0 that are provably not reducible to each other: for instance, our Corollary 3.3.11 states that $\lambda\mathsf{regHT}^{=2}[\mathsf{ap}]$ does not computably reduce RT^3_3, despite the two principles being equivalent to each other in RCA_0.
Moreover, the converse is also true, meaning that some computable reductions can not be formalized in RCA_0, as the former might use derivations that can not be inferred from the limited axioms of RCA_0.
Observe that, even if the notion of $\leq_c$ is based on algorithmic transformations, they may not be uniform, meaning that they can possibly be different for any instance and for any solution.
Finally, notice that the procedure transforming $\widehat{Y}$ into Y is allowed to access the instance of the original problem. However, one can decide not to endow the algorithm with such a capability, thus obtaining a stronger version of computable reducibility.

Definition 1.2.8 (Strong computable reduction, $\leq_{sc}$)**.** *Let* P, Q *be* Π^1_2*-principles. Then* Q *is* strongly computably reducible *to* P *(denoted* $\mathsf{Q} \leq_{sc} \mathsf{P}$*) if for every instance* X *of* Q *there exists an* X*-computable instance* $\widehat{X}$ *of* P *such that, if* $\widehat{Y}$ *is a solution to* P *for* $\widehat{X}$*, then there exists an* $\widehat{Y}$*-computable solution to* Q *for* X*.*

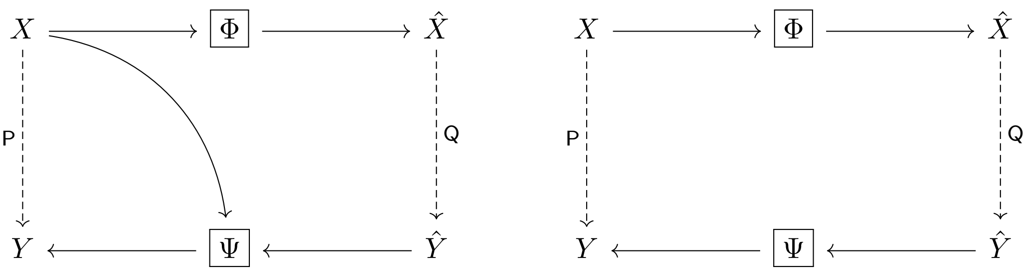

Figure 1.2. Diagram of Weihrauch reduction (left) and strong Weihrauch reduction (right) of a principle P to a principle Q, witnessed by Turing functionals Φ and Ψ.

Clearly, strong computable reducibility implies computable reducibility.

On the other hand, weaker notions of reducibility can be given. For instance, Hirschfeldt and Jockusch [HJ] formalized the ω-reducibility as follows, using the concept of ω-model, that is, a model of RCA_0 with domain ω.

Definition 1.2.9 (ω-reducibility, $\leq_\omega$)**.** *Let* P *and* Q *be* Π^1_2*-principles. Then* Q *is* ω-reducible *to* P *(denoted* $\mathsf{Q} \leq_\omega \mathsf{P}$*) if any* ω*-model satisfying* P *also satisfies* Q.

In the next chapters, we will focus exclusively on stronger kinds of reductions, so we do not provide further details about ω-reducibility; we just point out that ω-reducibility to a principle P is implied not only by computable reducibility to P, but even by the weaker notion of computable reducibility to m applications of P, for any $m \geq 1$ (see Proposition 4.6.9 in [DM]).

1.2.4 Weihrauch reductions

As we noted above, computable reducibility of a principle to another may be witnessed by different computable procedures for any instance of the original problem and for any solution. Then, it makes sense to require the two procedures to be uniform in the instances and in the solutions to the problem, hence giving a stronger notion of reducibility.
Despite the idea of uniform transformations was first formalized by Weihrauch [Wei1, Wei2], in the context of Reverse Mathematics such a concept is always defined according to the version given by Dorais, Dzhafarov, Hirst, Mileti and Shafer in [DDH$^+$], as the original formulation provided by Weihrauch used notions from computable analysis whose adoption is not particularly convenient in the field of Reverse Mathematics, especially when dealing with combinatorial principles. However, in [DDH$^+$] it is also shown how the two versions are equivalent to each other within our scope of interest.
Then, in the next Definition and throughout the rest of the thesis, we stick to the formulation given in [DDH$^+$] and based on the notion of Turing functional, that is, a computable function from sets to sets.

Definition 1.2.10 (Weihrauch reduction, $\leq_{\mathrm{W}}$)**.** *Let* P, Q *be* Π^1_2*-principles. Then* Q *is* Weihrauch reducible *to* P *(denoted* $\mathsf{Q} \leq_{\mathrm{W}} \mathsf{P}$*) if there exist Turing functionals* Φ

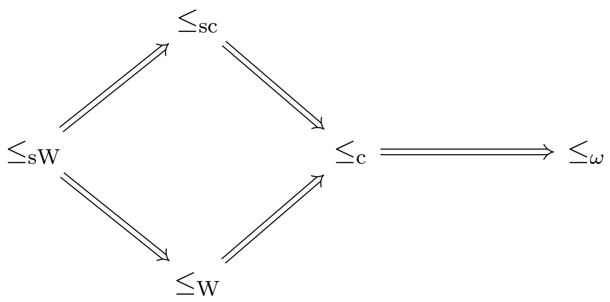

Figure 1.3. Diagram of implications between notions of reducibility. No other relations hold but the ones given by transitivity of implication.

and Ψ *such that, for every instance* X *of* Q*, we have that* $\Phi(X)$ *is an instance of* P*, and if* $\hat{Y}$ *is a solution to* P *for* $\Phi(X)$*, then* $\Psi(X \oplus \hat{Y})$ *is a solution to* Q *for* X*.*

While $\leq_{W}$ trivially implies $\leq_{c}$, it does not necessarily implies $\leq_{sc}$. As suggested by Hirschfeldt and Jockusch in [HJ], one can easily check that $\mathsf{Q} \leq_{W} \mathsf{P}$ while $\mathsf{Q} \not\leq_{sc} \mathsf{P}$ when we take P as $\forall X \exists Y (Y = \emptyset)$ and Q as $\forall X \exists Y (X = Y)$. The converse is also true, meaning that $\leq_{sc}$ does not entail $\leq_{W}$.

Similarly to computable reductions, one can require the functional Ψ to use an oracle just for $\hat{Y}$ and not for both $\hat{Y}$ and X.

Definition 1.2.11 (Strong Weihrauch reduction, $\leq_{sW}$)**.** *Let* P, Q *be* Π^1_2*-principles. Then* Q *is* strongly Weihrauch reducible *to* P *(denoted* $\mathsf{Q} \leq_{sW} \mathsf{P}$*) if there exist Turing functionals* Φ *and* Ψ *such that, for every instance* X *of* Q*, we have that* $\Phi(X)$ *is an instance of* P*, and if* $\hat{Y}$ *is a solution to* P *for* $\Phi(X)$*, then* $\Psi(\hat{Y})$ *is a solution to* Q *for* X*.*

Strong Weihrauch reductions imply all the previously discussed notions of reducibility. Figure 1.3 summarizes all the existing relations between the different kinds of reductions presented above. Hirschfeldt and Jockusch showed [HJ] that none of such implications can be reversed.

In the next chapters, we establish several reductions concerning various versions of Ramsey's Theorem and Hindman's Theorem: some of them use uniform functionals for transforming instances and solutions, hence witnessing Weihrauch reductions.

Chapter 2

Ramsey-type Theorems and Well-Ordering Principles

2.1 Ramsey's Theorem and its variations

Among combinatorial principles, Ramsey's Theorem is certainly one of the most studied in Proof Theory and Reverse Mathematics. Even if a finite version of such theorem does exist, here we are only interested in the infinitary version and in its several variations. All the definitions below are formalized in the language of Z_2. For $X \subseteq \mathbf{N}$ and $n \geq 1$ we denote by $[X]^n$ the set of subsets of X of cardinality n. For $k \geq 1$ we identify $\{0, \dots, k-1\}$ with k. Accordingly, for $S \subseteq \mathbf{N}$, $c : [S]^n \to k$ indicates a colouring of $[S]^n$ in k colours, and $f(x_1, \dots, x_n)$ means $f(\{x_1, \dots, x_n\})$ with $x_1 < \cdots < x_n$.

Definition 2.1.1 (Ramsey's Theorem)**.** *Let $n, k \geq 1$. We denote by RT^n_k the following principle. For all $c : [\mathbf{N}]^n \to k$ there exists an infinite set $H \subseteq \mathbf{N}$ such that c is constant on $[H]^n$. The set H is called* homogeneous *or* monochromatic *for c. Also, we use RT^n to denote $(\forall k \geq 1)\,\mathsf{RT}^n_k$ and RT to denote $(\forall n \geq 1)\,\mathsf{RT}^n$.*

This principle has been widely studied in Reverse Mathematics throughout the last few decades, hence several well-known results about its strength can be found in the literature. For instance, it is known that, for any $k \geq 1$, RT^1_k is provable in RCA_0, while RT^1 is equivalent to $\mathsf{B}\Sigma^0_2$ over RCA_0 [Hir1]. As noted in [DM], the latter result is quite surprising, for it shows the equivalence between a first order and a second order principle. Moreover, for $n \geq 3$ and $k \geq 2$, Simpson proved the following equivalence, starting from former results due to Jockusch [Joc].

Theorem 2.1.2 (Simpson [Sim2])**.** *Over RCA_0, for $n \geq 3$ and $k \geq 2$, RT^n_k is equivalent to ACA_0.*

The case $n = 2$, instead, is a bit anomalous, and even the strength of RT^2_2 has represented a longstanding question in Reverse Mathematics, until Liu [Liu] finally proved that RT^2_2 – which was already known to be weaker than ACA_0 [SS] but still not derivable in WKL_0 [Joc] – does not imply WKL_0: therefore, RT^2_2 and WKL_0 are independent from each other, meaning that the former does not imply the latter, and vice versa.

It is worth pointing out that, while the equivalence over RCA_0 between RT^n_j and RT^n_k can be proved for any positive j, k, n by a straightforward argument, the number of colours does matter when it comes to computable reductions: when $j < k$, neither $\mathsf{RT}^n_k \leq_{\mathrm{sc}} \mathsf{RT}^n_j$ nor $\mathsf{RT}^n_k \leq_{\mathrm{W}} \mathsf{RT}^n_j$ holds (see [DDH$^+$], [HJ], [BR]), while for $n \geq 2$, Patey [Pat] recently refined the corresponding (negative) results by proving that $\mathsf{RT}^n_k \not\leq_{\mathrm{c}} \mathsf{RT}^n_j$.

As for the "full" principle, i.e. RT, we have the following equivalence.

Theorem 2.1.3 (McAloon [McA])**.** *Over* RCA_0, RT *is equivalent to* ACA_0'.

The stronger system ACA_0^+, instead, is equivalent to Ramsey's Theorem for 2-colourings of exactly large sets, that is formally defined below.

Definition 2.1.4. *A set S is* exactly large *if $|S| = \min(S) + 1$. Moreover, we denote by $[X]^{!\omega}$ the collection of all exactly large sets $S \subseteq X$.*

Definition 2.1.5 (Ramsey's Theorem for exactly large sets)**.** *Let $k \geq 1$. We denote by $\mathsf{RT}^{!\omega}_k$ the following principle. For all $c : [\mathbf{N}]^{!\omega} \to k$ there exists an infinite set $H \subseteq \mathbf{N}$ such that c is constant on $[H]^{!\omega}$. The set H is called* homogeneous *or* monochromatic *for c.*

$\mathsf{RT}^{!\omega}_2$ has been originally proved by Pudlàk and Rödl [PR] and, independently, by Farmaki and Negrepontis [FN], while Carlucci and Zdanowski [CZ2] showed its equivalence with ACA_0^+.

Theorem 2.1.6 (Carlucci-Zdanowski [CZ2])**.** *Over* RCA_0, $\mathsf{RT}^{!\omega}_2$ *is equivalent to* ACA_0^+.

In Section 2.2 we will obtain the implications from ACA_0, ACA_0' and ACA_0^+ to the respective versions of Ramsey's Theorem via a combinatorial approach.

Besides those mentioned so far, there are several other variations of the original principle (e.g., Polarized Ramsey's Theorem, Stable Ramsey's Theorem, Rainbow Ramsey's Theorem, etc.), but they are out of the scope of this thesis. Many original results about some of these further versions can be found in [Pat], most of which are also discussed in [DM] together with many other pertinent results.

2.2 Ramsey's Theorem and Well-Ordering Principles

In this section, we relate well-ordering principles to different variations of Ramsey's Theorem, namely RT^3_2, RT and $\mathsf{RT}^{!\omega}_2$, hence giving a sort of ordinal analysis of such theorems, in the sense described in the last part of Sec. 1.1.4. As already mentioned in the previous sections, these Ramsey-theoretic statements have already been proved to be equivalent, respectively, to ACA_0, ACA_0' and ACA_0^+, and in turn these theories have already been associated with well-ordering principles of increasing strength. However, here we give a method for obtaining direct, combinatorial implications over RCA_0 and – at the same time – Weihrauch reductions between these three versions of Ramsey's Theorem and the corresponding well-ordering principles.

In order to apply the scheme of reductions to well-ordering principles, it is convenient to consider their contrapositive form, already presented and explained in Section 1.2, that is:

$$\forall \mathcal{X}(\neg \mathrm{WO}(f(\mathcal{X})) \to \neg \mathrm{WO}(\mathcal{X})),$$

where the instances of the problem are sequences witnessing that $f(\mathcal{X})$ is not well-ordered, while the solutions are sequences witnessing that $\mathcal{X}$ is not well-ordered. In all our proofs, moreover, the solution sequence is not only computable from the instance sequence, but it consists only of terms appearing as sub-terms in the instance sequence.

RT^3_2, ACA_0, and base-ω exponentiation. We start by giving a direct combinatorial argument showing that Ramsey's Theorem for triples and two colours implies $\mathsf{WOP}(\mathcal{X} \mapsto f(\mathcal{X}))$, where f is the operator of base-ω exponentiation defined below. The needed colouring is adapted from Loebl and Nešetril's [LN] combinatorial proof of the unprovability of Paris-Harrington principle in Peano Arithmetic. Our proof – that also establishes a Weihrauch reduction of the aforementioned well-ordering principle to RT^3_2 – is based on the same argument applied in [CZ1], where a direct implication from RT^3_3 to $\mathsf{WOP}(\mathcal{X} \mapsto f(\mathcal{X}))$ is already obtained. However, here we use just two colours: this minimizes the number of colours needed, hence proving the implication (and the reduction) from the usual principle RT^3_2 often used in Reverse Mathematics; moreover, this makes the proofs of Theorem 2.2.3 and Theorem 2.2.5 more readable.
When dealing with linear orders, we use the same notations as in [MM]. In particular, we define the operator $\boldsymbol{\omega}$ as follows. Given a linear order $\mathcal{X}$, the field of $\boldsymbol{\omega}^{\mathcal{X}}$ is the set of finite (possibly empty) sequences $\langle x_0, x_1, \ldots, x_k\rangle$ of elements of $\mathcal{X}$, where $x_0 \geq_{\mathcal{X}} x_1 \geq_{\mathcal{X}} \cdots \geq_{\mathcal{X}} x_k$. When $\mathcal{X}$ is an ordinal, the intended meaning of $\langle x_0, x_1, \ldots, x_k\rangle$ is $\omega^{x_0} + \cdots + \omega^{x_k}$. The order $\leq_{\boldsymbol{\omega}^{\mathcal{X}}}$ on $\boldsymbol{\omega}^{\mathcal{X}}$ is the lexicographic order. For any ordering $\mathcal{Y} = (Y, \leq_{\mathcal{Y}})$, we just use the symbol $\leq$ in place of $\leq_{\mathcal{Y}}$ when there is no risk of ambiguity.
If $\alpha \in \boldsymbol{\omega}^{\mathcal{X}}$ is the empty sequence (that we will occasionally denote by 0), then we set $lh(\alpha) = 0$, while we leave $e_i(\alpha)$ undefined for any i; otherwise, for any $\alpha = \langle x_0, x_1, \ldots, x_k\rangle \in \boldsymbol{\omega}^{\mathcal{X}}$, we define $lh(\alpha) = k+1$ and $e_i(\alpha) = x_i$ for any $i < lh(\alpha)$. In more intuitive terms, we sometimes call $lh(\alpha)$ the *length* of α and $e_i(\alpha)$ the i-th *exponent* of α. Also, for any $\beta = \langle y_0, y_1, \ldots, y_l\rangle \in \boldsymbol{\omega}^{\mathcal{X}}$ different from α, with $lh(\alpha) > 0$ and $lh(\beta) > 0$, we denote by $\Delta(\alpha, \beta)$ the least index at which α and β differ, i.e. the minimum i such that $e_i(\alpha) \neq e_i(\beta)$ if β is not an initial segment of α or vice versa, otherwise $\Delta(\alpha, \beta) = min(k, l) + 1$. If either α or β is the empty sequence, or $\alpha = \beta$, then we set $\Delta(\alpha, \beta) = 0$.

We show the following.

Theorem 2.2.1. *Over* RCA_0*,* RT^3_2 *implies* $\mathsf{WOP}(\mathcal{X} \mapsto \boldsymbol{\omega}^{\mathcal{X}})$*. Moreover, we have* $\mathsf{WOP}(\mathcal{X} \mapsto \boldsymbol{\omega}^{\mathcal{X}}) \leq_{\mathrm{W}} \mathsf{RT}^3_2$*.*

Proof. Assume RT^3_2 and, by way of contradiction, suppose $\neg\mathrm{WO}(\boldsymbol{\omega}^{\mathcal{X}})$. We show $\neg\mathrm{WO}(\mathcal{X})$. We define a σ-computable colouring $C^{(\sigma)} : [\mathbf{N}]^3 \to 2$ with an explicit sequence parameter σ of intended type $\sigma : \mathbf{N} \to field(\boldsymbol{\omega}^{\mathcal{X}})$ as follows:

$$C^{(\sigma)}(i,j,k) = \begin{cases} 0 & \text{if } \Delta(\sigma_i,\sigma_j) > \Delta(\sigma_j,\sigma_k), \\ 1 & \text{otherwise.} \end{cases}$$

Let $\alpha : \mathbf{N} \to field(\boldsymbol{\omega}^{\mathcal{X}})$ be an infinite descending sequence in $\boldsymbol{\omega}^{\mathcal{X}}$. Let $H = \{h_0 < h_1 < \cdots\}$ be an infinite $C^{(\alpha)}$-homogeneous set. Consider $(\beta_i)_{i\in\mathbf{N}}$, where $\beta_i = \alpha_{h_i}$. We reason by cases.

Case 1. The colour of $C^{(\alpha)}$ on $[H]^3$ is 0. Then the sequence:

$$\Delta(\beta_0,\beta_1) > \Delta(\beta_1,\beta_2) > \cdots$$

is strictly descending in $\mathbf{N}$, contradicting $\mathrm{WO}(\omega)$.

Case 2. The colour of $C^{(\alpha)}$ on $[H]^3$ is 1. Thus, for any $i < j < k$, we can derive $e_{\Delta(\beta_i,\beta_j)}(\beta_i) >_{\mathcal{X}} e_{\Delta(\beta_j,\beta_k)}(\beta_j)$, since $\beta_i >_{\boldsymbol{\omega}^{\mathcal{X}}} \beta_j >_{\boldsymbol{\omega}^{\mathcal{X}}} 0$ and $\Delta(\beta_i,\beta_j) \leq \Delta(\beta_j,\beta_k)$. Therefore we have:

$$e_{\Delta(\beta_0,\beta_1)}(\beta_0) >_{\mathcal{X}} e_{\Delta(\beta_1,\beta_2)}(\beta_1) >_{\mathcal{X}} \cdots$$

In other words, $\alpha' : \mathbf{N} \to \mathcal{X}$ defined by $i \mapsto e_{\Delta(\alpha_{h_i},\alpha_{h_{i+1}})}(\alpha_{h_i})$ is an infinite descending sequence in $\mathcal{X}$.

Notice that α' is obtained by first transforming the sequence α into $C^{(\alpha)}$, and then transforming the solution set H into the solution sequence α' (using information from α as well). Both transformations are uniform in the instance and in the solution, as required by the definition of Weihrauch reduction. □

We have the following immediate corollary, from Theorem 1.1.16 and Theorem 2.2.1.

Corollary 2.2.2. *Over* RCA_0, RT^3_2 *implies* ACA_0.

RT, ACA_0', and iterated base-ω exponentiation. We generalize the result from the previous paragraph to Ramsey's Theorem with internal universal quantification over all dimensions. This result is also in [CZ1], yet here we give a slightly different proof using an argument more clearly formalizable in RCA_0. Also, we observe that a Weihrauch reduction can be obtained.

Given a linear ordering $\mathcal{X}$, we define $\boldsymbol{\omega}^{\langle 0,\mathcal{X}\rangle} = \mathcal{X}$, and $\boldsymbol{\omega}^{\langle n+1,\mathcal{X}\rangle} = \boldsymbol{\omega}^{\boldsymbol{\omega}^{\langle n,\mathcal{X}\rangle}}$.

Theorem 2.2.3. *Over* RCA_0, RT *implies* $\forall h(\mathsf{WOP}(\mathcal{X} \to \boldsymbol{\omega}^{\langle h,\mathcal{X}\rangle}))$. *Moreover,* $\forall h(\mathsf{WOP}(\mathcal{X} \to \boldsymbol{\omega}^{\langle h,\mathcal{X}\rangle})) \leq_{\mathrm{W}} \mathsf{RT}$.

Proof. The case $h = 0$ is trivial, while the case $h = 1$ holds by Theorem 2.2.1. So, by way of contradiction, let $\alpha : \mathbf{N} \to field(\boldsymbol{\omega}^{\langle h,\mathcal{X}\rangle})$ be an infinite descending sequence in $\boldsymbol{\omega}^{\langle h,\mathcal{X}\rangle}$, for some $h \geq 2$ and for some well ordering $\mathcal{X}$. We show that we can construct an infinite descending sequence in $\mathcal{X}$, hence contradicting $\mathrm{WO}(\mathcal{X})$.

First, let $\sigma : \mathbf{N} \to field(\boldsymbol{\omega}^{\langle h,\mathcal{X}\rangle}) \cup \{\#\}$, where # is an additional symbol used for the sake of readability. For any $j \in \mathbf{N}$, any $I = \{i_0 < \cdots < i_k\}$ with $k \geq 1$, and any $n \leq k$, let us denote by $\sigma_j^{(n),I}$ the result of the n-th iteration in the process of extracting the comparing exponent of σ_j using indexes in I; formally, if $succ_I(j)$ is the least element in I greater than j (or 0 if no such element exists), then $\sigma_j^{(0),I} = \sigma_j$ and for $m < n$:

$$\sigma_j^{(m+1),I} = \begin{cases} e_{\Delta(\sigma_j^{(m),I}, \sigma_{succ_I(j)}^{(m),I})}(\sigma_j^{(m),I}) & \text{if } j \in \{i_0, \dots, i_{k-m-1}\} \text{ and} \\ & e_{\Delta(\sigma_j^{(m),I}, \sigma_{succ_I(j)}^{(m),I})}(\sigma_j^{(m),I}) \text{ exists}, \\ \# & \text{otherwise}, \end{cases}$$

We then denote by $\sigma^{(n),I}$ the sequence $(\sigma_j^{(n),I})_{j\in\mathbf{N}}$.

Now, let $C_1^{(\sigma)} : [\mathbf{N}]^3 \to 3$ be the same colouring defined in the proof of Theorem 2.2.1, with the additional property that $C_1^{(\sigma)}(i,j,k) = \#$ if at least one out of $\sigma_i, \sigma_j, \sigma_k$ is #. Also, if $I = \{i_0 < \dots < i_k\}$ with $k \geq 3$, for any $j = 0, \dots, k-3$, let:

$$v_j^{(\sigma),I} = (C_1^{(\sigma(j),I)}(i_0, i_1, i_2), C_1^{(\sigma(j),I)}(i_1, i_2, i_3), \dots, C_1^{(\sigma(j),I)}(i_{k-j-3}, i_{k-j-2}, i_{k-j-1})),$$

and

$$w_j^{(\sigma),I} = (C_1^{(\sigma(j),I)}(i_1, i_2, i_3), C_1^{(\sigma(j),I)}(i_2, i_3, i_4), \dots, C_1^{(\sigma(j),I)}(i_{k-j-2}, i_{k-j-1}, i_{k-j})).$$

Finally, we can define a colouring $C_h^{(\sigma)} : [\mathbf{N}]^{h+2} \to d(h)$, where d is a primitive recursive function (uniform in the proof) such that $d(h)$ is large enough to cover all the cases defined below. For $I = \{i_0 < \dots < i_{h+1}\}$, let us define:

$$C_h^{(\sigma)}(I) = \begin{cases} (v_0^{(\sigma),I},\ w_0^{(\sigma),I}) & \text{if } \neg(v_0^{(\sigma),I} = w_0^{(\sigma),I} = (1,\dots,1)) \\ (v_1^{(\sigma),I},\ w_1^{(\sigma),I}) & \text{if } \neg(v_1^{(\sigma),I} = w_1^{(\sigma),I} = (1,\dots,1)) \\ \dots & \dots \\ (v_{h-2}^{(\sigma),I},\ w_{h-2}^{(\sigma),I}) & \text{if } \neg(v_{h-2}^{(\sigma),I} = w_{h-2}^{(\sigma),I} = 1) \\ C_1^{(\sigma^{(h-1),I})}(i_0, i_1, i_2) & \text{otherwise} \end{cases}$$

where each case of $C_h^{(\sigma)}(I)$ is defined assuming that the conditions describing the previous cases do not hold.

By $\mathsf{RT}^{h+2}_{d(h)}$, let H be an infinite $C_h^{(\alpha)}$-homogeneous set. We show how to compute an $\mathcal{X}$-descending sequence given α and H. Let $\{s_0 < s_1 < \cdots\}$ be an enumeration of H in increasing order. We first show that the colour of H must be 1. To exclude the other cases, we argue as follows. Let $I = \{s_{i_0} < \cdots < s_{i_{h+1}}\}$ and $J = \{s_{i_1} < \cdots < s_{i_{h+2}}\}$ be in $[H]^{h+2}$, and let (v, w) be the colour of H, where $v = (v_0, \dots, v_l)$ and $w = (w_0, \dots, w_l)$, for some $l < h-1$. Also, let $l' = (h-2) - l$. Informally, l' is the number of times we "extract" the exponent of the elements of any $(h+2)$-tuple in $[H]^{h+2}$ in order to get the colour of the tuple; in other words, l' is such that $C_h^{(\alpha)}(I) = (v, w) = (v_{l'}^{(\alpha),I},\ w_{l'}^{(\alpha),I})$. Notice that l' is constant for every tuple in $[H]^{h+2}$, since it only depends on h and l, the latter being constant by homogeneity of H.

Case 1. The colour is (v, w), with $v \neq w$. This is easily seen to be impossible, since $C_h^{(\alpha)}(I) = (v, w)$ and, by definition of J, the first component of the colour of J is exactly w. But then $v = w$ should hold by homogeneity.

Case 2. The colour is (v, w) for some $v = w \neq (1, \dots, 1)$. We have three subcases.

Case 2.1. $v \neq (t, \dots, t)$, with $t \in \{\#, 0, 1\}$. Then, by hypothesis, we have $v_0 = w_0, v_1 = w_1, \dots, v_l = w_l$ and, by definition of the vectors v and w, we also have $w_0 = v_1, w_1 = v_2, \dots, w_{l-1} = v_l$. Thus we have $v_0 = v_1 = \cdots = v_l$. Contradiction.

Case 2.2. $v = (\#, \dots, \#)$. Then $v_{l'}^{(\alpha),I} = (\#, \dots, \#)$. However, no components of $v_{l'}^{(\alpha),I}$ can be $\#$, since α is a total strictly descending sequence and $v_j^{(\alpha),I} = (1, \dots, 1)$ for any $j < l'$, which implies that $e_{\Delta(\alpha_i^{(j),I}, \alpha_{succ_I(i)}^{(j),I})}(\alpha_i^{(j),I})$ exists for each $i \in I$. So this case can not occur.

Case 2.3. $v = (0, \dots, 0)$. Then the sequence

$$(\Delta(\alpha_{s_n}^{(l'),\{s_n,\dots,s_{n+l'}\}}, \alpha_{s_{n+1}}^{(l'),\{s_{n+1},\dots,s_{n+l'+1}\}}))_{n\in\mathbf{N}}$$

is strictly descending in $\mathbf{N}$, contradicting $\mathrm{WO}(\omega)$.

Since Case 1 and Case 2 cannot occur, for any $I = \{i_0 < \cdots < i_{h+1}\} \subset H$ we have

$$C_h^{(\alpha)}(I) = C_1^{(\alpha^{(h-1),I})}(i_0, i_1, i_2) \in \{\#, 0, 1\}.$$

We can discard colours # and 0 as in Cases 2.2 and 2.3, so the only possible colour for H is 1. Therefore, we can construct the sequence $(\alpha_{s_n}^{(h),\{s_n,\dots,s_{n+h}\}})_{n\in\mathbf{N}}$, which is an infinite descending sequence in $\mathcal{X}$. To see this, we can easily prove by Δ_0^0-induction on j that:

$$(\forall j \leq h)\left(\alpha_{s_n}^{(j),I} \ >_{\boldsymbol{\omega}\langle h-j,\mathcal{X}\rangle} \ \alpha_{s_{n+1}}^{(j),I'}\right)$$

where n is a free variable, $I = \{s_n, \dots, s_{n+h}\}$ and $I' = \{s_{n+1}, \dots, s_{n+h+1}\}$. The case $j = 0$ is trivial, while the case $j > 0$ is guaranteed by the fact that

$$\left(\alpha_{s_n}^{(j-1),I} \ >_{\boldsymbol{\omega}\langle h-j+1,\mathcal{X}\rangle} \ \alpha_{s_{n+1}}^{(j-1),I'}\right)$$

by induction hypothesis, and $C_1^{(\alpha^{(j-1),I})}(s_n, s_{n+1}, s_{n+2}) = 1$ due to the colour of H. Then, both $\alpha_{s_{n+1}}^{(j),I}$ and $\alpha_{s_n}^{(j),I'}$ exist, with the former being $<_{\boldsymbol{\omega}\langle h-j,\mathcal{X}\rangle}$ than the latter by the same argument used in the proof of Theorem 2.2.1.

The proof above witnesses a Weihrauch reduction: in particular, notice that both h and $d(h)$ can be computed uniformly from the original instance. □

The following corollary is immediate from Theorem 1.1.17 and Theorem 2.2.3.

Corollary 2.2.4. *Over* RCA_0, RT *implies* ACA_0'.

Large Ramsey's Theorem, ACA_0^+ and the ε-ordering. We now apply the proof technique from the previous paragraphs to $\mathsf{RT}^{!\omega}_2$, i.e. to the extension of Ramsey's Theorem to 2-colourings of exactly large sets. Such a result is conjectured in [CZ1], yet a proof is missing.
For the sake of readability of the argument below, here we slightly redefine the notion of largeness given in Definition 2.1.4, calling a set X *exactly large* if $|X| = \min(X)+3$. We do not reserve a specific notation for the principle $\mathsf{RT}^{!\omega}_2$ using this alternative definition of largeness, since it is equivalent over RCA_0 to the original version. Indeed, the implication from this novel version to the original one can be proved by a simple forgetful argument, while for the converse we can fix a colouring f of sets X such that $|X| = \min(X) + 3$, and then define:

$$g(i_0, i_1, \ldots, i_{i_0}) = \begin{cases} 0 & \text{if } i_0 < 2, \\ f(i_0 - 2, i_1 - 2, \ldots, i_{(i_0-2)+2} - 2) & \text{otherwise.} \end{cases}$$

Therefore, if $H = \{h_0 < h_1 < \cdots\}$ is any solution for g (of colour $c < 2$) to the original version of $\mathsf{RT}^{!\omega}_2$, then $H' = \{h'_0 < h'_1 < \cdots\} = \{h - 2 \mid h \in (H \setminus \{0, 1\})\}$ is a solution for f to this new version of the principle, as shown by the following chain of equalities holding for any $h'_{s_0} < h'_{s_1} < \cdots < h'_{h'_{s_0}+2}$ in H':

$$\begin{aligned} f(h'_{s_0}, h'_{s_1}, \ldots, h'_{h'_{s_0}+2}) &= f(h_{t_0} - 2, h_{t_1} - 2, \ldots, h_{h_{t_0}-2+2} - 2) = \\ &= f(h_{t_0} - 2, h_{t_1} - 2, \ldots, h_{h_{t_0}} - 2) = \\ &= g(h_{t_0}, h_{t_1}, \ldots, h_{h_{t_0}}) = c, \end{aligned}$$

where $h_{t_0} < h_{t_1} < \cdots < h_{h_{t_0}}$ are chosen in H. The argument above also witnesses a strong Weihrauch reduction.

Now, we are ready to show a direct implication (and a Weihrauch reduction) between $\mathsf{RT}^{!\omega}_2$ and $\mathsf{WOP}(\mathcal{X} \mapsto \boldsymbol{\varepsilon}_{\mathcal{X}})$, where $\boldsymbol{\varepsilon}$ is defined as in [MM]: hence, if $\mathcal{X}$ is a linear ordering, then the terms of $\boldsymbol{\varepsilon}_{\mathcal{X}}$ are 0, ε_x for any $x \in \mathcal{X}$, $\gamma_1 + \gamma_2$ for any γ_1 and γ_2 in $\boldsymbol{\varepsilon}_{\mathcal{X}}$, and ω^γ for any γ in $\boldsymbol{\varepsilon}_{\mathcal{X}}$. The normal form of each term and the relation $<_{\boldsymbol{\varepsilon}_{\mathcal{X}}}$ are defined together as follows. A term $\gamma = \gamma_0 + \cdots + \gamma_n$ in $\boldsymbol{\varepsilon}_{\mathcal{X}}$ is in normal form if either $n = 0$, or $\gamma_0 \geq_{\boldsymbol{\varepsilon}_{\mathcal{X}}} \gamma_1 \geq_{\boldsymbol{\varepsilon}_{\mathcal{X}}} \cdots \geq_{\boldsymbol{\varepsilon}_{\mathcal{X}}} \gamma_n >_{\boldsymbol{\varepsilon}_{\mathcal{X}}} 0$ and each γ_i is either ε_{x_i} for some $x_i \in \mathcal{X}$ or $\omega^{\delta'_i}$ for some δ'_i in normal form and different from ε_x for any $x \in \mathcal{X}$; also, for each $\gamma = \gamma_0 + \cdots + \gamma_n$ and $\delta = \delta_0 + \cdots + \delta_m$ in $\boldsymbol{\varepsilon}_{\mathcal{X}}$ both in normal form, $\gamma \leq_{\boldsymbol{\varepsilon}_{\mathcal{X}}} \delta$ if one of the following conditions holds:

- $\gamma = 0$;
- $\gamma = \varepsilon_x$ and, for some $y \geq_{\mathcal{X}} x$, ε_y occurs in δ;
- $\gamma = \omega^{\gamma'}$, $\delta_0 = \varepsilon_y$ and $\gamma' \leq_{\boldsymbol{\varepsilon}_{\mathcal{X}}} \varepsilon_y$;
- $\gamma = \omega^{\gamma'}$, $\delta_0 = \omega^{\delta'}$ and $\gamma' \leq_{\boldsymbol{\varepsilon}_{\mathcal{X}}} \delta'$;
- $n = 0$ and $\gamma_0 \leq_{\boldsymbol{\varepsilon}_{\mathcal{X}}} \delta_0$;
- $n > 0$ and $\gamma_0 = \delta_0$, $m > 0$ and $\gamma_1 + \cdots + \gamma_n \leq_{\boldsymbol{\varepsilon}_{\mathcal{X}}} \delta_1 + \cdots + \delta_m$.

Henceforth, we assume all the terms in $\boldsymbol{\varepsilon}_{\mathcal{X}}$ to be written in normal form. Also, for any $\gamma \in \boldsymbol{\varepsilon}_{\mathcal{X}}$ and $n \in \mathbf{N}$, let:

- γ_i be the i-th term of the normal form of γ if such term exists, that is, $\gamma_0 + \cdots + \gamma_n$ is the normal form of γ and $\gamma_i = 0$ whenever $i > n$;
- $e_i(\gamma)$ be the exponent of γ_i if such exponent does exist – that is, $e_i(\gamma) = \gamma'$ if $\gamma = \omega^{\gamma'}$ – otherwise $e_i(\gamma) = 0$;
- $b_i(\gamma) = min\{x \in \mathcal{X} \mid \gamma_i < \varepsilon_{x+1}\}$;
- $ht(\gamma) = 0$ if $\gamma < \varepsilon_0$ or $\gamma_0 = \varepsilon_{b_0(\gamma)}$, otherwise $ht(\gamma) = 1 + ht(\gamma')$, with $\gamma_0 = \omega^{\gamma'}$.

Informally, when $\gamma_n \geq \varepsilon_0$, $b_n(\gamma)$ is the largest $x \in \mathcal{X}$ such that ε_x appears in γ_n, while $ht(\gamma)$ is the maximum height at which $\varepsilon_{b_0(\gamma)}$ appears in γ_0 and therefore in γ. Moreover, if $\delta \in \boldsymbol{\varepsilon}_{\mathcal{X}}$, we indicate by $\Delta(\gamma, \delta)$ the index of the first term at which γ and δ differs – or 0 if $\gamma = \delta$.
Finally, notice that $\boldsymbol{\omega}$ and $\boldsymbol{\varepsilon}$ operators are compatible to each other in the sense of Lemma 2.6 in [MM], so we can refer to results from the previous paragraphs while dealing with $\boldsymbol{\varepsilon}$.

Theorem 2.2.5. *Over* RCA_0*,* $\mathsf{RT}^{!\omega}_2$ *implies* $\mathsf{WOP}(\mathcal{X} \to \boldsymbol{\varepsilon}_{\mathcal{X}})$*. Moreover, we have* $\mathsf{WOP}(\mathcal{X} \to \boldsymbol{\varepsilon}_{\mathcal{X}}) \leq_{\mathrm{W}} \mathsf{RT}^{!\omega}_2$*.*

Proof. Suppose WO($\mathcal{X}$) but $\neg$WO($\boldsymbol{\varepsilon}_{\mathcal{X}}$). Without loss of generality, we assume $0 \in \mathcal{X}$. We first define a colouring $C^{(\sigma)}_1 : [\mathbf{N}]^3 \to 5$ with an explicit sequence parameter σ of intended type $\sigma : \mathbf{N} \to field(\boldsymbol{\varepsilon}_{\mathcal{X}}) \cup \{\#\}$. Each case of $C^{(\sigma)}_1$ is defined assuming that the conditions describing the previous cases do not hold.

$$C^{(\sigma)}_1(i,j,k) = \begin{cases} \# & \text{if } \sigma_i = \# \ \vee \ \sigma_j = \# \ \vee \ \sigma_k = \# \\ 0 & \text{if } \Delta(\sigma_i, \sigma_j) > \Delta(\sigma_j, \sigma_k) \\ 1 & \text{if } (\sigma_i)_{\Delta(\sigma_i,\sigma_j)} \leq_{\boldsymbol{\varepsilon}_{\mathcal{X}}} \varepsilon_0 \\ 2 & \text{if } b_{\Delta(\sigma_i,\sigma_j)}(\sigma_i) >_{\mathcal{X}} b_{\Delta(\sigma_j,\sigma_k)}(\sigma_j) \\ 3 & \text{otherwise} \end{cases}$$

First, we claim that $C^{(\sigma)}_1$ has the following useful property.

Claim 1. *For any* $i < j < k$ *such that* $\sigma_i >_{\boldsymbol{\varepsilon}_{\mathcal{X}}} \sigma_j$*, if* $C^{(\sigma)}_1(i,j,k) = 3$ *then the comparing exponent of* α_i *with* α_j *does exist, formally* $e_{\Delta(\sigma_i,\sigma_j)}(\sigma_i) \neq 0$*.*

Proof. Assume otherwise. Since the colour is not #, all the elements in the rest of the proof are defined. Also, $C^{(\sigma)}_1(i,j,k) \neq 0$ implies $(\sigma_i)_{\Delta(\sigma_i,\sigma_j)} >_{\boldsymbol{\varepsilon}_{\mathcal{X}}} (\sigma_j)_{\Delta(\sigma_j,\sigma_k)}$. By $C^{(\sigma)}_1(i,j,k) \neq 1$ plus our assumption that the exponent of $(\sigma_i)_{\Delta(\sigma_i,\sigma_j)}$ is 0, instead, we can derive that $(\sigma_i)_{\Delta(\sigma_i,\sigma_j)} = \varepsilon_x$ for some $x >_{\mathcal{X}} 0$ in $\mathcal{X}$. Then, by $C^{(\sigma)}_1(i,j,k) \neq 2$, we have that $b_{\Delta(\sigma_j,\sigma_k)}(\sigma_j) = x$, so $(\sigma_j)_{\Delta(\sigma_j,\sigma_k)}$ must contain a term ε_x: however, such a term should appear at a lower height with respect to its height in $(\sigma_i)_{\Delta(\sigma_i,\sigma_j)}$, but this is impossible due to the latter being exactly ε_x. Thus, we have a contradiction, which concludes the proof of the claim.

Then, for any $h \geq 2$, we define $C^{(\sigma)}_h : [\mathbf{N}]^{h+2} \to d(h)$ as in Theorem 2.2.3, but using $C^{(\sigma)}_1$ as defined above: therefore, d must be adapted according to the larger range of

$C_1^{(\sigma)}$, and the condition $\neg(v_0^{(\sigma),I} = w_0^{(\sigma),I} = (1,\ldots,1))$ in the definition of $C_h^{(\sigma)}$ must be replaced with $\neg(v_0^{(\sigma),I} = w_0^{(\sigma),I} = (3,\ldots,3))$.

Lastly, we define a colouring $C^{(\sigma)} : [\mathbf{N}]^{!\omega} \to 2$ as follows:

$$C^{(\sigma)}(i_0,\ldots,i_{i_0+2}) = \begin{cases} 0 & \text{if } C_{i_0}^{(\sigma)}(i_1,\ldots,i_{i_0+2}) = 3 \\ 1 & \text{otherwise} \end{cases}$$

Now we are ready to prove that, by assuming $\neg\mathrm{WO}(\boldsymbol{\varepsilon}_{\mathcal{X}})$, we can construct an infinite descending sequence in $\mathcal{X}$, thus contradicting $\mathrm{WO}(\mathcal{X})$.
Let $\alpha : \mathbf{N} \to field(\boldsymbol{\varepsilon}_{\mathcal{X}})$ be an infinite descending sequence in $\boldsymbol{\varepsilon}_{\mathcal{X}}$ and, by $\mathsf{RT}_2^{!\omega}$, let $H = \{h_0 < h_1 < \ldots\}$ be an infinite $C^{(\alpha)}$-homogeneous set. Without loss of generality, we assume $h_0 > 1$ (this avoids some problems deriving from a different behaviour of $C_1^{(\alpha)}$ with respect to $C_h^{(\alpha)}$ for $h \geq 2$).
First, notice that the $C^{(\alpha)}$-colour of H is 0, otherwise, for any choice of $h \in H$, we could colour $H \setminus [0,h]$ using $C_h^{(\alpha)}$ and, by $\mathsf{RT}_{d(h)}^{h+2}$, we would obtain an infinite $C_h^{(\alpha)}$-homogeneous set whose colour is different from 3, hence contradicting the proof of Theorem 2.2.3, or rather its version adapted in order to manage the two additional colours of the base colouring $C_1^{(\alpha)}$. We can apply $\mathsf{RT}_{d(h)}^{h+2}$ since it is implied by $\mathsf{RT}_2^{!\omega}$.
Now, for the sake of readability of the argument below, we slightly redefine the notation used in Theorem 2.2.3. Let us denote by $\alpha_i^{(n),H}$, with $i \in H \setminus \{h_0\}$ and $n \leq prec_H(i) = max\{h \in H \mid h < i\}$, the result of the n-th iteration in the process of extracting the comparing exponent of α_i using indexes in H, i.e. for any $m < n$:

$$\alpha_i^{(0),H} = \alpha_i$$

$$\alpha_i^{(m+1),H} = e_{\Delta(\alpha_i^{(m),H},\alpha_{succ_H(i)}^{(m),H})}(\alpha_i^{(m),H})$$

Each term $\alpha_i^{(n),H}$ is well-defined, since the comparing exponent in the definition of $\alpha_i^{(m+1),H}$ does exist. This derives from the fact that $C_{prec_H(i)}^{(\alpha)}(i, h_{j_0},\ldots,h_{j_{prec_H(i)}}) = 3$ for any choice of $j_0 < j_1 < \cdots < j_{prec_H(i)}$ such that $h_{j_0} > i$, and from the following claim.

Claim 2. *For any $h \in H$, $m \leq h$ and $I = \{i_0 < \cdots < i_{n+1}\} \subseteq H \setminus [0,m]$, if $C_h^{(\alpha)}(I) = 3$ then the comparing exponent of $\alpha_{i_0}^{(m),H}$ with $\alpha_{i_1}^{(m),H}$ does exist, formally $e_{\Delta(\alpha_{i_0}^{(m),H},\alpha_{i_1}^{(m),H})}(\alpha_{i_0}^{(m),H}) \neq 0$.*

Proof. By $C_h^{(\alpha)}(I) = 3$ and by definition of $v_j^{(\alpha),I}$ we can derive $C_1^{(\alpha(j),I)}(i_0,i_1,i_2) = 3$ for any $j < h$. Then, by Claim 1 and the fact that α is decreasing, we can easily show that $e_{\Delta(\alpha_{i_0}^{(j),H},\alpha_{i_1}^{(j),H})}(\alpha_{i_0}^{(j),H}) \neq 0$ holds for any $j < h$, and therefore holds for m.

Now, we construct an $\mathcal{X}$-descending sequence τ using α and H. The idea is the following:

1. We put in τ the $<_{\mathcal{X}}$-largest x such that ε_x appears in α_0; clearly, ε_x must appear in the first term of α_0, at some height z.

2. Now, we know that any subsequent element of α cannot contain ε_x at height larger than z, since α is decreasing: therefore, if we switch to the sequence of the comparing exponents of height $z+1$ (as we have done in the proof of Theorem 2.2.3), we will only find ε-terms $\varepsilon_{x'}$ where $x' <_{\mathcal{X}} x$. The existence of such a sequence is guaranteed by Claim 2 above. Then, we can repeat the procedure from step 1 replacing α_0 with $\alpha_h^{(z+1),H}$.

We now formalize the idea above, yet starting our procedure from α_{h_1} – rather than from α_0 – just to avoid an abuse of notation: of course, the argument remains valid. Then, let us define the sequence $\tau : \mathbf{N} \to field(\mathcal{X})$ as follows:

$$\tau_i = b_0(\alpha_{n_i}^{(t_i),H})$$

for any $i \geq 0$, where $t_0 = h_0$, $n_0 = h_1$ and, for any $j \geq 0$,

$$t_{j+1} = t_j + ht(\alpha_{n_j}^{(t_j),H}) + 1$$

$$n_{j+1} = succ_H(min\{h \in H \mid h \geq t_{j+1}\})$$

Notice that each term τ_i is well-defined, since $t_i \leq prec_H(n_i)$, as required by the definition of $\alpha_{n_i}^{(t_i),H}$.
Finally, since the sequence $(\alpha_{n_{i+k}}^{(t_i),H})_{k\in\mathbf{N}}$ is decreasing by construction (cf. proof of Theorem 2.2.3), we have:

$$\tau_i = b_0(\alpha_{n_i}^{(t_i),H}) \overset{(*)}{=} b_0(\alpha_{n_{i+1}}^{(t_i),H}) \overset{(**)}{>_{\mathcal{X}}} b_0(\alpha_{n_{i+1}}^{(t_{i+1}),H}) = \tau_{i+1}$$

where $(*)$ is due to the $C_{t_i}^{(\alpha)}$-colour of tuples in $[H \setminus [0, n_i)]^{t_i+2}$ being different from 2, while $(**)$ is guaranteed by the choice of t_{i+1}: since ε_{τ_i} is the maximum ε-term in $\alpha_{n_i}^{(t_i),H}$ and it appears at height $h = ht(\alpha_{n_i}^{(t_i),H})$, then no terms ε_x with $x \geq \tau_i$ can appear at height $h' > h$ in $\alpha_{n_{i+1}}^{(t_i),H}$. Hence, no such terms can appear in $\alpha_{n_{i+1}}^{(t_{i+1}),H}$ either. So $(**)$ holds.

Therefore, τ is an infinite descending sequence in $\mathcal{X}$.
Moreover, both $C^{(\alpha)}$ and τ are obtained from computable transformations uniform in α and in $\alpha \oplus H$, respectively: then, they witness a Weihrauch reduction. ☐

We have the following corollary, yielding an alternative proof of Theorem 3.6 in [CZ2].

Corollary 2.2.6. *Over* RCA_0, $\mathsf{RT}_2^{!\omega}$ *implies* ACA_0^+.

Proof. From Theorem 1.1.18 and Theorem 2.2.5. ☐

$\mathsf{RT}_2^{!\omega}$ generalizes to a version of Ramsey's Theorem for bicolorings of exactly α-large sets [FN], and we conjecture that the method presented here can be extended to relate such general version of the theorem to the systems $\Pi^0_{\omega^\beta}$-CA_0 for every $\beta \in \omega^{\text{CK}}$ by using the characterization of the latter systems in terms of the well-ordering principles $(\mathsf{WOP}(\mathcal{X} \mapsto \boldsymbol{\varphi}(\beta, \mathcal{X}))$ [MM].

2.3 Hindman's Theorem and its variations

Hindman's Theorem is another relevant principle in Reverse Mathematics that is derived from Combinatorial Theory, and in particular from Ramsey's Theorem: indeed, it was first conjectured by Graham and Rothschild while listing a number of Ramsey-related open questions [GR]. It was then proved by Hindman [Hin1], whose argument – despite being more complex than other successive proofs – was later used by Blass, Hirst and Simpson [BHS] to give a first (and still unimproved) result concerning this principle from the point of view of Reverse Mathematics. Thereafter, proof-theoretic results about Hindman's Theorem have prospered in the literature and, similarly to Ramsey's Theorem, several variations of the original principle have been formalized and studied, some of which are presented below. Then, in the next section, we formulate a novel version of Hindman's Theorem with the interesting property of being in some sense "symmetric" with the analogous variation of Ramsey's Theorem.
We start by giving the statement of the original Hindman's Theorem, which requires the following Definition. As in the previous sections, we use the language of Z_2.

Definition 2.3.1 (FS)**.** *For any $X \subseteq \mathbf{N}$, we denote by $\mathrm{FS}(X)$ the set of all non-empty finite sums of distinct elements in X.*

Definition 2.3.2 (Hindman's Theorem)**.** *Let $k \geq 1$. We denote by HT_k the following principle. For all $c : \mathbf{N} \to k$ there exists an infinite set $H \subseteq \mathbf{N}$ such that c is constant on $\mathrm{FS}(H)$. We denote by HT the principle $(\forall k \geq 1)\, \mathsf{HT}_k$.*

For technical convenience and without loss of generality, Hindman's Theorem (as well as any of its variations) is sometimes stated replacing $\mathbf{N}$ with $\mathbf{N}^+ = \mathbf{N} \setminus \{0\}$. Also, this principle is often stated using an equivalent formulation based on the notions of finite union and block sequence, that are formalized in the next Definitions. We use $\mathrm{FIN}(X)$ to denote the set of all non-empty finite subsets of $X \subseteq \mathbf{N}$.

Definition 2.3.3 (FU)**.** *Let $\mathcal{B} = (B)_{i\in\mathbf{N}}$ be a sequence of non-empty finite sets. Then we denote by $\mathrm{FU}(\mathcal{B})$ the set of non-empty finite unions of $\mathcal{B}$.*

Definition 2.3.4 (Block sequence)**.** *Let $\mathcal{B} = (B_i)_{i\in\mathbf{N}}$ be a sequence of non-empty finite sets. Then $\mathcal{B}$ satisfies the* block condition *if, for any i, $\max(B_i) < \min(B_{i+1})$ – in short, $B_i < B_{i+1}$. In that case, we call $\mathcal{B}$ a* block sequence.

Using the previous Definitions, we can then restate Hindman's Theorem (which in that case is called *Finite Union Theorem*) as follows.

Definition 2.3.5 (Finite Union Theorem)**.** *Let $k \geq 1$. We denote by FUT_k the following principle. For all $c : \mathrm{FIN}(\mathbf{N}) \to k$ there exists an infinite block sequence $\mathcal{B}$ such that c is constant on $\mathrm{FU}(\mathcal{B})$. We denote by FUT the principle $(\forall k \geq 1)\, \mathsf{FUT}_k$.*

The block condition is needed by FUT in order to be equivalent (over RCA_0) to HT: when this condition is dropped, Hindman's Finite Unions Theorem becomes much weaker (in particular, provable in RCA_0) as shown by Hirst [Hir3]. We can then introduce the corresponding property in the finite sums setting. This property is already implicit in Hindman's original proof [Hin2] and is usually called *apartness*. Let $n \in \mathbf{N}^+$. If $n = 2^{t_1} + \cdots + 2^{t_p}$ with $0 \leq t_1 < \cdots < t_p$ and $p \geq 1$, then we set $\lambda(n) = t_1$ and $\mu(n) = t_p$ (the notation is from [BHS]). We set $\lambda(0) = \mu(0) = 0$.

Definition 2.3.6 (Apartness Condition)**.** *A set X satisfies the* apartness condition *if for all $x, x' \in X$ such that $x < x'$, we have $\mu(x) < \lambda(x')$. If X satisfies the apartness condition we say that X is* apart*.*

If P is a Hindman-type principle, we denote by P *with apartness* or $\mathsf{P}[\mathsf{ap}]$ the principle P with the apartness condition imposed on the solution set.
In Hindman's original proof the apartness condition is ensured by a simple counting argument (Lemma 2.2 in [Hin1]) on any solution to the Finite Sums Theorem, i.e., on any infinite $H \subseteq \mathbf{N}$ such that $\mathrm{FS}(H)$ is monochromatic (Lemma 2.3 in [Hin1]). As noted in [BHS], the proof shows that a solution satisfying the apartness condition can be obtained computably in any such solution. In the Reverse Mathematics setting, one needs to be slightly more careful to establish that HT implies HT *with apartness* over RCA_0.
We first check that Lemma 2.2 in [Hin1] holds in RCA_0.

Lemma 2.3.7. *The following is provable in RCA_0: For all ℓ, for all k, for all finite sets X, if X has cardinality 2^k and is such that $\lambda(x) = \ell$ for all $x \in X$, then there exists $Y \subseteq X$ such that $\lambda(\sum_{y \in Y} y) \geq \ell + k$.*

Proof. The Lemma is established by a straightforward induction on k. We give the details for completeness.
For the base case, let $k = 0$ and let $X = \{x\}$ be a finite set of cardinality 2^0 such that $\lambda(x) = \ell$. Obviously choosing $Y = X$ gives the desired solution.
For the inductive step, let $k \geq 0$ and let X be a set of cardinality 2^{k+1} such that for all $x \in X$ we have $\lambda(x) = \ell$. Let A and B be two disjoint subsets of X each of cardinality 2^k. By inductive hypothesis there exists $A' \subseteq A$ such that $\lambda(\sum_{a \in A'} a) \geq \ell + k$ and there exists $B' \subseteq B$ such that $\lambda(\sum_{b \in B'} b) \geq \ell + k$. We distinguish the following cases. If $\lambda(\sum_{a \in A'} a) = \ell + k$ and $\lambda(\sum_{b \in B'} b) = \ell + k$ then $\lambda(\sum_{c \in A' \cup B'} c) \geq \ell + k + 1$. If either $\lambda(\sum_{a \in A'} a) > \ell + k$ or $\lambda(\sum_{b \in B'} b) > \ell + k$ then we are done.
The argument can be carried out in RCA_0 since quantification over finite sets formally means quantification over their numerical codes and the set Y is a finite subset of the finite set X, so that the existential quantifier over Y is bounded. The induction predicate is then Π^0_1, and Π^0_1-induction holds in RCA_0 (Cor. II.3.10 in [Sim2]). ◻

The following Lemma appears as Lemma 9.9.6 in Dzhafarov and Mummert [DM], but the proof proposed there uses the wrong assumption that the element denoted by x_2 is in $\mathrm{FS}(I)$. Then we give an alternative argument, using Lemma 2.3.7.

Lemma 2.3.8. *The following is provable in $\mathsf{RCA}_0 + \mathsf{RT}^1$: For every $m \in \mathbf{N}$ and every infinite $I \subseteq \mathbf{N}$, there exists $x \in \mathrm{FS}(I)$ with $\lambda(x) \geq m$.*

Proof. Fix m and I and suppose that every $x \in \mathrm{FS}(I)$ satisfies $\lambda(x) < m$. In particular this implies that every $x \in I$ satisfies $\lambda(x) < m$, since $I \subseteq \mathrm{FS}(I)$. By RT^1 there exists an $\ell < m$ and an infinite set $J \subseteq I$ such that $\lambda(x) = \ell$ for all $x \in J$.
Since $\ell < m$ there exists k such that $\ell + k = m$. Pick a subset $X \subseteq J$ of cardinality 2^k. Then by Lemma 2.3.7 there exists a $Y \subseteq X$ such that $\lambda(\sum_{y \in Y} y) \geq \ell + k = m$. This contradicts the hypothesis that $\lambda(x) < m$ for all $x \in \mathrm{FS}(I)$. ◻

As a corollary one obtains the following Proposition, which will be used to show that HT self-strengthens to $\mathsf{HT}[\mathsf{ap}]$ over RCA_0.

Proposition 2.3.9.

1. *The following is provable in* $\mathsf{RCA}_0 + \mathsf{RT}^1$*: For every infinite set* $I \subseteq \mathbf{N}$*, there is an infinite set* J *such that* J *is apart and* $\mathrm{FS}(J) \subseteq \mathrm{FS}(I)$.
2. *For all infinite set* $I \subseteq \omega$ *there exists an infinite set* $J \subseteq \omega$ *computable in* I *such that* J *is apart and* $\mathrm{FS}(J) \subseteq \mathrm{FS}(I)$.

Proof. Define a sequence of elements $x_0 < x_1 < \cdots$ in $\mathrm{FS}(I)$ recursively as follows. Let $x_0 = \min(I)$ and, for $i \in \mathbf{N}$, let x_{i+1} be the least element of $\mathrm{FS}(I \setminus [0, x_i])$ such that $\lambda(x_{i+1}) > \mu(x_i)$. The existence of x_{i+1} follows from Lemma 2.3.8. Let $J = \{x_i : i \in \mathbf{N}\}$. By construction J is apart and $\mathrm{FS}(J) \subseteq \mathrm{FS}(I)$. □

Proposition 2.3.9 is close in both statement and proof to Corollary 9.9.8 in [DM] but ensures $\mathrm{FS}(J) \subseteq \mathrm{FS}(I)$ rather than $J \subseteq \mathrm{FS}(I)$ as in [DM]. This stronger condition is indeed needed in the proof of the following corollary, which appears as Theorem 9.9.9 in [DM]. The proof of the latter contains an error, for it wrongly claims that $J \subseteq \mathrm{FS}(I)$ implies $\mathrm{FS}(J) \subseteq \mathrm{FS}(I)$.

Corollary 2.3.10. HT *implies* $\mathsf{HT}[\mathsf{ap}]$ *over* RCA_0*. Moreover* $\mathsf{HT} \geq_{\mathrm{sW}} \mathsf{HT}[\mathsf{ap}]$.

Proof. From Proposition 2.3.9 and the fact that HT trivially implies RT^1 over RCA_0. More precisely, let $c : \mathbf{N} \to k$. Let I be a solution to HT for c. By Proposition 2.3.9 there exists an infinite apart set J such that $\mathrm{FS}(J) \subseteq \mathrm{FS}(I)$. Then J is a solution to $\mathsf{HT}[\mathsf{ap}]$ for c.
It is clear from the proof of Proposition 2.3.9 that there is a Turing functional that computes J from I uniformly. This is sufficient to establish the claimed strong Weihrauch reduction. □

Despite the fact that – as we have just proved – HT self-strengthens to $\mathsf{HT}[\mathsf{ap}]$, the exact proof-theoretic strength of Hindman's Theorem is unsettled to date: the best known lower bound and upper bound for this principle are still the ones provided by the seminal work of Blass, Hirst and Simpson [BHS], where HT is shown to be between ACA_0 and ACA_0^+.

Over the years, a number of restrictions of the original principle have been formulated, whose provability in systems below ACA_0^+ is mostly unknown as well, even if they appear to be weaker than Hindman's Theorem. Among these alternative versions, substantial attention has been devoted by many authors to restrictions of HT requiring the monochromaticity condition only for particular families of finite sums, mainly obtained by limiting the number of terms in the sums. We then introduce the needed terminology. For $n \geq 1$ and $X \subseteq \mathbf{N}$, we denote by $\mathrm{FS}^{\leq n}(X)$ the set of all non-empty sums of at most n distinct elements of X, while we denote by $\mathrm{FS}^{=n}(X)$ the set of all sums of exactly n distinct elements of X.
Using these definitions, two natural families of restrictions of Hindman's Theorem are obtained.

Definition 2.3.11 (Bounded Hindman's Theorems)**.** *Let $n, k \geq 1$. We denote by $\mathsf{HT}^{\leq n}_k$ (resp. $\mathsf{HT}^{=n}_k$) the following principle. For every $c : \mathbf{N} \to k$ there exists an infinite set $H \subseteq \mathbf{N}$ such that $\mathrm{FS}^{\leq n}(H)$ (resp. $\mathrm{FS}^{=n}(H)$) is monochromatic for c. We use $\mathsf{HT}^{\leq n}$ (resp. $\mathsf{HT}^{=n}$) to denote $(\forall k \geq 1)\, \mathsf{HT}^{\leq n}_k$ (resp. $(\forall k \geq 1)\, \mathsf{HT}^{=n}_k$).*

Note that $\mathsf{HT}^{\leq 1}_k$, $\mathsf{HT}^{=1}_k$ and RT^1_k are all equivalent and strongly Weihrauch inter-reducible (by identity). For all the other cases, i.e. $\mathsf{HT}^{\leq n}_k$ and $\mathsf{HT}^{=n}_k$ with $n \geq 2$ and $k \geq 1$, a good overview is given in [CKLZ] and in [Car2]. One remarkable point is that, even if $\mathsf{HT}^{\leq n}_k$ with $n, k \geq 2$ seems to be weaker than HT, the best known upper bound for the former principle is still ACA^+_0, that is, the same upper bound of the original Hindman's Theorem. It is also worth pointing out that Bounded Hindman's Theorems are not known to self-strengthen to their own versions with apartness: the proof of Corollary 2.3.10, indeed, strongly relies on the use of sums of arbitrary (finite) length. Finally, a noteworthy result about Bounded Hindman's Theorem is that, over RCA_0, $\mathsf{HT}^{=n}_k[\mathsf{ap}]$ is equivalent to RT^n_k for any $n \geq 3$ and any $k \geq 2$ [CKLZ], and therefore is equivalent to ACA_0.

2.4 Hindman's Theorem and Well-Ordering Principles

We establish a connection between Hindman-type theorems and well-ordering principles, along the lines of our previous results. For the sake of readability, here we adopt the alternative formulation of Hindman's Theorem based on the concept of finite unions (see Definition 2.3.5): hence, we use $\mathsf{FUT}^{=n}_k$ in place of $\mathsf{HT}^{=n}_k[\mathsf{ap}]$, with the two principles being equivalent over RCA_0 and strongly Weihrauch-reducible to each other [CKLZ, Proposition 2.5].
Since $\mathsf{FUT}^{=3}_2$ implies ACA_0 over RCA_0, (see [CKLZ]), by Theorem 1.1.16 we know that $\mathsf{FUT}^{=3}_2$ implies $\mathsf{WOP}(\mathcal{X} \mapsto \boldsymbol{\omega}^{\mathcal{X}})$ over RCA_0. We give a new proof of this implication (actually, a slightly more general version of it) by a direct argument that furthermore establishes a Weihrauch reduction. Its interest also lies in the connection between Hindman's Theorem and principles related to transfinite ordinals.

We proceed as follows. Let $\mathcal{X}$ be a linear ordering and let $\alpha = (\alpha_i)_{i \in \mathbf{N}}$ be an infinite decreasing sequence in $\boldsymbol{\omega}^{\mathcal{X}}$. We show, using $\mathsf{FUT}^{=n}_k$ for $n \geq 3$ and $k \geq 2$, that there exists an infinite decreasing sequence in $\mathcal{X}$. The proof uses ideas from the proof of $\mathsf{FUT}^{=3}_2 \to \mathsf{ACA}_0$ [CKLZ, Proposition 3.1] adapted to the present context, based on the following analogy between deciding the Halting Set K and computing an infinite descending sequence in $\mathcal{X}$. Given an enumeration of K and a number n, RCA_0 knows that there is an ℓ such that all numbers in K below n appear within ℓ steps of the enumeration, but is not able to compute this ℓ. Similarly, given an element α_i in the infinite decreasing sequence α in $\boldsymbol{\omega}^{\mathcal{X}}$, RCA_0 knows that there is an ℓ such that if an exponent of α_i ever decreases – meaning that there exist j, p such that $e_p(\alpha_i) >_{\mathcal{X}} e_p(\alpha_{i+j})$ – it will do so by the ℓ-th element of the infinite descending sequence, but it is unable to compute such an ℓ. Therefore, we can not computably locate the leftmost "decreasing" exponent of γ, and the point in the sequence where it first decreases. However, an appropriately designed colouring will ensure that the information about such an ℓ can be read off from the elements of a solution to Hindman's Theorem.

We start with the following simple Lemma, stating that any element of an infinite descending sequence in $\boldsymbol{\omega}^{\mathcal{X}}$ contains an exponent that eventually decreases.

Lemma 2.4.1. *The following is provable in* RCA_0*: if* $\alpha = (\alpha_i)_{i \in \mathbf{N}}$ *is an infinite descending sequence in* $\boldsymbol{\omega}^{\mathcal{X}}$*, then*

$$(\forall n)\,(\exists n')\,(\exists m < lh(\alpha_n))\,\big[(n' > n)\,\wedge\,(e_m(\alpha_n) >_{\mathcal{X}} e_m(\alpha_{n'}))\big].$$

Proof. Assume by way of contradiction that none of the exponents of α_n, for some n, will ever decrease. By definition of $<_{\boldsymbol{\omega}^{\mathcal{X}}}$, for any distinct $\gamma, \delta \in \boldsymbol{\omega}^{\mathcal{X}}$, we have $\gamma <_{\boldsymbol{\omega}^{\mathcal{X}}} \delta$ if and only if either (1.) γ is a proper initial segment of δ, or (2.) there exists $m < lh(\delta)$ such that $e_m(\gamma) <_{\mathcal{X}} e_m(\delta)$ and $e_{m'}(\gamma) = e_{m'}(\delta)$ for each $m' < m$. Then we can prove by Δ^0_1-induction the following claim:

$$\forall p\ (p \geq n \rightarrow (\alpha_{p+1} \text{ is a proper initial segment of both } \alpha_p \text{ and } \alpha_n)).$$

The case $p = n$ is trivial, since $\alpha_n >_{\mathcal{X}} \alpha_{n+1}$ and (2.) cannot hold by assumption. For $p > n$, by induction hypothesis we know that α_p is a proper initial segment of α_n. Since $\alpha_{p+1} >_{\mathcal{X}} \alpha_p$, α_{p+1} must be a proper initial segment of α_p, otherwise the leftmost exponent differing between α_{p+1} and α_p – i.e. the exponent of α_{p+1} with index m witnessing (2.) – would contradict our assumption, for we would have $m < lh(\alpha_p)$ and $e_m(\alpha_{p+1}) <_{\mathcal{X}} e_m(\alpha_p) = e_m(\alpha_n)$.
So α_{p+1} must be a proper initial segment of α_p and, by induction hypothesis, it must be a proper initial segment of α_n as well.
The claim above implies that:

$$\forall p\ (p \geq n \rightarrow lh(\alpha_p) > lh(\alpha_{p+1}))$$

hence contradicting $\mathrm{WO}(\omega)$. This concludes the proof. □

Theorem 2.4.2. *Let* $n \geq 3, k \geq 2$*. Over* RCA_0*,* $\mathsf{FUT}^{=n}_k$ *implies* $\mathsf{WOP}(\mathcal{X} \rightarrow \boldsymbol{\omega}^{\mathcal{X}})$*. Moreover,* $\mathsf{WOP}(\mathcal{X} \rightarrow \boldsymbol{\omega}^{\mathcal{X}}) \leq_{\mathrm{W}} \mathsf{FUT}^{=n}_k$*.*

Proof. Assume by way of contradiction $\neg\mathrm{WO}(\boldsymbol{\omega}^{\mathcal{X}})$, and let $\alpha = (\alpha_n)_{n \in \mathbf{N}}$ be an infinite descending sequence in $\boldsymbol{\omega}^{\mathcal{X}}$. For this proof, it is convenient to use the α-computable sequence β of all the exponents of the terms α_n, enumerated in order of "appearance", i.e. $\beta = \langle e_0(\alpha_0), e_1(\alpha_0), \dots, e_{lh(\alpha_0)-1}(\alpha_0), e_0(\alpha_1), e_1(\alpha_1), \dots \rangle$. Formally we construct such sequence by first defining $\theta : \mathbf{N} \times \mathbf{N} \rightarrow \mathbf{N}$ as follows: $\theta(n, m) = m + \sum_{k<n} lh(\alpha_k)$ whenever $m < lh(\alpha_n)$, otherwise it is undefined. The partial function θ is clearly bijective (meaning that it is a bijection between its domain of definition and its codomain), hence its inverse is a well-defined injective total function. We can then fix functions $t : \mathbf{N} \rightarrow \mathbf{N}$ and $p : \mathbf{N} \rightarrow \mathbf{N}$ such that for each $n \in \mathbf{N}$ we have $\theta(t(n), p(n)) = n$. The sequence $(\beta_h)_{h \in \mathbf{N}}$ of exponents of terms in α is then defined by setting $\beta_h = \alpha_{t(h),p(h)}$. Intuitively, $t(h)$ is the element of α from which β_h has been "extracted", while $p(h)$ is the position of the exponent β_h within $\alpha_{t(h)}$.
We call i *decreasible* if there exists $j > i$ such that $p(j) = p(i)$ and $\beta_i > \beta_j$. In that case, we say that j *decreases* i and that j is a *decreaser* of i. Using this terminology, Lemma 2.4.1 states that each element of α contains at least one decreasible exponent.

Now suppose that $f : \mathbf{N} \to \mathbf{N}$ is a function with the following property.

Property P: For all $i \in \mathbf{N}$, if i is decreasible, then it is decreased by some $j \leq f(i)$.

We first show that given such an f we can compute (in f and β) an infinite descending sequence $(\sigma_i)_{i\in\mathbf{N}}$ in $\mathcal{X}$ as follows.

Step 0. Let i_0 be the least decreasible index of β, and let j_0 be the least decreaser of i_0. By Lemma 2.4.1, β_{i_0} must be an exponent of α_0, i.e. $t(i_0) = 0$, so we can find i_0 by just taking the least decreasible $i^* < lh(\alpha_0)$. Notice that we can decide whether i^* is decreasible by just inspecting β up to the index $f(i^*)$, since f has the property P.
We set $\sigma_0 = \beta_{j_0}$ and observe that $p(i) \geq p(i_0)$ for each decreasible $i > i_0$. Suppose otherwise as witnessed by i^*, and let j^* be the least decreaser of i^*. By definition of decreaser, $p(j^*) = p(i^*) < p(i_0)$ and $\beta_{i^*} > \beta_{j^*}$. However i_0 is the least decreasible β-index of exponents of α_0, hence the β-index of $e_{p(i^*)}(\alpha_0)$ will never decrease due to $p(i^*) < p(i_0)$, so $\beta_{i^*} = \beta_z$, where $z = \theta(0, p(i^*))$. Therefore, $\beta_z > \beta_{j^*}$ and $p(z) = p(i^*) = p(j^*)$, but in that case z would be decreasible (by j^*) and $p(z) < p(i_0)$, which implies $z < i_0$ since $t(z) = t(i_0) = 0$, thus contradicting the minimality of i_0.

Step $s+1$. Suppose i_s, j_s, σ_s are defined, $(\sigma_t)_{t\leq s}$ is decreasing in $\mathcal{X}$ and $p(i) \geq p(i_s)$ for each decreasible $i > i_s$.
Let i_{s+1} be the least decreasible index of β larger than or equal to j_s, and let j_{s+1} be the least decreaser of i_{s+1}. By Lemma 2.4.1 and the fact that no decreasible $i > i_s$ can have $p(i) < p(i_s) = p(j_s)$, $\beta_{i_{s+1}}$ must be an exponent of $\alpha_{t(j_s)}$, namely the leftmost whose β-index is decreasible. So we can find i_{s+1} by just taking the least decreasible $i^* \in [j_s,\ j_s + lh(\alpha_{t(j_s)}) - p(j_s))$. Notice that we can decide whether i^* is decreasible by just inspecting β up to the index $f(i^*)$, since f has the property P.
We set $\sigma_{s+1} = \beta_{j_{s+1}}$. As we noted above, $\beta_{i_{s+1}}$ must be either β_{j_s} or an exponent of $\alpha_{t(j_s)}$ on the right of β_{j_s}, i.e. $t(i_{s+1}) = t(j_s)$ and $p(i_{s+1}) \geq p(j_s)$, so $\beta_{j_s} \geq \beta_{i_{s+1}}$. Then, $\sigma_s > \sigma_{s+1}$ because $\sigma_s = \beta_{j_s} \geq \beta_{i_{s+1}} > \beta_{j_{s+1}} = \sigma_{s+1}$. Finally, we observe that the last part of the inductive invariant is guaranteed as well, since $p(i) \geq p(i_{s+1})$ for each decreasible $i > i_{s+1}$. Suppose otherwise as witnessed by i^*, and let j^* be the least decreaser of i^*. By definition of decreaser, $p(j^*) = p(i^*) < p(i_{s+1})$ and $\beta_{i^*} > \beta_{j^*}$. However $\beta_{i_{s+1}}$ is the leftmost exponent of $\alpha_{t(j_s)}$ whose β-index is decreasible, then $\beta_{i^*} = \beta_z$, where $z = \theta(t(j_s), p(i^*))$. Hence, $\beta_z > \beta_{j^*}$, but in that case z would be decreasible (by j^*) and $p(z) = p(i^*) < p(i_{s+1})$, which implies $z < i_{s+1}$ since $t(z) = t(j_s) = t(i_{s+1})$, thus contradicting the minimality of i_{s+1}.

We now show how to obtain a function satisfying the property P from a solution of $\mathsf{FUT}^{=n}_k$ for a suitable colouring.
First, we define a number $j \in [0, r]$ *important in* $S = \{n_0 < \cdots < n_r\}$ if the following condition holds:

$$(\exists i < n_0)\left[(\exists i' \in [n_{j-1}, n_j))\,(i' \text{ decreases } i) \wedge (\neg\exists i'' < n_{j-1})\,(i'' \text{ decreases } i)\right]$$

where we set $n_{-1} = 0$. Informally, j is important if in $[n_{j-1}, n_j)$ appears the first decreaser of some element less than n_0.

Then, let $g : \mathrm{FIN}(\mathbf{N}^+) \to k$ as follows:

$$g(S) = card\{j \mid j \text{ is important in } S\} \bmod k.$$

By $\mathsf{FUT}_k^{=n}$ let $\mathcal{B} = \{B_0 < B_1 < B_2 < \ldots\}$ be an infinite block sequence such that $FU^{=n}(\mathcal{B})$ is monochromatic under g, and let $c < k$ be the colour of $\mathcal{B}$.

Claim 3. *Given $S_0 < \ldots < S_{n-3}$ in $\mathcal{B}$, there exists $S \in \mathcal{B}$ such that $S_{n-3} < S$ and $g(S_0 \cup \ldots \cup S_{n-3} \cup S) = c$.*

Fix $S_0 < \ldots < S_{n-3}$ in $\mathcal{B}$ and let ℓ be the actual upper bound of the minimal indexes decreasing all the decreasible $j < \min(S_0)$. By this we mean the ℓ given by the following instance of strong Σ^0_1-bounding (in RCA_0):

$$\forall m\, \exists \ell\, (\forall j < m)\, \left[\exists d\, (d \text{ decreases } j) \to (\exists d < \ell)\, (d \text{ decreases } j)\right],$$

where we can take $m = \min(S_0)$.
Since $\mathcal{B}$ is an infinite block sequence, there exists $S > S_{n-3}$ in $\mathcal{B}$ such that $\max(S) > \ell$. Then, for any $T > S$ in $\mathcal{B}$, we have $g(S_0 \cup \ldots \cup S_{n-3} \cup S) = g(S_0 \cup \ldots \cup S_{n-3} \cup S \cup T)$, since no elements in T are important in $S_0 \cup \ldots \cup S_{n-3} \cup S \cup T$. Also, by monochromaticity of $\mathcal{B}$, $g(S_0 \cup \ldots \cup S_{n-3} \cup S \cup T) = c$, so $g(S_0 \cup \ldots \cup S_{n-3} \cup S) = c$, hence proving the Claim.

Now, we define $f : \mathbf{N} \to \mathbf{N}$ as $f(i) = \max(B_q)$, where q is minimal such that $B_{p+n-3} < B_q$ and $g(B_p \cup B_{p+1} \cup \ldots \cup B_{p+n-3} \cup B_q) = c$, with p minimal such that $i < \min(B_p)$. Notice that q exists by Claim 3. Also, f has the Property P, i.e., each decreasible i is decreased by some $j \leq f(i)$. In order to prove this, assume by way of contradiction that i is decreasible, p is minimal such that $i < \min(B_p)$, and q is minimal such that $B_{p+n-3} < B_q$ and $g(B_p \cup B_{p+1} \cup \ldots \cup B_{p+n-3} \cup B_q) = c$, but i is decreasible only by numbers larger than $f(i) = \max(B_q)$.
By strong Σ^0_1-bounding, let ℓ be the actual upper bound of the minimal indexes decreasing all the decreasible $j < \min(B_p)$. Since $\mathcal{B}$ is an infinite block sequence, there exists $B > B_q$ in $\mathcal{B}$ such that $\min(B) > \ell$. Now, consider:

$$B_p \cup \ldots \cup B_{p+n-3} \cup B_q \cup B = \{b_0, \ldots, b_r, b_{r+1}, \ldots b_t\},$$

where $b_0 = \min(B_p) = \min(B_p \cup \ldots \cup B_{p+n-3} \cup B_q \cup B)$, $b_r = \max(B_q)$ and $b_{r+1} = \min(B)$. Clearly, $j \leq t$ is important in $B_p \cup \ldots \cup B_{p+n-3} \cup B_q \cup B$ if and only if either $j \leq r$ and j is important in $B_p \cup \ldots \cup B_{p+n-3} \cup B_q$, or $j = r+1$; hence, $g(B_p \cup \ldots \cup B_{p+n-3} \cup B_q) \neq g(B_p \cup \ldots \cup B_{p+n-3} \cup B_q \cup B) = c$, contra our assumption that $g(B_p \cup \ldots \cup B_{p+n-3} \cup B_q) = c$. □

We obtain the following immediate corollary.

Corollary 2.4.3. *Let $n \geq 3, k \geq 2$. Over RCA_0, $\mathsf{FUT}_k^{=n}$ implies ACA_0.*

Proof. From Theorem 1.1.16 and Theorem 2.4.2. □

The proof of Theorem 2.4.2 can be easily adapted to $\mathsf{FUT}_k^{\leq 2}$ (in place of $\mathsf{FUT}_k^{=n}$). Yet, while $\mathsf{FUT}_k^{=n}$ is provably equivalent to ACA_0 and so Theorem 2.4.2 is an optimal result, we do not know the actual strength of $\mathsf{FUT}_k^{\leq 2}$, since we only know that $\mathsf{ACA}_0 \leq \mathsf{FUT}_k^{\leq 2} \leq \mathsf{FUT} \leq \mathsf{ACA}_0^+$. Hence, extending the above approach – adapted to $\mathsf{FUT}_k^{\leq 2}$ – to stronger well-ordering principles would improve the known lower bound on $\mathsf{FUT}_k^{\leq 2}$ and, a fortiori, on full Hindman's Theorem.

Chapter 3

Infinitely many colours

3.1 Canonical and Regressive Ramsey's Theorem

One natural question about Ramsey's Theorem is whether this principle can be extended to colourings of tuples into infinitely many colours. Clearly, in that case, the homogeneous condition on the solution set needs to be changed: by having an infinite number of colours available, it is indeed quite easy to come up with an instance of the problem – i.e., a function $c : [\mathbf{N}]^n \to \mathbf{N}$, for some $n > 0$ – that admits no infinite monochromatic sets.

However, by carefully relaxing the constraint on the solution set, one can indeed obtain an actual generalization of Ramsey's Theorem for infinitely many colours, as shown by Erdős and Rado [ER]. The idea is to impose the monochromatic condition not on every tuple with elements in the solution set, but rather just on the tuples sharing the same values in a uniform subset of their coordinates. To better formalize this idea, we need to introduce the following notion. For $n > 0$, $S \subseteq \{1, \dots, n\}$, $I = \{i_1 < \cdots < i_n\} \subseteq \mathbf{N}$ and $J = \{j_1 < \cdots < j_n\} \subseteq \mathbf{N}$ we say that I and J *agree* on S if and only if for all $s \in S$, $i_s = j_s$. Note that if S is empty then all n-sized subsets of $\mathbf{N}$ agree on S.

Using this notion, we can then state the so-called Canonical Ramsey's Theorem.

Definition 3.1.1 (Erdős and Rado's Canonical Ramsey's Theorem). *Let $n > 0$. We denote by canRT^n the following principle. For all $c : [\mathbf{N}]^n \to \mathbf{N}$ there exists an infinite set $H \subseteq \mathbf{N}$ and a finite (possibily empty) set $S \subseteq \{1, \dots, n\}$ such that for all $I, J \in [H]^n$ the equality $c(I) = c(J)$ holds if and only if I and J agree on S. The set H is called* canonical *for c. We use canRT to denote $(\forall n \geq 1)\,\mathsf{canRT}^n$.*

It is interesting to notice that, if $S = \emptyset$, we obtain a monochromatic solution set (therefore, H is a solution for c to the standard Ramsey's Theorem). If $S = \{1, \dots, n\}$, instead, H is usually called a *rainbow set*, since each tuple in $[H]^n$ has a distinct colour, while if $S = \{1\}$, the colour of the tuples in $[H]^n$ only depends on the minimum of each tuple. However, this case should not be confused with the notion of *min-homogeneous set* given in Definition 3.1.3, for the latter allows two tuples with different minimum to share the same colour.

The Reverse Mathematics of canRT^n is studied in [Mil], where it is denoted by CAN^n. The different notation adopted here is due to the necessity of better distinguishing

the Canonical version of Ramsey's Theorem from the analogous version of Hindman's Theorem (presented in Sec. 3.2).
As observed in [Mil] (Proposition 8.5), canRT^1 is equivalent to RT^1 over RCA_0.
Kanamori and McAloon [KM] isolated a straightforward corollary of the Canonical Ramsey's Theorem inspired by Fodor's Lemma in Uncountable Combinatorics. To state Kanamori-McAloon's principle we need the following definitions.

Definition 3.1.2 (Regressive function)**.** *Let $n \geq 1$. A function $c : [\mathbf{N}]^n \to \mathbf{N}$ is called* regressive *if and only if, for all $I \in [\mathbf{N}]^n$, $c(I) < \min(I)$ if $\min(I) > 0$, otherwise $c(I) = 0$.*

Definition 3.1.3 (Min-homogeneity)**.** *Let $n \geq 1$, $c : [\mathbf{N}]^n \to \mathbf{N}$ and $H \subseteq \mathbf{N}$ an infinite set. The set H is* min-homogeneous *for c if and only if the following condition holds: for any $I, J \in [H]^n$, if $\min(I) = \min(J)$ then $c(I) = c(J)$.*

It is worth noting that – as mentioned before – two tuples with elements in a min-homogeneous set can share the same colour even if their minimum is different.

We can now formalize the so-called Regressive Ramsey's Theorem.

Definition 3.1.4 (Regressive Ramsey's Theorem)**.** *Let $n \geq 1$. We denote by regRT^n the following principle. For all regressive $c : [\mathbf{N}]^n \to \mathbf{N}$ there exists an infinite min-homogeneous set $H \subseteq \mathbf{N}$. We denote by regRT the principle $(\forall n \geq 1)\, \mathsf{regRT}^n$.*

The Reverse Mathematics of regRT^n is studied in [Mil], where it is denoted by REG^n.
Again, we use a different notation to better distinguish the Regressive version of Ramsey's Theorem from the analogous version of Hindman's Theorem discussed in Sec. 3.2.
Note that regRT^1 is trivial. A finite first-order miniaturization of regRT was proved by Kanamori and McAloon [KM] to be independent from Peano Arithmetic and is often considered one of the most mathematically natural examples of statements independent from that system.
The next theorems (both stated as in [Mil]) summarize the main known results about the Reverse Mathematics of the Canonical and the Regressive versions of Ramsey's Theorem.

Theorem 3.1.5. *The following are equivalent over RCA_0.*

1. ACA_0.
2. canRT^n, *for any fixed $n \geq 2$.*
3. regRT^n, *for any fixed $n \geq 2$.*
4. RT^n, *for any fixed $n \geq 3$.*
5. RT^n_k, *for any fixed $n \geq 3$ and $k \geq 2$.*

Proof. See [Mil, Proposition 8.2]. □

$$\begin{array}{ccccccc}
\mathsf{regRT} & \longleftrightarrow & \mathsf{canRT} & \longleftrightarrow & \mathsf{RT} & \longleftrightarrow & \mathsf{ACA}_0' \\
\Downarrow & & \Downarrow & & \Downarrow & & \Downarrow \\
\mathsf{regRT}^n, n \geq 2 & \longleftrightarrow & \mathsf{canRT}^n, n \geq 2 & \longleftrightarrow & \mathsf{RT}^n, n \geq 3 & \longleftrightarrow & \mathsf{ACA}_0 \\
 & & & & \updownarrow & & \\
 & & & & \mathsf{RT}^n_k, n \geq 3, k \geq 2 & &
\end{array}$$

Figure 3.1. Implications over RCA_0. Double arrows indicate strict implications. The equivalences with ACA_0 are from Theorem 3.1.5. For the other implications we refer the reader to [Mil].

The proof above does not directly relate (2) and (3) with (4) and (5), i.e., it does not highlight any straight relationship between the standard Ramsey's Theorems and their versions using infinitely many colours. However, Theorem 6.14 in [Hir1] gives an implication (and a strong Weihrauch reduction) from RT^{2n-1}_2 to regRT^n for any fixed $n \geq 2$, while a simple forgetful function argument proves RT^n from regRT^{n+1}. In addition to these implications, in Sec. 3.3 (Proposition 3.3.5) we enrich the relationship between these two principles by giving a simple direct and exponent-preserving proof of RT^n from regRT^n, which seems to be missing in the literature.

Theorem 3.1.6. *The following are equivalent over* RCA_0*.*

1. ACA_0'.
2. canRT.
3. regRT.
4. RT.

Proof. See [Mil, Proposition 8.4]. □

The main relations among Canonical, Regressive and standard Ramsey's Theorems with respect to implication over RCA_0 are visualized in Figure 3.1.

3.2 Canonical and Regressive Hindman's Theorem

Similarly to Ramsey's Theorem, a variation of Hindman's Theorem for colourings using infinitely many colours can be stated. In [Tay], Taylor proved the following "canonical" version of Hindman's Theorem, analogous to the Canonical Ramsey's Theorem by Erdős and Rado (Definition 3.1.1). Recall that, for any $X \subseteq \mathbf{N}$, we denote by $\mathrm{FIN}(X)$ the set of all non-empty finite subsets of X.

Definition 3.2.1 (Taylor's Canonical Hindman's Theorem)**.** *We denote by* canHT *the following principle. For all* $c : \mathbf{N} \to \mathbf{N}$ *there exists an infinite set* $H = \{h_0 < h_1 < \cdots\} \subseteq \mathbf{N}$ *such that one of the following holds:*

1. *For all* $I, J \in \mathrm{FIN}(\mathbf{N})$, $c(\sum_{i\in I} h_i) = c(\sum_{j\in J} h_j)$.

2. *For all* $I, J \in \mathrm{FIN}(\mathbf{N})$, $c(\sum_{i\in I} h_i) = c(\sum_{j\in J} h_j)$ *if and only if* $I = J$.

3. *For all* $I, J \in \mathrm{FIN}(\mathbf{N})$, $c(\sum_{i\in I} h_i) = c(\sum_{j\in J} h_j)$ *if and only if* $\min(I) = \min(J)$.

4. *For all* $I, J \in \mathrm{FIN}(\mathbf{N})$, $c(\sum_{i\in I} h_i) = c(\sum_{j\in J} h_j)$ *if and only if* $\max(I) = \max(J)$.

5. *For all* $I, J \in \mathrm{FIN}(\mathbf{N})$, $c(\sum_{i\in I} h_i) = c(\sum_{j\in J} h_j)$ *if and only if* $\min(I) = \min(J)$ *and* $\max(I) = \max(J)$.

The set H *is called* canonical *for* c.

We can easily show that Taylor's Theorem implies Hindman's Theorem, just as the Canonical Ramsey's Theorem implies Ramsey's Theorem.

Proposition 3.2.2. canHT *implies* HT *over* RCA_0. *Moreover,* $\mathsf{canHT} \geq_{\mathrm{sW}} \mathsf{HT}$.

Proof. Let $c : \mathbf{N} \to k$ be a finite colouring of $\mathbf{N}$, with $k \geq 1$. By canHT there exists an infinite set $H \subseteq \mathbf{N}$ such that one of the five canonical cases in Definition 3.2.1 occurs. It is easy to see that, since c is a colouring in k colours, only case (1) of Definition 3.2.1 can occur. Thus $\mathrm{FS}(H)$ is homogeneous for c. The argument obviously establishes a strong Weihrauch reduction. □

As a Hindman-type principle, canHT can be endowed with the apartness condition. However, the argument adopted to prove Corollary 2.3.10 does not immediately apply to the case of Taylor's Theorem. Indeed, what the min-term (or max-term) of a number is depends on whether that number is seen as a sum of elements of I or as a sum of elements of J, in the notation of Corollary 2.3.10.
Nevertheless it is true that Taylor's Theorem implies its own self-strengthening with apartness, as we next prove.

Theorem 3.2.3. canHT *implies* $\mathsf{canHT[ap]}$ *over* RCA_0. *Moreover,* $\mathsf{canHT} \geq_{\mathrm{sW}} \mathsf{canHT[ap]}$.

Proof. Given $c : \mathbf{N} \to \mathbf{N}$, let $H = \{h_0 < h_1 < \cdots\}$ be a solution to canHT for c. Let $H' = \{h'_1 < h'_2 < \cdots\}$ be an infinite apart set such that $\mathrm{FS}(H') \subseteq \mathrm{FS}(H)$ (defined as the set J in the proof of Prop. 2.3.9; notice that, for the implication over RCA_0, RT^1 is required, but by Prop. 3.2.2 and the fact that $\mathsf{HT}^{=1}$ is equivalent to RT^1 over RCA_0, we can derive RT^1 from canHT).
For each $i \in \mathbf{N}$, let $A_i \in \mathrm{FIN}(\mathbf{N})$ be such that $\sum_{a\in A_i} h_a = h'_i$ and $h_{\min(A_i)} > h'_{i-1}$ if $i > 0$. A non-empty set with these properties exists by definition of H'. We fix a uniform computable method to select A_i if more than one choice exists (for instance, we take the set A that satisfies the conditions above and that minimizes $\sum_{a\in A} 2^a$). Then, we can state the following three Claims.

Claim 4. *For any set of indexes* $I = \{i_0 < i_1 < \cdots < i_m\} \in \mathrm{FIN}(\mathbf{N})$, *the following properties hold:*

(i) $A_{i_0} < A_{i_1} < \cdots < A_{i_m}$.

(ii) $\min(\bigcup_{i\in I} A_i) = \min(A_{i_0})$.

(iii) $\max(\bigcup_{i\in I} A_i) = \max(A_{i_m})$.

(iv) $\sum_{i\in I} h'_i = \sum_{s\in\bigcup_{i\in I} A_i} h_s$.

Proof. (i) derives from the fact that, for any $s \in (0, m]$, $h_{\min(A_{i_s})} > h'_{i_s-1} \geq h'_{i_{s-1}} \geq h_{\max(A_{i_{s-1}})}$, which implies $\min(A_{i_s}) > \max(A_{i_{s-1}})$ because H is enumerated in increasing order.
(ii), (iii), and (iv) are trivial consequences of (i). □

Claim 5. *For any* $I = \{i_0 < i_1 < \cdots < i_m\}$ *and* $J = \{j_0 < j_1 < \cdots < j_n\}$ *in* $\mathrm{FIN}(\mathbf{N})$, $\min(I) = \min(J)$ *if and only if* $\min(\bigcup_{i\in I} A_i) = \min(\bigcup_{j\in J} A_j)$.

Proof. ($\Longrightarrow$) By hypothesis, $i_0 = j_0$, hence $A_{i_0} = A_{j_0}$ and $\min(A_{i_0}) = \min(A_{j_0})$. Then, by Claim 4.(ii), $\min(\bigcup_{i\in I} A_i) = \min(\bigcup_{j\in J} A_j)$.

($\Longleftarrow$) By hypothesis, $\min(\bigcup_{i\in I} A_i) = \min(\bigcup_{j\in J} A_j)$ so, by Claim 4.(ii), we have $\min(A_{i_0}) = \min(A_{j_0})$ and then $h_{\min(A_{i_0})} = h_{\min(A_{j_0})}$. Thus, we can show that $i_0 = j_0$, i.e., $\min(I) = \min(J)$. Assume otherwise, and suppose $i_0 < j_0$ (the case $i_0 > j_0$ is analogous). By definition of A_{j_0}, we can derive $h_{\min(A_{j_0})} > h'_{j_0-1} \geq h'_{i_0} \geq h_{\min(A_{i_0})}$, hence contradicting $h_{\min(A_{i_0})} = h_{\min(A_{j_0})}$. □

Claim 6. *For any* $I = \{i_0 < i_1 < \cdots < i_m\}$ *and* $J = \{j_0 < j_1 < \cdots < j_n\}$ *in* $\mathrm{FIN}(\mathbf{N})$, $\max(I) = \max(J)$ *if and only if* $\max(\bigcup_{i\in I} A_i) = \max(\bigcup_{j\in J} A_j)$.

Proof. ($\Longrightarrow$) By hypothesis, $i_m = j_n$, hence $A_{i_m} = A_{j_n}$ and $\max(A_{i_m}) = \max(A_{j_n})$. Then, by Claim 4.(iii), $\max(\bigcup_{i\in I} A_i) = \max(\bigcup_{j\in J} A_j)$.

($\Longleftarrow$) By hypothesis, $\max(\bigcup_{i\in I} A_i) = \max(\bigcup_{j\in J} A_j)$ so, by Claim 4.(iii), we have $\max(A_{i_m}) = \max(A_{j_n})$ and then $h_{\max(A_{i_m})} = h_{\max(A_{j_n})}$. Thus, we can show that $i_m = j_n$, i.e., $\max(I) = \max(J)$. Assume otherwise, and suppose $i_m < j_n$ (the case $i_m > j_n$ is analogous). By definition of A_{j_n}, we can derive $h_{\max(A_{j_n})} \geq h_{\min(A_{j_n})} > h'_{j_n-1} \geq h'_{i_m} \geq h_{\max(A_{i_m})}$, hence contradicting $h_{\max(A_{i_m})} = h_{\max(A_{j_n})}$. □

Now we can show that H' is a solution to canHT for c by analyzing each case of Definition 3.2.1.

Case 1. For any $I, J \in \mathrm{FIN}(\mathbf{N})$, by homogeneity of H and by Claim 4.(iv), $c(\sum_{i\in I} h'_i) = c(\sum_{s\in\bigcup_{i\in I} A_i} h_s) = c(\sum_{t\in\bigcup_{j\in J} A_j} h_t) = c(\sum_{j\in J} h'_j)$.

Case 2. Let $I, J \in \mathrm{FIN}(\mathbf{N})$. If $I = J$, then $c(\sum_{i\in I} h'_i) = c(\sum_{j\in J} h'_j)$. Now assume $I \neq J$, as witnessed by $w \in I \setminus J$ (the case $w \in J \setminus I$ is analogous). By Claim 4.(i) applied to $J \cup \{w\}$, we have that $A_w \cap A_j = \emptyset$ for all $j \in J$, therefore $\bigcup_{i\in I} A_i \neq \bigcup_{j\in J} A_j$.

Then, $c(\sum_{i\in I} h'_i) = c(\sum_{s\in\bigcup_{i\in I} A_i} h_s) \neq c(\sum_{t\in\bigcup_{j\in J} A_j} h_t) = c(\sum_{j\in J} h'_j)$, where the two equalities hold by Claim 4.(iv), while the inequality holds by Case 2 of Definition 3.2.1, since c is applied to sums of different elements in H on the two sides of the equality, as we noted above.

Case 3. Let $I, J \in \mathrm{FIN}(\mathbf{N})$. If $\min(I) = \min(J)$, then we have $c(\sum_{i\in I} h'_i) = c(\sum_{s\in\bigcup_{i\in I} A_i} h_s) = c(\sum_{t\in\bigcup_{j\in J} A_j} h_t) = c(\sum_{j\in J} h'_j)$, where the first and the last equality hold by Claim 4.(iv), while the second equality holds by Case 3 of Definition 3.2.1, since in both sides of the equality, c is applied to sums of elements in H having the same minimum term by Claim 5. Similarly, if $\min(I') \neq \min(J')$, we have $c(\sum_{i\in I} h'_i) = c(\sum_{s\in\bigcup_{i\in I} A_i} h_s) \neq c(\sum_{t\in\bigcup_{j\in J} A_j} h_t) = c(\sum_{j\in J} h'_j)$.

Case 4. The proof is similar to the proof of Case 3, but using Claim 6 in place of Claim 5.

Case 5. The proof is analogous to the proof of Cases 3 and 4. □

As observed in the previous Section, when the Canonical Ramsey's Theorem is applied to regressive functions the Regressive Ramsey's Theorem is obtained. Similarly, a regressive version of Hindman's Theorem follows from Taylor's Theorem. We introduce the suitable versions of the notions of regressive function and min-homogeneous set.

Definition 3.2.4 (λ-regressive function)**.** *A function $c : \mathbf{N} \to \mathbf{N}$ is called* λ-regressive *if and only if, for all $n \in \mathbf{N}$, $c(n) < \lambda(n)$ if $\lambda(n) > 0$, while $c(n) = 0$ if $\lambda(n) = 0$.*

Obviously every λ-regressive function is regressive since $\lambda(n) \leq n$ for any n.

Definition 3.2.5 (Min-term-homogeneity for FS)**.** *Let $c : \mathbf{N} \to \mathbf{N}$ and $H = \{h_0 < h_1 < \cdots\} \subseteq \mathbf{N}$. We call* $\mathrm{FS}(H)$ min-term-homogeneous *for c if and only if, for all $I, J \in \mathrm{FIN}(\mathbf{N})$, if $\min(I) = \min(J)$ then $c(\sum_{i\in I} h_i) = c(\sum_{j\in J} h_j)$.*

We can then state the following principle, which is an analogue of Kanamori-McAloon's Regressive Ramsey's Theorem in the spirit of Hindman's Theorem.

Definition 3.2.6 (Regressive Hindman's Theorem)**.** *We denote by $\lambda\mathsf{regHT}$ the following principle. For all λ-regressive $c : \mathbf{N} \to \mathbf{N}$ there exists an infinite $H \subseteq \mathbf{N}$ such that $\mathrm{FS}(H)$ is min-term-homogeneous.*

We start by observing how Taylor's Theorem implies the Regressive Hindman's Theorem just as the Canonical Ramsey's Theorem implies the Kanamori-McAloon Regressive Ramsey's Theorem.

Proposition 3.2.7. canHT *implies* $\lambda\mathsf{regHT}$ *over* RCA_0. *Moreover,* $\mathsf{canHT} \geq_{\mathrm{sW}} \lambda\mathsf{regHT}$.

Proof. Let $c : \mathbf{N} \to \mathbf{N}$ be a λ-regressive function. By canHT there exists an infinite set $H \subseteq \mathbf{N}$ such that one of the five canonical cases occurs for $\mathrm{FS}(H)$. It is easy to see that, since c is λ-regressive, only case (1) and case (3) of Definition 3.2.1 can occur. Thus $\mathrm{FS}(H)$ is min-term-homogeneous for c. □

Similarly to Hindman's Theorem and Taylor's Theorem, the Regressive Hindman's Theorem self-improves to its own version with apartness, as shown below. We first show that $\lambda\mathsf{regHT}$ implies Ramsey's Theorem for singletons.

Lemma 3.2.8. $\lambda\mathsf{regHT}$ *implies* RT^1 *over* RCA_0.

Proof. Let $f : \mathbf{N} \to k$, with $k \geq 1$, and let $g : \mathbf{N} \to \mathbf{N}$ be defined as follows:

$$g(n) = \begin{cases} \lambda'(n) & \text{if } \lambda'(n) < k, \\ f(n) & \text{otherwise,} \end{cases}$$

where $\lambda'(n) = \lambda(n) - 1$ if $\lambda(n) > 0$, otherwise $\lambda'(n) = 0$.
Clearly, g is f-computable and λ-regressive, so let $H = \{h_0 < h_1 < \cdots\}$ be a solution to $\lambda\mathsf{regHT}$ for g. First, we prove the following Claim.

Claim. *There exists an infinite* $H' = \{h'_0 < h'_1 < \cdots\} \subseteq H$ *such that* $\lambda'(h'_{n_1} + h'_{n_2} + h'_{n_3} + h'_{n_4}) \geq k$ *for all* $n_1 < n_2 < n_3 < n_4$.

Proof. Let us define $J = \{j \in H \mid \lambda'(j) < k\}$. If J contains finitely many elements, then $(H \setminus J)$ witnesses the existence of H'. Thus, let us assume $J = \{j_0 < j_1 < \cdots\}$ is infinite.
Notice that the sequence $\lambda'(j_0), \lambda'(j_1), \ldots$ never decreases: suppose otherwise by way of contradiction, and let $j, j' \in J$ be such that $j < j'$ and $\lambda'(j) > \lambda'(j')$. Then we have $g(j) = \lambda'(j) > \lambda'(j') = \lambda'(j + j') = g(j + j')$; this contradicts the min-term-homogeneity of $\mathrm{FS}(H)$. Hence λ' on J is a bounded non-decreasing function on an infinite set.
Then we have two cases. Either for any $j \in J$ there exists $j' > j$ in J such that $\lambda'(j') > \lambda'(j)$, or there exists $j \in J$ such that, for any $j' > j$ in J, $\lambda'(j) \geq \lambda'(j')$. The former case can not hold, since by definition of J, $\lambda'(j) < k$ for any $j \in J$.
In the latter case, instead, we have some $m \in J$ such that $\lambda'(m) \geq \lambda'(j)$ for any j in J. Since $\lambda'(j_0), \lambda'(j_1), \ldots$ is non-decreasing, $\lambda'(j) = \lambda'(m)$ holds for each j in the infinite set $J' = J \setminus [0, m)$. Finally, we can show that J' witnesses the existence of H'. Assume otherwise by way of contradiction. Then, there exist $j, j', j'', j''' \in J'$ such that $j < j' < j'' < j'''$ and $\lambda'(j + j' + j'' + j''') < k$. Thus $g(j + j' + j'' + j''') = \lambda'(j + j' + j'' + j''')$ by definition of g. On the other hand, since $j \in J' \subseteq J$, $\lambda'(j) < k$ and therefore $g(j) = \lambda'(j)$ by definition of g. Moreover, $\lambda'(j) = \lambda'(j') = \lambda'(j'') = \lambda'(j''')$ since $j, j', j'', j''' \in J'$. Therefore we have the following inequality

$$g(j + j' + j'' + j''') = \lambda'(j + j' + j'' + j''') > \lambda'(j) = g(j),$$

contradicting the min-term-homogeneity of $\mathrm{FS}(H)$. This completes the proof of the Claim. Notice that, while $\lambda(x) = \lambda(y)$ implies $\lambda(x + y) > \lambda(x)$ for any $x, y \in \mathbf{N}^+$, the same implication does not hold when using λ': hence, sums of 4 elements are required in the argument above.

In order to prove the lemma, let $H' = \{h'_0 < h'_1 < \cdots\}$ be as in the previous Claim. Then, for any $n_0 < n_1 < n_2$ in $\mathbf{N}^+$, we have

$$\begin{aligned} f(h'_0 + h'_{n_0} + h'_{n_1} + h'_{n_2}) &= g(h'_0 + h'_{n_0} + h'_{n_1} + h'_{n_2}) \\ &= g(h'_0 + h'_1 + h'_2 + h'_3) \\ &= f(h'_0 + h'_1 + h'_2 + h'_3), \end{aligned}$$

where the first and the last equalities hold by the previous Claim and by definition of g, while the second equality holds by min-term-homogeneity of $\mathrm{FS}(H)$.
Hence $\{(h'_0 + h'_{n_0} + h'_{n_1} + h'_{n_2}) \mid 0 < n_1 < n_2 < n_3\}$ is an infinite homogeneous set for f. □

Although the proof above does not witness any Weihrauch reduction, in the next Section we will observe that RT^1_k can be Weihrauch-reduced to some restriction of $\lambda\mathsf{regHT}$ with apartness – hence, a fortiori, it can be Weihrauch-reduced to $\lambda\mathsf{regHT}$ (see Proposition 3.3.4 *infra*).
We can now derive $\lambda\mathsf{regHT}[\mathsf{ap}]$ from $\lambda\mathsf{regHT}$.

Proposition 3.2.9. $\lambda\mathsf{regHT}$ *implies* $\lambda\mathsf{regHT}[\mathsf{ap}]$ *over* RCA_0. *Moreover,* $\lambda\mathsf{regHT} \geq_{\mathrm{sW}} \lambda\mathsf{regHT}[\mathsf{ap}]$.

Proof. The proof of Theorem 3.2.3 adapts *verbatim* to the case of $\lambda\mathsf{regHT}$.
Lemma 3.2.8 takes care of the use of RT^1 in that proof, which is only needed for the implication over RCA_0. □

It is easy to see that the proof of Lemma 3.2.8 uses only sums of at most 4 terms. However, this does not help in extending the previous Proposition to some restriction of $\lambda\mathsf{regHT}$ (see section 3.3), since the argument (which is derived from the proof of Theorem 3.2.3) still requires sums of arbitrary length.

The following proposition shows that the Regressive Hindman's Theorem implies Hindman's Theorem.

Proposition 3.2.10. $\lambda\mathsf{regHT}$ *implies* HT *over* RCA_0.

Proof. Given $f : \mathbf{N} \to k$, with $k \geq 1$, and let $g : \mathbf{N} \to k$ be as follows:

$$g(n) = \begin{cases} f(n) & \text{if } f(n) < \lambda(n), \\ 0 & \text{otherwise.} \end{cases}$$

The function g is λ-regressive by construction and obviously f-computable. Let $H = \{h_0 < h_1 < \cdots\}$ be an infinite set such that $FS(H)$ is min-term-homogeneous for g. By Proposition 3.2.9 we can assume that H is apart. Let i be the minimum such that $\lambda(h_i) > k$. Let $H^- = H \setminus \{h_0, \ldots, h_i\}$. By choice of H^-, g behaves like f on $\mathrm{FS}(H^-)$. Let g^- be the k-colouring of numbers induced by g on H^-.
By RT^1_k (which we can assume by Lemma 3.2.8) let $H' = \{h'_0 < h'_1 < \cdots\}$ be an infinite subset of H^- homogeneous for g^-. Then, for $\{s_1, \ldots, s_m\}$ and $\{t_1, \ldots, t_n\}$ non-empty subsets of H', we have

$$\begin{aligned} f(s_1 + \cdots + s_m) &= g(s_1 + \cdots + s_m) \\ &= g(s_1) = g^-(s_1) \\ &= g^-(t_1) = g(t_1) \\ &= g(t_1 + \cdots + t_n) \\ &= f(t_1 + \cdots + t_n), \end{aligned}$$

since $\mathrm{FS}(H^-)$ is min-term-homogeneous for g and g coincides with f on $\mathrm{FS}(H^-)$. □

We do not know if the implication in Proposition 3.2.10 can be reversed. In fact, obtaining such a result is probably not trivial: in particular, that seems to be as hard as proving any other implication from HT to whatever principle stronger than ACA_0 (as might be the case for $\lambda\mathsf{regHT}$, since its restricted versions – presented in the next section – are already equivalent to ACA_0). Even though proving the equivalence between $\lambda\mathsf{regHT}$ and HT would not immediately improve the lower bound on the latter – since, to date, our best result about the lower bound on $\lambda\mathsf{regHT}$ is still ACA_0 – such an equivalence may provide an additional tool to achieve this goal: as we will see in Sec. 3.3, using $\lambda\mathsf{regHT}$ in place of HT may indeed result in more natural arguments, due to the infinite number of colours available.

3.3 The strength of Bounded Regressive Hindman's Theorem

As already mentioned in Sec. 2.3, restrictions of Hindman's Theorem relaxing the monochromaticity condition received substantial attention in recent years. To formulate analogous restrictions of $\lambda\mathsf{regHT}$ we extend the definition of min-term-homogeneity in the natural way. For $n \geq 1$, we denote by $\mathrm{FIN}^{\leq n}(\mathbf{N})$ (resp. $\mathrm{FIN}^{=n}(\mathbf{N})$) the set of all non-empty subsets of $\mathbf{N}$ of cardinality at most n (resp. of cardinality exactly n).

Definition 3.3.1 (Min-term-homogeneity for $\mathrm{FS}^{\leq n}, \mathrm{FS}^{=n}$)**.** *Let $n \geq 1$. Let $c : \mathbf{N} \to \mathbf{N}$ be a colouring and $H = \{h_0 < h_1 < \cdots\}$ an infinite subset of $\mathbf{N}$. We call $\mathrm{FS}^{\leq n}(H)$ (resp. $\mathrm{FS}^{=n}(H)$)* min-term-homogeneous for c *if and only if, for all $I, J \in \mathrm{FIN}^{\leq n}(\mathbf{N})$ (resp. $I, J \in \mathrm{FIN}^{=n}(\mathbf{N})$), if $\min(I) = \min(J)$ then $c(\sum_{i \in I} h_i) = c(\sum_{j \in J} h_j)$.*

We can then formulate the natural restrictions of the Regressive Hindman's Theorem obtained by relaxing the min-term-homogeneity requirement from $\mathrm{FS}(H)$ to $\mathrm{FS}^{\leq n}(H)$ or $\mathrm{FS}^{=n}(H)$.

Definition 3.3.2 (Bounded λ-Regressive Hindman's Theorems)**.** *Let $n \geq 1$. We denote by $\lambda\mathsf{regHT}^{\leq n}$ (resp. $\lambda\mathsf{regHT}^{=n}$) the following principle. For all λ-regressive $c : \mathbf{N} \to \mathbf{N}$ there exists an infinite $H \subseteq \mathbf{N}$ such that $\mathrm{FS}^{\leq n}(H)$ (resp. $\mathrm{FS}^{=n}$) is min-term-homogeneous for c.*

Note that $\lambda\mathsf{regHT}^{\leq 1}$ and $\lambda\mathsf{regHT}^{=1}$ are trivial. We also point out the following obvious relations: $\lambda\mathsf{regHT}$ yields $\lambda\mathsf{regHT}^{\leq n}$ which yields $\lambda\mathsf{regHT}^{=n}$ for all n (both in RCA_0 and by strong Weihrauch reductions), and the same chain of implications and reductions holds for the versions with the apartness condition. Also, for $m > n$, $\lambda\mathsf{regHT}^{\leq m}$ obviously yields $\lambda\mathsf{regHT}^{\leq n}$, while $\lambda\mathsf{regHT}^{=m}$ yields $\lambda\mathsf{regHT}^{=n}$ if m is a multiple of n (see the analogous results for Hindman's Theorem for sums of exactly n terms in [CKLZ], Proposition 3.5).
We compare the bounded versions of our regressive Hindman's Theorem with other prominent Ramsey-type and Hindman-type principles.
We start with the following simple Lemma showing that, for any $n \geq 2$, $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$ implies RT^1. Note that in Lemma 3.2.8 we established that $\lambda\mathsf{regHT}$ *without apartness* implies RT^1 and we later used this result to show that $\lambda\mathsf{regHT}$ implies $\lambda\mathsf{regHT}[\mathsf{ap}]$ (Proposition 3.2.9).

Lemma 3.3.3. *Let $n \geq 2$. Over* RCA_0*,* $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$ *implies* RT^1*. Moreover, for any $k \geq 1$, we have* $\mathsf{RT}^1_k \leq_{\mathrm{sW}} \lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$.

Proof. We give the proof for $n = 2$ for ease of readability. Let $k \geq 1$ and $f : \mathbf{N} \to k$ be given. Define $g : \mathbf{N} \to k$ as follows.

$$g(m) = \begin{cases} 0 & \text{if } \lambda(m) \leq k, \\ f(\mu(m)) & \text{otherwise.} \end{cases}$$

Clearly g is λ-regressive and f-computable in a uniform way.
Let $H = \{h_0 < h_1 < \cdots\}$ be an infinite apart set such that $\mathrm{FS}^{=2}(H)$ is min-term-homogeneous for g.
By the apartness condition, for all $h \in H \setminus \{h_0, h_1, \ldots, h_k\}$ we have $g(h) = f(\mu(h))$. Then it is easy to see that $M = \{\mu(h_{k+2}), \mu(h_{k+3}), \ldots\}$ is an infinite f-homogeneous set of colour $f(\mu(h_{k+2}))$ since, for any i, $f(\mu(h_{k+2+i})) = g(h_{k+1} + h_{k+2+i}) = g(h_{k+1} + h_{k+2}) = f(\mu(h_{k+2}))$. □

The next proposition relates the principles $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$ (respectively $\lambda\mathsf{regHT}^{\leq n}[\mathsf{ap}]$) with the principles $\mathsf{HT}^{=n}_k[\mathsf{ap}]$ (respectively $\mathsf{HT}^{\leq n}_k[\mathsf{ap}]$). The argument is essentially the same as the proof of Proposition 3.2.10.

Proposition 3.3.4. *Let $n \geq 2$.*

1. $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$ *implies* $\mathsf{HT}^{=n}[\mathsf{ap}]$ *over* RCA_0*. Also,* $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}] \geq_{\mathrm{c}} \mathsf{HT}^{=n}_k[\mathsf{ap}]$ *for any $k \geq 1$.*
2. $\lambda\mathsf{regHT}^{\leq n}[\mathsf{ap}]$ *implies* $\mathsf{HT}^{\leq n}[\mathsf{ap}]$ *over* RCA_0*. Also,* $\lambda\mathsf{regHT}^{\leq n}[\mathsf{ap}] \geq_{\mathrm{c}} \mathsf{HT}^{\leq n}_k[\mathsf{ap}]$ *for any $k \geq 1$.*

Proof. We prove the second point, the proof of the first point being completely analogous. Given $k \geq 1$ and $f : \mathbf{N} \to k$, let $g : \mathbf{N} \to k$ be as follows:

$$g(m) = \begin{cases} f(m) & \text{if } f(m) < \lambda(m), \\ 0 & \text{otherwise.} \end{cases}$$

The function g is λ-regressive and f-computable. By $\lambda\mathsf{regHT}^{\leq n}[\mathsf{ap}]$ let $H = \{h_0 < h_1 < \cdots\} \subseteq \mathbf{N}$ be an infinite apart set such that $\mathrm{FS}^{\leq n}(H)$ is min-term-homogeneous for g. Let $g' : H \setminus \{h_0, \ldots, h_{k-1}\} \to k$ be defined as $g'(h_i) = g(h_i + h_{i+1} + \cdots + h_{i+n-1})$. By RT^1_k, let $H' \subseteq H$ be an infinite homogeneous set for g'. For the sake of establishing the implication over RCA_0, recall that RT^1 follows from $\lambda\mathsf{regHT}^{=2}[\mathsf{ap}]$ by Lemma 3.3.3 and therefore also from $\lambda\mathsf{regHT}^{\leq n}[\mathsf{ap}]$ for any $n \geq 2$. For the sake of the computable reduction result, just notice that for each fixed $k \geq 1$, RT^1_k is computably true. Then, for $\{s_1, \ldots, s_p\}$ and $\{t_1, \ldots, t_q\}$ non-empty subsets of H', with $p, q \leq n$ and $s_1 < \cdots < s_p$, $t_1 < \cdots < t_q$, we have

$$\begin{aligned} f(s_1 + \cdots + s_p) &= g(s_1 + \cdots + s_p) \\ &\overset{(*)}{=} g(s_1) = g'(s_1) \\ &= g'(t_1) = g(t_1) \\ &\overset{(**)}{=} g(t_1 + \cdots + t_q) \\ &= f(t_1 + \cdots + t_q), \end{aligned}$$

where the equalities dubbed by $(*)$ and $(**)$ hold by the min-term-homogeneity of $\mathrm{FS}^{\leq n}(H)$ for g. This shows that H' is an apart solution to $\mathsf{HT}^{\leq n}_k$ for f. ☐

Remark 1. *The previous proof gives us a hint as to how extend the reduction to* $\mathsf{HT}^{\leq n}[\mathsf{ap}]$*, i.e. to the universally-quantified principles* $(\forall k \geq 1)$ $\mathsf{HT}^{\leq n}_k[\mathsf{ap}]$*. In that case, the number of colours is not given as part of the instance, and it cannot be computably inferred from the instance* X *of the principle* $\mathsf{HT}^{\leq n}[\mathsf{ap}]$ *(see the discussion in [DM] p. 54 for more details on this issue). Nevertheless, we can easily obtain a computable reduction by just observing that the proof of Proposition 3.3.4 provides us, for any* $k \geq 1$*, with both an* X*-computable procedure giving us an instance* $\widehat{X}$ *of* $\lambda\mathsf{regHT}^{\leq n}[\mathsf{ap}]$*, and an* $(X \oplus \widehat{Y})$*-computable procedure transforming a solution* $\widehat{Y}$ *for* $\widehat{X}$ *to a solution for* X*: so, even if we do not know the actual value of* k*, we know that the two procedures witnessing the computable reduction do exist. Thus, we can conclude that for any* $n \geq 2$*,* $\lambda\mathsf{regHT}^{\leq n}[\mathsf{ap}] \geq_{\mathrm{c}} \mathsf{HT}^{\leq n}[\mathsf{ap}]$*. It is not straightforward to improve this result to a Weihrauch reduction.*
The same argument also applies to the case of $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$*, so that we have that for any* $n \geq 2$*,* $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}] \geq_{\mathrm{c}} \mathsf{HT}^{=n}[\mathsf{ap}]$*.*

Also, we point out that a proof of $\lambda\mathsf{regHT}^{\leq 2}$ that does not also prove HT (or, more technically, a separation over RCA_0 of these two principles) would answer Question 12 from [HLS].
It is worth noticing that a further slight adaptation of the proof of Proposition 3.3.4 gives a direct proof of RT^n from regRT^n and also shows that $\mathsf{regRT}^n \geq_{\mathrm{c}} \mathsf{RT}^n_k$. The following definition can be used for computably reducing RT^n_k to regRT^n (for $n \geq 2$ and $k \geq 1$). Given $k \geq 1$ and $c : [\mathbf{N}]^n \to k$, let $c^+ : [\mathbf{N}]^n \to k$ be as follows:

$$c^+(x_1, \ldots, x_n) = \begin{cases} 0 & \text{if } x_1 \leq k, \\ c(x_1, \ldots, x_n) & \text{otherwise.} \end{cases}$$

We can thus state the following Proposition.

Proposition 3.3.5. *For any* $n \geq 2$ *and* $k \geq 1$*,* $\mathsf{RT}^n_k \leq_{\mathrm{c}} \mathsf{regRT}^n$*.*

Note that by $\mathsf{HT}^{=n}_k[\mathsf{ap}] \leq_{\mathrm{sW}} \mathsf{RT}^n_k$ (see [CKLZ]), the above also implies $\mathsf{HT}^{=n}_k[\mathsf{ap}] \leq_{\mathrm{c}} \mathsf{regRT}^n$ for any $n \geq 2$ and $k \geq 1$.

Equivalents of ACA_0. Proposition 3.3.4, coupled with the fact that $\mathsf{HT}^{=3}_2[\mathsf{ap}]$ implies ACA_0 (Theorem 3.3 in [CKLZ]), yields the following corollary.

Corollary 3.3.6. $\lambda\mathsf{regHT}^{=3}[\mathsf{ap}]$ *implies* ACA_0 *over* RCA_0*.*

Proof. From Theorem 3.3 in [CKLZ] and Proposition 3.3.4 above. ☐

We have the following reversal, showing that $\lambda\mathsf{regHT}^{=3}[\mathsf{ap}]$ is a "weak yet strong" restriction of Taylor's Theorem in the sense of [Car1]. The result is analogous to the implication from RT^n_k to $\mathsf{HT}^{=n}_k$ (see [CKLZ]).

Theorem 3.3.7. *Let* $n \geq 1$*.* ACA_0 *proves* $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$*. Also,* $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}] \leq_{\mathrm{sW}} \mathsf{regRT}^n$*.*

Proof. We give the proof for $n = 2$ for ease of readability.
Let $f : \mathbf{N} \to \mathbf{N}$ be λ-regressive. Let $g : [\mathbf{N}]^2 \to \mathbf{N}$ be defined as follows: $g(x, y) = f(2^x + 2^y)$. The function g is regressive since f is λ-regressive. Recall that regRT^2 is provable in ACA_0. Let $H \subseteq \mathbf{N}$ be a min-homogeneous solution to regRT^2 for g. Let $\hat{H} = \{2^h : h \in H\}$. Obviously $\hat{H}$ is apart. It is easy to see that $\mathrm{FS}^{=2}(\hat{H})$ is min-term-homogeneous for f: let $2^h < 2^{h'} < 2^{h''}$ be elements of $\hat{H}$. Then

$$f(2^h + 2^{h'}) = g(h, h') = g(h, h'') = f(2^h + 2^{h''}).$$

□

We do not know if the reduction in Theorem 3.3.7 can be reversed.
We next show that $\lambda\mathsf{regHT}^{=2}[\mathsf{ap}]$ already implies Arithmetical Comprehension. The proof is reminiscent of the proof that $\mathsf{HT}_2^{\leq 2}[\mathsf{ap}]$ implies ACA_0 in [CKLZ], but the use of λ-regressive colourings allows us to avoid the parity argument used in that proof. As happens in the proofs of independence of combinatorial principles from Peano Arithmetic [KM], in the present setting the use of regressive colourings simplifies the combinatorics.
Recall from Sec. 1.1.3 that we denote by RAN the Π^1_2-principle stating that for every injective function $f : \mathbf{N} \to \mathbf{N}$ the range of f (denoted by $rg(f)$) exists. It is well-known that RAN is equivalent to ACA_0 (see Theorem 1.1.7).

Theorem 3.3.8. *Let $n \geq 2$. $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$ implies ACA_0 over RCA_0. Moreover, $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}] \geq_{\mathrm{W}} \mathsf{RAN}$.*

Proof. We give the proof for $n = 2$. The easy adaptation to larger values is left to the reader.
Let $f : \mathbf{N} \to \mathbf{N}$ be injective. For technical convenience and without loss of generality we assume that f never takes the value 0. We show, using $\lambda\mathsf{regHT}^{=2}[\mathsf{ap}]$, that $rg(f)$ does exist.
Define $c : \mathbf{N} \to \mathbf{N}$ as follows. If m is a power of 2 then $c(m) = 0$. Else $c(m) =$ the unique x such that $x < \lambda(m)$ and there exists $j \in [\lambda(m), \mu(m))$ such that $f(j) = x$ and for all $j < j' < \mu(m)$, $f(j') \geq \lambda(m)$. If no such x exists, we set $c(m) = 0$.
Intuitively c checks whether there are values below $\lambda(m)$ in the range of f restricted to $[\lambda(m), \mu(m))$. If any, it returns the latest one, i.e., the one obtained as image of the maximal $j \in [\lambda(m), \mu(m))$ that is mapped by f below $\lambda(m)$). In other words, x is the "last" element below $\lambda(m)$ in the range of f restricted to $[\lambda(m), \mu(m))$.
The function c is computable in f and λ-regressive.
Let $H = \{h_0 < h_1 < \cdots\} \subseteq \mathbf{N}$ be an apart solution to $\lambda\mathsf{regHT}^{=2}$ for c. Without loss of generality we can assume that $\lambda(h_0) > 1$, since H is apart. Let $h_i \in H$.
We claim that if $x < \lambda(h_i)$ and x is in the range of f then x is in the range of f restricted to $[0, \mu(h_{i+1}))$.
We prove the claim as follows. Suppose, by way of contradiction, that there exist $h_i \in H$ and $x < \lambda(h_i)$ such that $x \in rg(f)$ but $x \notin f([0, \mu(h_{i+1}))$. Let b be the true bound for the elements in the range of f smaller than $\lambda(h_i)$, i.e., b is such that if $n < \lambda(h_i)$ and $n \in rg(f)$, then $n < b$. The existence of b follows in RCA_0 from strong Σ^0_1-bounding (see [Sim2], Exercise II.3.14):

$$\forall n \exists b \forall i < n (\exists j (f(j) = i) \to \exists j < b (f(j) = i)),$$

where we take $n = \lambda(h_i)$.
Let h_j in H be such that $h_j > h_{i+1}$ and $\mu(h_j) \geq b$. Such an h_j exists since H is infinite.
Then, by min-term-homogeneity of $\mathrm{FS}^{=2}(H)$, $c(h_i+h_{i+1}) = c(h_i+h_j)$. But by choice of h_i, x and h_j, and the definition of c, it must be the case that $c(h_i+h_{i+1}) \neq c(h_i+h_j)$. To see this, first note that, by apartness of H, the following equalities hold:

$$\lambda(h_i + h_{i+1}) = \lambda(h_i) = \lambda(h_i + h_j),\ \mu(h_i + h_{i+1}) = \mu(h_{i+1}),\ \mu(h_i + h_j) = \mu(h_j).$$

Then observe that $c(h_i + h_j) > 0$: by hypothesis $f^{-1}(x) \in [\mu(h_{i+1}), b)$ (recall that f is injective), therefore x is a value of f below $\lambda(h_i+h_j)$ whose pre-image under f is in $[\lambda(h_i+h_j), \mu(h_i+h_j))$, i.e. in $[\lambda(h_i), \mu(h_j))$. Suppose now that $c(h_i+h_{i+1}) = z > 0$. Then, by definition of c, it must be the case that $z < \lambda(h_i + h_{i+1})$, i.e., $z < \lambda(h_i)$, and $f^{-1}(z)$ is in $[\lambda(h_i + h_{i+1}), \mu(h_i + h_{i+1}))$, i.e. in $[\lambda(h_i), \mu(h_{i+1}))$. This z cannot be the value of $c(h_i + h_j)$, since by hypothesis and by choice of b, we have $x < \lambda(h_i)$ and $f^{-1}(x)$ is in $[\mu(h_{i+1}), b)$, hence in $[\lambda(h_i + h_j), \mu(h_i + h_j))$. Thus z cannot be the value of f below $\lambda(h_i)$ with maximal pre-image under f in $[\lambda(h_i + h_j), \mu(h_i + h_j))$ as the definition of $c(h_i + h_j)$ requires, since $f^{-1}(z) < \mu(h_{i+1}) \leq f^{-1}(x)$ and f is injective. This concludes our reasoning by way of contradiction and hence establishes the claim that values in the range of f below $\lambda(h_i)$ appear as values of f applied to arguments smaller than $\mu(h_{i+1})$.
In view of the just established claim it is easy to see that the range of f can be decided computably in H as follows. Given x, pick any $h_i \in H$ such that $x < \lambda(h_i)$ and check whether x appears in $f([0, \mu(h_{i+1}))$. □

Theorem 3.3.8 for the case of $n = 2$ should be contrasted with the fact that $\mathsf{HT}^{=2}_2[\mathsf{ap}]$ follows easily from RT^2_2 and is therefore strictly weaker than ACA_0, while $\mathsf{HT}^{=3}_2[\mathsf{ap}]$ implies ACA_0 as proved in [CKLZ]. The situation matches the one among regRT^2, RT^3_2 and RT^2_2 (see Theorem 3.1.5).
The proof of Theorem 3.3.8 can be recast in a straightforward way to show that there exists a computable λ-regressive colouring such that all apart solutions to $\lambda\mathsf{regHT}^{=2}$ for that colouring compute the first Turing Jump $\emptyset'$. Analogously, the reduction can be cast in terms of the Π^1_2-principle $\forall X \exists Y (Y = (X)')$ expressing closure under the Turing Jump, rather than in terms of RAN.

The next theorem summarizes the equivalences – over RCA_0 – between the Regressive Hindman's theorems for sums of exactly n elements and other prominent Ramsey-theoretic principles (see also Figure 3.2).

Theorem 3.3.9. *The following are equivalent over* RCA_0.

1. ACA_0.
2. regRT^n, *for any fixed* $n \geq 2$.
3. RT^n_k, *for any fixed* $n \geq 3$, $k \geq 2$.
4. $\mathsf{HT}^{=n}_k[\mathsf{ap}]$, *for any fixed* $n \geq 3$, $k \geq 2$.
5. $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$, *for any fixed* $n \geq 2$.

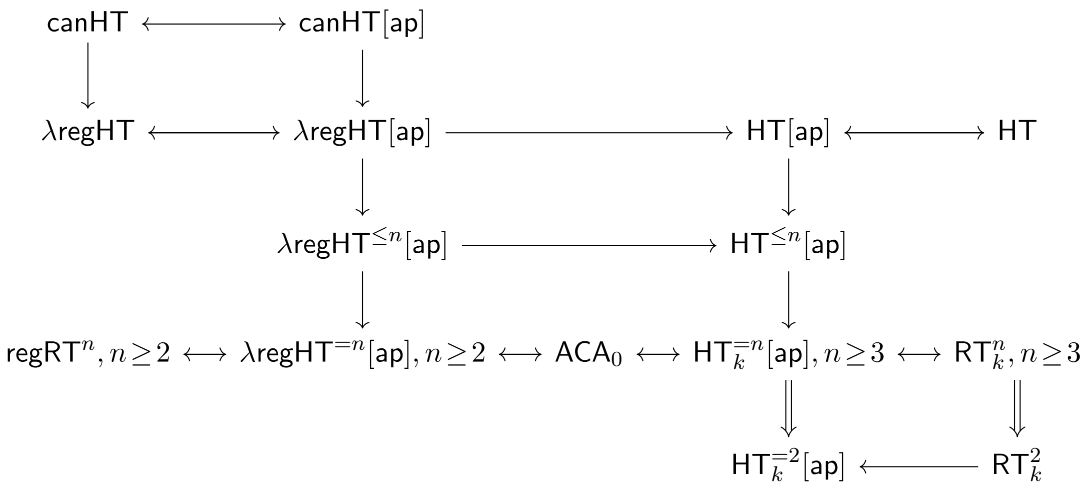

Figure 3.2. Implications over RCA_0. Double arrows indicate strict implications. The equivalence of $\mathsf{canHT}[\mathsf{ap}]$ and canHT is from Theorem 3.2.3. The implication from canHT to $\lambda\mathsf{regHT}$ is from Proposition 3.2.7 and similarly for the versions with apartness. The equivalence between $\lambda\mathsf{regHT}$ and $\lambda\mathsf{regHT}[\mathsf{ap}]$ is from Proposition 3.2.9. The implication from $\lambda\mathsf{regHT}$ to HT is from Proposition 3.2.10. The implication from $\lambda\mathsf{regHT}^{\leq n}[\mathsf{ap}]$ to $\mathsf{HT}^{\leq n}[\mathsf{ap}]$ is from Proposition 3.3.4. Finally, all the equivalences at level ACA_0 are from Theorem 3.3.9.

Proof. The equivalences between points (1), (2) and (3) are as in Theorem 3.1.5. The equivalence of (1) and (4) is from Proposition 3.4 in [CKLZ]. Then the equivalence of (1) and (5) follows from Theorem 3.3.7 and Theorem 3.3.8. ☐

In terms of computable reductions we have the following, for $n \geq 2$ and $k \geq 1$:

$$\mathsf{RT}_2^{2n-1} \geq_{\mathrm{sW}} \mathsf{regRT}^n \geq_{\mathrm{c}} \mathsf{RT}_k^n,$$

where the first inequality is due to Hirst [Hir1] and the second inequality is from Proposition 3.3.5. Furthermore we have that

$$\mathsf{regRT}^n \geq_{\mathrm{W}} \lambda\mathsf{regHT}^{=n}[\mathsf{ap}] \geq_{\mathrm{c}} \mathsf{HT}_k^{=n}[\mathsf{ap}],$$

from Theorem 3.3.7 and Proposition 3.3.4.
Moreover, whereas $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}] \geq_{\mathrm{W}} \mathsf{RAN}$ for any $n \geq 2$ (Theorem 3.3.8), we have that $\mathsf{HT}_k^{=n}[\mathsf{ap}] \geq_{\mathrm{W}} \mathsf{RAN}$ only for $n \geq 3$ and $k \geq 2$ (by an easy adaptation of the proof of Theorem 3.3 in [CKLZ]). Also note that $\mathsf{RT}_k^n \geq_{\mathrm{sW}} \mathsf{HT}_k^{=n}[\mathsf{ap}]$ by a straightforward reduction (see [CKLZ]).
Some non-reducibility results can be gleaned from the above and known non-reducibility results from the literature. First, Dorais, Dzhafarov, Hirst, Mileti, and Shafer showed that $\mathsf{RT}_k^n \not\leq_{\mathrm{sW}} \mathsf{RT}_j^n$ when $k > j$ (Theorem 3.1 of [DJSW]). Then $\mathsf{RT}_k^n \not\leq_{\mathrm{W}} \mathsf{RT}_j^n$ for $k > j$ was proved by Brattka and Rakotoniaina [BR] and, independently, by Hirschfeldt and Jockusch [HJ]. Patey further improved this result by showing that the computable reduction does not hold either [Pat]; i.e., $\mathsf{RT}_k^n \not\leq_{\mathrm{c}} \mathsf{RT}_j^n$ for all $n \geq 2$, $k > j \geq 2$. We can derive, among others, the following corollaries.

Corollary 3.3.10. *For each* $n, k \geq 2$, $\mathsf{regRT}^n \not\leq_{\mathrm{c}} \mathsf{RT}^n_k$.

Proof. From Proposition 3.3.5 we know that $\mathsf{RT}^n_{k+1} \leq_{\mathrm{c}} \mathsf{regRT}^n$, so if we had $\mathsf{regRT}^n \leq_{\mathrm{c}} \mathsf{RT}^n_k$ we could transitively obtain $\mathsf{RT}^n_{k+1} \leq_{\mathrm{c}} \mathsf{RT}^n_k$, hence contradicting the fact that $\mathsf{RT}^n_{k+1} \not\leq_{\mathrm{c}} \mathsf{RT}^n_k$ proved by Patey [Pat]. □

Corollary 3.3.11. $\mathsf{RT}^3_3 \not\leq_{\mathrm{c}} \lambda\mathsf{regHT}^{=2}[\mathsf{ap}]$.

Proof. It is known from [Pat] that $\mathsf{RT}^3_3 \not\leq_{\mathrm{c}} \mathsf{RT}^3_2$. On the other hand $\lambda\mathsf{regHT}^{=2}[\mathsf{ap}] \leq_{\mathrm{W}} \mathsf{RT}^3_2$, since $\lambda\mathsf{regHT}^{=2}[\mathsf{ap}] \leq_{\mathrm{W}} \mathsf{regRT}^2$ (Theorem 3.3.7) and $\mathsf{regRT}^2 \leq_{\mathrm{sW}} \mathsf{RT}^3_2$ (from the proof of Theorem 6.14 in [Hir1]) and since the involved reducibilities satisfy the following inclusions and are transitive: $\leq_{\mathrm{sW}} \subseteq \leq_{\mathrm{W}} \subseteq \leq_{\mathrm{c}}$. □

As proved in [CKLZ], restrictions of Hindman's Theorem have intriguing connections with the so-called Increasing Polarized Ramsey's Theorem for pairs (IPT^2_2) of Dzhafarov and Hirst [DH]. For example, $\mathsf{HT}^{=2}_2[\mathsf{ap}] \geq_{\mathrm{W}} \mathsf{IPT}^2_2$ (Theorem 4.2 in [CKLZ]). By this result and Proposition 3.3.4 we have the following corollary.

Corollary 3.3.12. $\mathsf{IPT}^2_2 \leq_{\mathrm{c}} \lambda\mathsf{regHT}^{=2}[\mathsf{ap}]$.

Note that IPT^2_2 is the strongest known lower bound for $\mathsf{HT}^{=2}_2[\mathsf{ap}]$ in terms of reductions. Some interesting lower bounds on $\mathsf{HT}^{=2}$ *without* apartness are in [CDH$^+$]. We haven't investigated $\lambda\mathsf{regHT}^{=n}$ *without* the apartness condition; we conjecture that the lower bounds on $\mathsf{HT}^{=2}$ (*without* apartness) from [CDH$^+$] can be adapted to $\lambda\mathsf{regHT}^{=2}$.

3.4 Bounded Regressive Hindman's Theorem and Well-Ordering Principles.

Another relevant result that can be inferred from the theorems above (precisely, from Theorem 3.3.8) is that, for any $n \geq 2$, $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$ implies $\mathsf{WOP}(\mathcal{X} \mapsto \boldsymbol{\omega}^{\mathcal{X}})$, due to the aforementioned equivalence of the latter principle with ACA_0. However, such an argument would give neither a direct implication nor a computable reduction. Nevertheless, it is possible to obtain $\mathsf{WOP}(\mathcal{X} \mapsto \boldsymbol{\omega}^{\mathcal{X}})$ from $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$ via a straightforward argument, which also witnesses a Weihrauch reduction.
The proof is quite similar to the proof of Theorem 2.4.2. The idea is, again, to extract an infinite descending sequence τ in $\mathcal{X}$ starting from an infinite descending sequence σ in $\boldsymbol{\omega}^{\mathcal{X}}$: this can be done in RCA_0 (and via uniform computable transformations) by bounding the research for candidate elements of τ using a suitable function f. The existence of such a function can be proved using the min-term-homogeneity of a solution to $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$ for a λ-regressive colouring derived from σ. The interesting point, here, is that showing the existence of f is a little easier than obtaining the same result in the proof of Theorem 2.4.2, due to the infinite number of colours available in λ-regressive colourings (indeed, here we can even avoid the definition of "important number"). This highlights how sometimes regressive colourings can make arguments easier, as is also the case with the proof of Theorem 3.3.8 (with

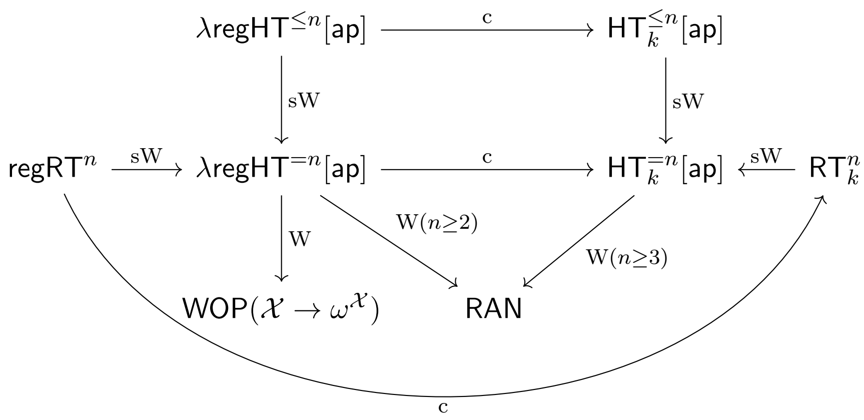

Figure 3.3. Diagram of reductions. $\mathsf{HT}^{\leq n}[\mathsf{ap}] \leq_{\mathrm{c}} \lambda\mathsf{regHT}^{\leq n}[\mathsf{ap}]$ is from Proposition 3.3.4. That the versions with sums of exactly n terms reduce to the corresponding versions for sums of $\leq n$ terms is a trivial observation. The Weihrauch reduction of $\mathsf{WOP}(\mathcal{X} \to \omega^{\mathcal{X}})$ to $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$ for $n \geq 2$ is Theorem 3.4.1. The reduction $\mathsf{RAN} \leq_{\mathrm{W}} \lambda\mathsf{regHT}^{=n}$ for $n \geq 2$ is Theorem 3.3.8. The reduction $\mathsf{RAN} \leq_{\mathrm{W}} \mathsf{HT}^{=n}_k[\mathsf{ap}]$ for $n \geq 3, k \geq 2$ is from [CKLZ]. The reduction $\mathsf{RT}^n_k \leq_{\mathrm{c}} \mathsf{regRT}^n$ is from Proposition 3.3.5. The reduction $\mathsf{HT}^{=n}_h[\mathsf{ap}] \leq_{\mathrm{sW}} \mathsf{RT}^n_k$ is folklore.

respect to the proof of Proposition 5 in [CKLZ], which uses $\mathsf{HT}^{\leq 2}_2[\mathsf{ap}]$ in place of $\lambda\mathsf{regHT}^{=2}[\mathsf{ap}]$) and with Kanamori and McAloon's proof of the independence of a first order version of regRT from PA (with respect to other independence results from PA of analogous principles). Therefore, further investigations on $\lambda\mathsf{regHT}$ aimed at obtaining its equivalence with Hindman's Theorem might be helpful not only in order to improve our knowledge about both principles, but even to possibly give a better answer to the longstanding question about the actual strength of HT.
For the sake of completeness, we now give the complete proof of the direct implication of $\mathsf{WOP}(\mathcal{X} \mapsto \omega^{\mathcal{X}})$ from $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$, which also establish a Weihrauch reduction.

Theorem 3.4.1. *Let $n \geq 2$. $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$ implies $\mathsf{WOP}(\mathcal{X} \mapsto \omega^{\mathcal{X}})$ over RCA_0. Moreover, $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}] \geq_{\mathrm{W}} \mathsf{WOP}(\mathcal{X} \mapsto \omega^{\mathcal{X}})$.*

Proof. Assume by way of contradiction $\neg\mathrm{WO}(\boldsymbol{\omega}^{\mathcal{X}})$, and let $\alpha = (\alpha_n)_{n\in\mathbf{N}}$ be an infinite descending sequence in $\boldsymbol{\omega}^{\mathcal{X}}$. We define the α-computable sequence $\beta = (\beta_n)_{n\in\mathbf{N}}$ of exponents of elements in α as in the proof of Theorem 2.4.2, i.e., $\beta_n = \alpha_{t(n),p(n)}$, where t and p are such that $\theta(t(n), p(n)) = n$, with $\theta(n, m) = m + \sum_{k<n} lh(\alpha_k)$ whenever $m < lh(\alpha_n)$, otherwise it is undefined. Intuitively, $t(n)$ is the element of α from which β_n has been "extracted", while $p(n)$ is the position of the exponent β_n within $\alpha_{t(n)}$.
Recall that i is *decreasible* if there exists $j > i$ such that $p(j) = p(i)$ and $\beta_i > \beta_j$. In that case, we say that j *decreases* i and that j is a *decreaser* of i. Also, by Lemma 2.4.1, each element of α contains at least one decreasible exponent.

Now suppose that $f : \mathbf{N} \to \mathbf{N}$ is a function with the following property.

Property P: For all $i \in \mathbf{N}$, if i is decreasible, then it is decreased by some $j \leq f(i)$.

We first show that given such an f we can compute (in f and β) an infinite descending sequence $(\sigma_i)_{i\in\mathbf{N}}$ in $\mathcal{X}$ as follows.

Step 0. Let i_0 be the least decreasible index of β, and let j_0 be the least decreaser of i_0. By Lemma 2.4.1, β_{i_0} must be an exponent of α_0, i.e. $t(i_0) = 0$, so we can find i_0 by just taking the least decreasible $i^* < lh(\alpha_0)$. Notice that we can decide whether i^* is decreasible by just inspecting β up to the index $f(i^*)$, since f has the property P.
We set $\sigma_0 = \beta_{j_0}$ and observe that $p(i) \geq p(i_0)$ for each decreasible $i > i_0$. Suppose otherwise as witnessed by i^*, and let j^* be the least decreaser of i^*. By definition of decreaser, $p(j^*) = p(i^*) < p(i_0)$ and $\beta_{i^*} > \beta_{j^*}$. However i_0 is the least decreasible β-index of exponents of α_0, hence the β-index of $e_{p(i^*)}(\alpha_0)$ will never decrease due to $p(i^*) < p(i_0)$, so $\beta_{i^*} = \beta_z$, where $z = \theta(0, p(i^*))$. Therefore, $\beta_z > \beta_{j^*}$ and $p(z) = p(i^*) = p(j^*)$, but in that case z would be decreasible (by j^*) and $p(z) < p(i_0)$, which implies $z < i_0$ since $t(z) = t(i_0) = 0$, thus contradicting the minimality of i_0.

Step $s + 1$. Suppose i_s, j_s, σ_s are defined, $(\sigma_t)_{t\leq s}$ is decreasing in $\mathcal{X}$ and $p(i) \geq p(i_s)$ for each decreasible $i > i_s$.
Let i_{s+1} be the least decreasible index of β larger than or equal to j_s, and let j_{s+1} be the least decreaser of i_{s+1}. By Lemma 2.4.1 and the fact that no decreasible $i > i_s$ can have $p(i) < p(i_s) = p(j_s)$, $\beta_{i_{s+1}}$ must be an exponent of $\alpha_{t(j_s)}$, namely the leftmost whose β-index is decreasible. So we can find i_{s+1} by just taking the least decreasible $i^* \in [j_s,\ j_s + lh(\alpha_{t(j_s)}) - p(j_s))$. Notice that we can decide whether i^* is decreasible by just inspecting β up to the index $f(i^*)$, since f has the property P.
We set $\sigma_{s+1} = \beta_{j_{s+1}}$. As we noted above, $\beta_{i_{s+1}}$ must be either β_{j_s} or an exponent of $\alpha_{t(j_s)}$ on the right of β_{j_s}, i.e. $t(i_{s+1}) = t(j_s)$ and $p(i_{s+1}) \geq p(j_s)$, so $\beta_{j_s} \geq \beta_{i_{s+1}}$. Then, $\sigma_s > \sigma_{s+1}$ because $\sigma_s = \beta_{j_s} \geq \beta_{i_{s+1}} > \beta_{j_{s+1}} = \sigma_{s+1}$. Finally, we observe that the last part of the inductive invariant is guaranteed as well, since $p(i) \geq p(i_{s+1})$ for each decreasible $i > i_{s+1}$. Suppose otherwise as witnessed by i^*, and let j^* be the least decreaser of i^*. By definition of decreaser, $p(j^*) = p(i^*) < p(i_{s+1})$ and $\beta_{i^*} > \beta_{j^*}$. However $\beta_{i_{s+1}}$ is the leftmost exponent of $\alpha_{t(j_s)}$ whose β-index is decreasible, then $\beta_{i^*} = \beta_z$, where $z = \theta(t(j_s), p(i^*))$. Hence, $\beta_z > \beta_{j^*}$, but in that case z would be decreasible (by j^*) and $p(z) = p(i^*) < p(i_{s+1})$, which implies $z < i_{s+1}$ since $t(z) = t(j_s) = t(i_{s+1})$, thus contradicting the minimality of i_{s+1}.

We now show how to obtain a function satisfying the property P from a solution of $\lambda\mathsf{regHT}^{=n}[\mathsf{ap}]$ for a suitable colouring. The argument is similar to the proof of Theorem 3.3.8.
Define $c : \mathbf{N} \to \mathbf{N}$ as follows: $c(x) =$ the unique $i < \lambda(x)$ satisfying the following conditions:

1. There exists j such that $\lambda(x) \leq j < \mu(x)$ and β_j is the least decreaser of β_i, and

2. For all j' such that $j < j' < \mu(x)$, if $\beta_{j'}$ is the least decreaser of $\beta_{i'}$ then $i' \geq \lambda(x)$.

If no such i exists, we set $c(x) = 0$.

The function c is computable in α and λ-regressive. Let $H = \{h_1 < h_2 < h_3 < \dots\}$ be an apart solution to $\lambda\mathsf{regHT}^{=n}$ for c. The following Claim ensures the existence of an $(\alpha \oplus H)$-computable function with Property P.

Claim 7. *For each $h_k \in H$ and each decreasible $i < \lambda(h_k)$, there exists $j < \mu(h_{k+n-1})$ such that j decreases i.*

Proof of Claim 7. Assume by way of contradiction that there is some $h_k \in H$ and some $i < \lambda(h_k)$ such that i is decreasible but not by any $j < \mu(h_{k+n-1})$.
Let b be such that if $i' < \lambda(h_k)$ is decreasible, then there exists $j' < b$ decreasesing i'. The existence of b can be proved in RCA_0 using the following instance of strong Σ^0_1-bounding (similarly as in the proof of Theorem 3.3.8):

$$\forall n \, \exists b \, (\forall i' < n) \, [\exists j'(j' \text{ decreases } i') \to (\exists j' < b)(j' \text{ decreases } i')].$$

Since H is infinite, there is an $h_{k'} \in H$ such that $h_{k'} > h_{k+n-1}$ and $\mu(h_{k'}) \geq b$. Then, by min-term-homogeneity, $c(h_k + \dots + h_{k+n-1}) = c(h_k + \dots + h_{k+n-2} + h_{k'})$. But by choice of h_k, $h_{k'}$ and the definition of c, we can show that $c(h_k + \dots + h_{k+n-1}) \neq c(h_k + \dots + h_{k+n-2} + h_{k'})$, yielding a contradiction.
To see this, we reason as follows. First observe that, by apartness of H, the following identities hold:

$$\lambda(h_k + \dots + h_{k+n-1}) = \lambda(h_k + \dots + h_{k+n-2} + h_{k'}) = \lambda(h_k),$$

and

$$\mu(h_k + \dots + h_{k+n-2} + h_{k'}) = \mu(h_{k'}).$$

Let $j \in [\mu(h_{k+n-1}), \mu(h_{k'}))$ be the least decreaser of i. Such a j exists for, by hypothesis, i is decreasible but not by any $j < \mu(h_{k+n-1})$, and by choice of h'_k, the least decreaser of i must be smaller than $\mu(h_{k'})$, since $i < \lambda(h_k)$.
First note that $c(h_k + \dots + h_{k+n-2} + h_{k'})$ cannot be 0, since this occurs if and only if there is no $i^* < \lambda(h_k)$ such that for some $j^* \in [\lambda(h_k), \mu(h_{k'}))$, j^* decreases i^*; but this is is false by choice of h_k and $h_{k'}$.
If $c(h_k + \dots + h_{k+n-1})$ takes some non-zero value $i^* < \lambda(h_k)$, then this same value cannot be taken by $c(h_k + \dots + h_{k+n-2} + h_{k'})$ under our assumptions. If it were, it would mean that i^* is decreased for the first time by some $j^* < \mu(h_{k'})$ that is also the maximal least decreaser below $\mu(h_{k'})$ of some q with $q < \lambda(h_k)$. This is impossible since the least decreaser of $i^* = c(h_k + \dots + h_{k+n-1})$ occurs earlier in the sequence β than the least decreaser of i since, by the definition of c, it must be that $j^* < \mu(h_k + \dots + h_{k+n-1})$ and the latter value, by apartness, equals $\mu(h_{k+n-1})$, as noted above. On the other hand, j is in $[\mu(h_{k+n-1}), \mu(h_{k'}))$, so that $j^* < j$. Thus j^* cannot be the maximal least decreaser below $\mu(h_{k'})$ of some $q < \lambda(h_k)$, as required by the definition of c, since j is such a least decreaser of i, and $i < \lambda(h_k)$.
This proves the Claim. ☐

Now it is sufficient to observe that the $(\alpha \oplus H)$-computable function f defined as follows has the Property P: on input n, pick the least k such that $\sum_{1 \leq n' \leq n} lh(\alpha_{n'}) < \lambda(h_k)$ and let $f(n)$ be the α-index of the $\mu(h_{k+n-1})$-th element in the sequence β of all components appearing in α, i.e., $f(n) = t(\mu(h_{k+n-1}))$. That this choice of f satisfies Property P is implied by Claim 7 above. This concludes the proof of the theorem. ☐

Chapter 4

The strength of Well-Orderings

As explained in Sec. 1.1.4, the outcome of ordinal analysis of a theory T carried out using well-ordering principles is typically a specific linear order operator f such that $\mathsf{WOP}(\mathcal{X} \mapsto f(\mathcal{X}))$ is equivalent to T over a weaker theory, while the Π^1_1-ordinal of T is the least transfinite ordinal whose existence can *not* be proved in T. Therefore, adding $\mathsf{WOP}(\mathcal{X} \mapsto f(\mathcal{X}))$ to T does never result in a system stronger than T, which instead is always strictly included in $T' = T + \mathrm{WO}(ord(T))$. Then, it is natural to wonder what is the strength of T', and – more generally – what is the strength of theories augmented with $\mathrm{WO}(\alpha)$, where α is at least their proof-theoretic ordinal. Even for weak theories like RCA_0, the answer to this question is not trivial and occasionally some confusion arises.[1] Even the standard argument for showing that ω^ω is an upper bound on the proof-theoretic ordinal of RCA_0 is somewhat indirect in that it hinges on the characterization of the provably recursive functions of RCA_0 rather than only on the computation of its Π^1_1-ordinal. A proper direct treatment approach to determining the proof-theoretic ordinal of theories of the form $\mathsf{RCA}_0 + \mathrm{WO}(\alpha)$ seems to be missing from the literature. The closest match is Sommer's [Som] model-theoretical treatment of first-order theories with transfinite induction restricted to various formula-classes and ordinals strictly below ε_0.
Therefore, in this Chapter – based on a joint work with Michael Rathjen – we fill an apparent gap in the literature by showing that, if δ is an ordinal satisfying $\omega \cdot \delta = \delta$, the Π^1_1-ordinal of the theory $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$ is δ^ω. Examples of relevant δ's are ω^ω, ω^{ω^ω}, etc. and ε_0. Such a result provides an answer to a more general question than just the strength of $\mathsf{RCA}_0 + \mathrm{WO}(ord(\mathsf{RCA}_0))$, allowing us to investigate the result of iterated strengthening of RCA_0 via well-orderings. Moreover, it also gives an ordinal analysis of several interesting theorems (see, e.g., [Sim1, HS, KY]) and even of Ramsey-related principles (e.g., [Fri1]), all of which being equivalent over RCA_0 to $\mathrm{WO}(\delta)$ for some $\delta = \omega \cdot \delta$.

The intuition behind our proof is the following. First, we switch to a first order theory so that we can make use of standard methods to carry out ordinal analysis: thus, starting from $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$, we construct the theory $\mathbf{T}_\delta$ by discarding the

[1] For example, in proving that a Π^1_1-version of Ramsey's Theorem called the Adjacent Ramsey Theorem is equivalent to $\mathrm{WO}(\varepsilon_0)$ over RCA_0, Lemma 2.2 in [Fri1] makes use of the false equivalence, over RCA_0, between $\mathrm{WO}(\varepsilon_0)$ and the Π^1_1-soundness of ACA_0. The presentation in the later [FP] avoids this pitfall but establishes a slightly different result.

second order axiom Δ^0_1-CA (which is not needed for our purposes) and by replacing both $\mathsf{I}\Sigma_1$ and $\mathrm{WO}(\delta)$ with a suitable first order axiom schema, namely the transfinite induction up to δ restricted to Π_1-formulas[2], that can be formalized as follows.

Definition 4.0.1 (Transfinite induction)**.** *Let* Γ *be a class of formulas and let* $\prec$ *be a primitive recursive linear ordering isomorphic to some transfinite ordinal* α*. Then the* transfinite induction up to α for Γ formulas *– in symbols,* $\mathrm{TI}(\alpha, \Gamma)$ *– is the following axiom schema:*

$$\forall x\,[\forall y\,(y \prec x \rightarrow \varphi(y)) \rightarrow \varphi(x)] \rightarrow \forall x\,\varphi(x)$$

for each $\varphi \in \Gamma$*.*

Then, we give an ordinal analysis of the theory $\mathbf{T}_\delta$, and finally we show that such theory proves the same Π^1_1-statements as $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$, where – in the first order context of $\mathbf{T}_\delta$ – a Π^1_1-sentence $\forall X F(X)$ with $F(X)$ arithmetic is expressed by the formula F in which every occurrence of the clause $x \in X$ is replaced by $\mathsf{U}(x)$, where U is a generic unary predicate symbol added to the language of the theory. However, in order to formalize well-orderings within $\mathbf{T}_\delta$, we do not translate Definition 1.1.10 using U: instead, we resort to the statement $Fund_\prec$ defined as follows:

$$\forall x\,[\forall y\,(y \prec x \rightarrow \mathsf{U}(y)) \rightarrow \mathsf{U}(x)] \rightarrow \forall x \mathsf{U}(x), \tag{4.1}$$

where $\prec$ is some ordering definable in the theory.
Therefore, by proof-theoretic ordinal of a first order theory T, here we mean the least transfinite ordinal α isomorphic to some primitive recursive linear ordering $\prec$ such that T does not prove $Fund_\prec$. Then, in Sec. 4.2, we eventually show that from the proof-theoretic ordinal (based on $Fund_\prec$) of $\mathbf{T}_\delta$ we can derive the Π^1_1-ordinal of our target theory, that is, the proof-theoretic ordinal (based on Definition 1.1.10) of $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$.
Ordinal analysis of $\mathbf{T}_\delta$ is carried out using the (one-sided) sequent calculus – a quite common choice in this field – since we need to work on derivations in $\mathbf{T}_\delta$ (and in other theories defined above) using methods made possible by this approach. Especially, we make use of the well-known technique of *cut elimination*, that consists in recasting a proof by removing all applications of the so-called *cut rule*, that is:

$$\frac{\Gamma, A \qquad \Gamma', \neg A}{\Gamma, \Gamma'} \tag{4.2}$$

This idea traces back directly to Gentzen's formalization of sequent calculus and it is widely used in Proof Theory, as any formal system admitting cut elimination (meaning that any derivation in that system can be obtained with no applications of the cut rule) can be easily proved to be consistent. In ordinal analysis, cut elimination is useful because several methods have been developed to extract an upper bound on the proof-theoretic ordinal of a theory starting from the length of specific *cut-free* derivations of certain formulas.

[2] Recall from Chapter 1 that the superscript 0 is omitted in the notation for classes in the arithmetical hierarchy when dealing with $\mathcal{L}_1$-formulas.

However, cut elimination easily applies only to pure logic: when we add axioms to formalize our theory, instead, cut elimination becomes "problematic". A typical workaround, then, is to perform partial cut elimination, that is, only cuts of formulas whose complexity is higher than the complexity of the axioms are removed, which is exactly what we do to our derivations in $\mathbf{T}_\delta$. Still, cuts of less complex formulas need to be eliminated in order to apply the aforementioned techniques, as they only work on cut-free derivations. Therefore, following a typical approach, we "embed" $\mathbf{T}_\delta$ into an infinitary theory PA_ω, i.e. a theory containing a rule with infinitely many premises, called ω-rule, that can be formalized as follows.

$$\frac{\Gamma, F(\bar{n}) \quad \text{for all } n}{\Gamma, \forall x\, F(x)} \tag{4.3}$$

where $\bar{n}$ denote the n-th numeral, that is, the term obtained by adding n times the successor function symbol to the term $\bar{0}$ for zero.
This rule can then replace axioms with a complexity too high for our purposes (like $\mathrm{TI}(\delta, \Pi_1)$), but it also turns deductions into infinite objects, for it requires to derive all its infinitely many premises before being applied. However, this way we can eventually obtain cut-free derivations[3], so that we can finally extract an upper bound for the proof-theoric ordinal of $\mathbf{T}_\delta$ and (by Π^1_1-conservativity) of $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$ as well, with the only constraint that $\omega \cdot \delta = \delta$.

Then, for the remainder of this Chapter, we fix an ordinal δ such that $\omega \cdot \delta = \delta$. The ordinal δ is assumed to be represented in a natural primitive recursive ordinal representation system. Also, we denote by $<_\delta$ the primitive recursive linear ordering on the ordinals smaller than δ to distinguish it from the usual ordering on the naturals.

4.1 Ordinal analysis of $\mathbf{T}_\delta$

The theory $\mathbf{T}_\delta$ is formalized using the language $\mathcal{L}_1$ augmented by the symbols for all the primitive recursive relations plus a unary predicate symbol U. Bounded quantifiers $\forall x \leq t$ and $\exists x \leq t$ are treated as quantifiers in their own right – that is, as specific constructions not to be intended as abbreviations for other formulas. Formulas containing only bounded quantifiers (or no quantifiers at all) are called Δ_0-*formulas.*
For our proof-theoretic purposes, derivations in $\mathbf{T}_\delta$ are formalized in a one sided sequent calculus, using negation normal forms following [Sch3] (this is also known as the Tait-calculus [Tai]).
For any term t, if t is closed (i.e., it does not contain free variables and can then be evaluated to a number), we denote by $t^{\mathbf{N}}$ the number n such that t evaluates to n (in the following we occasionally refer to $t^{\mathbf{N}}$ by t).
The axioms of $\mathbf{T}_\delta$ are sequents of two kinds. Let Γ be a finite set of formulas of $\mathbf{T}_\delta$.

[3]Actually, at this step, there still might be cuts of atomic formulas, so we need to take them into account (see Sec. 4.1.3).

(i) Let $R(t_1, \ldots, t_r)$ be an atomic closed formula, where R is a relation symbol for a primitive recursive relation $R^{\mathbf{N}}$. If $R^{\mathbf{N}}(t_1^{\mathbf{N}}, \ldots, t_r^{\mathbf{N}})$ is true, then

$$\Gamma, R(t_1, \ldots, t_r)$$

is an axiom. If $R^{\mathbf{N}}(t_1^{\mathbf{N}}, \ldots, t_r^{\mathbf{N}})$ is false, then

$$\Gamma, \neg R(t_1, \ldots, t_r)$$

is an axiom.

(ii) If $s^{\mathbf{N}} = t^{\mathbf{N}}$ holds for closed terms s and t, then

$$\Gamma, \mathsf{U}(s), \neg \mathsf{U}(t)$$

is an axiom.

The order relation for the ordering on δ is denoted by the same symbol $<_\delta$ used to denote the corresponding primitive recursive relation. Also, we fix some binary surjective coding function for pairs with inverses $(\,\cdot\,)_0\,,\,(\,\cdot\,)_1$.
Moreover, $\mathbf{T}_\delta$ includes the transfinite induction on δ for Π_1-formulas, which is expressed via the rule

$$\frac{\Gamma,\ \exists z\,((z)_0 <_\delta a \wedge \neg F((z)_1, (z)_0)),\ \forall x F(x,a)}{\Gamma,\ F(t,s)} \tag{4.4}$$

where t, s are arbitrary terms, Γ is an arbitrary finite set of formulas, $F(b,a)$ is Δ_0 and a is an eigenvariable, meaning that it does not occur free in the lower sequent. Intuitively, 4.4 asserts that $F(t,s)$ holds whenever the fact that $F(x,y)$ holds for any x and any $y <_\delta a$ implies $\forall x F(x,a)$.
Observe that Σ_1-induction is included in $\mathbf{T}_\delta$ as it follows from $\mathrm{TI}(\delta, \Pi_1)$, since $\mathsf{I}\Pi_1$ entails $\mathsf{I}\Sigma_1$ even over weak theories (e.g., see Prop. 6.1.5 in [DM]).
In order to perform partial cut eliminations, we define the rank of a formula A (and we denote it by $|A|$) as follows:

(i) if A is Δ_0, then $|A| = 0$; otherwise,

(ii) if $A = A_0 \wedge A_1$ or $A = A_0 \vee A_1$, then $|A| = \max(|A_0|, |A_1|) + 1$;

(iii) if $A = \forall x F(x)$ or $A = \exists x F(x)$, then $|A| = |F(0)| + 1$;

(iv) if $A = (\forall x \leq t)\, F(x)$ or $A = (\exists x \leq t)\, F(x)$, then $|A| = |F(0)| + 2$.

Note that point (iv) is needed due to our definition of bounded quantifiers within $\mathbf{T}_\delta$. As the rule (4.4) introduces a Δ_0-formula and the main formulas of the remaining axioms of $\mathbf{T}_\delta$ are Δ_0 as well, we can easily eliminate cuts of rank greater than 0. For any $m, k \in \omega$, we use the notation $\mathbf{T}_\delta \frac{\,m\,}{k}\, \Gamma$ to convey that Γ is deducible in $\mathbf{T}_\delta$ by a deduction of length at most m such that all cuts occurring in this deduction only act on formulas of a rank less than k. Thus $\mathbf{T}_\delta \frac{\,m\,}{1}\, \Gamma$ means that all cut formulas (if any) in the deduction of Γ are Δ_0-formulas.

Theorem 4.1.1. $\mathbf{T}_\delta \frac{\,n\,}{r+1}\, \Gamma \ \Rightarrow\ \exists m\, \mathbf{T}_\delta \frac{\,m\,}{1}\, \Gamma$.

Proof. By the usual cut elimination method of Gentzen's Hauptsatz. □

4.1.1 Embedding $\mathbf{T}_\delta$ in an infinitary system

Next we embed $\mathbf{T}_\delta$ into an infinitary system, called PA_ω, with ω-rule (basically the same as the system Z_∞ in [Sch3]; an explicit definition of PA_ω in a two-sided Gentzen calculus can be found in [Rat1]). The formulas of PA_ω are the closed formulas of $\mathbf{T}_\delta$, i.e. formulas without free variables. We assign a rank $|A|_\omega$ to a formula A of PA_ω as follows:

- $|A|_\omega = 0$ if A is atomic or a negated atom.
- $|A_0 \wedge A_1|_\omega = |A_0 \vee A_1|_\omega = \max(|A_0|_\omega, |A_1|_\omega) + 1$.
- $|(\exists x \leq t)\, F(x)|_\omega = |(\forall x \leq t)\, F(x)|_\omega = |F(0)|_\omega + 1$.
- $|\exists x\, F(x)|_\omega = |\forall x\, F(x)|_\omega = \max(\omega, |F(0)|_\omega + 1)$.

Note that $|A|_\omega < \omega$ exactly when A is Δ_0, and $|\exists x\, F(x)|_\omega = |\forall x\, F(x)|_\omega = \omega$ when $F(0)$ is Δ_0.

The axioms of PA_ω are the same as $\mathbf{T}_\delta$, but the transfinite induction rule (4.4) is discarded.
However, PA_ω includes the ω-rule as defined in 4.3, that is: if $\Gamma, F(\bar{n})$ is deducible for all n, then $\Gamma, \forall x\, F(x)$ holds.
Similarly to derivations in $\mathbf{T}_\delta$, we will use the notation $\mathsf{PA}_\omega \vert\frac{\alpha}{\beta}\, \Gamma$ to convey that Γ is deducible in PA_ω by a deduction of height at most α using only cuts of formulas with $|\cdot|_\omega$-rank $< \beta$.
However, in this case, α and β are ordinal numbers, hence we need to define the following operation (formalized as in [Poh], Exercise 3.3.14). Recall that any ordinal α has a unique normal form (see [Poh], Theorem 3.3.8), that is, there exist uniquely determined ordinals $\alpha_0, \ldots, \alpha_n$ such that $\alpha = \alpha_0 + \cdots + \alpha_n$; in that case, we write $\alpha =_{\mathrm{NF}} \alpha_0 + \cdots + \alpha_n$.

Definition 4.1.2 (Hessenberg sum)**.** *Let $\alpha =_{\mathrm{NF}} \alpha_0 + \cdots + \alpha_m$ and $\beta =_{\mathrm{NF}} \beta_0 + \cdots + \beta_n$. We set $\alpha \# \beta = \gamma_0 + \cdots + \gamma_p$, where $\gamma_0, \ldots, \gamma_p$ are the ordinals $\alpha_0, \ldots, \alpha_m, \beta_0, \ldots, \beta_n$ enumerated in non-increasing order, that is, $\gamma_i \geq \gamma_{i+1}$ for all $i < p = m + n$.*

Notice that the Hessenberg sum is commutative. Then, we can state the following Lemma.

Lemma 4.1.3 (Reduction Lemma)**.** *If $|B|_\omega \leq \omega$, $\mathsf{PA}_\omega \vert\frac{\alpha}{\omega}\, \Gamma, B$ and $\mathsf{PA}_\omega \vert\frac{\beta}{\omega}\, \Gamma, \neg B$, then $\mathsf{PA}_\omega \vert\frac{\alpha\#\beta}{\omega}\, \Gamma$.*

Proof. See [Sch3], p. 882. ∎

Now we can "embed" $\mathbf{T}_\delta$ into PA_ω.

Theorem 4.1.4 (Embedding Theorem)**.** *If $\mathbf{T}_\delta \vert\frac{m}{1}\, \Gamma$, then $\mathsf{PA}_\omega \vert\frac{\delta m}{\omega}\, \Gamma^*$, where Γ^* is the result of assigning closed terms to all free variables in Γ (the same term to the same variable).*

Proof. We proceed by induction on m. We only need to pay attention to the case where the last inference is an instance of the rule (4.4), since it is not included in PA_ω; all the other cases trivially derive from the definition of rank in $\mathbf{T}_\delta$ and in PA_ω. So let $\Gamma = \Theta, F(t,s)$ and assume

$$\mathbf{T}_\delta \,\big|\!\frac{m_0}{1}\, \Theta, \exists z\, ((z)_0 <_\delta a \wedge \neg F((z)_1,(z)_0)), \forall x\, F(x,a)$$

for some $m_0 < m$.
Let $*$ be an assignment. Inductively we have that, for all closed terms q:

$$\mathsf{PA}_\omega \,\big|\!\frac{\delta^{m_0}}{\omega}\, \Theta^*, \exists z\, ((z)_0 <_\delta q \wedge \neg F^*((z)_1,(z)_0)), \forall x\, F^*(x,q) \tag{4.5}$$

For any α in the field of $<_\delta$, we can show by transfinite induction that:

$$\mathsf{PA}_\omega \,\big|\!\frac{\delta^{m_0}\cdot\omega\cdot(\alpha+1)}{\omega}\, \Theta^*, \forall x\, F^*(x,\bar\alpha) \tag{4.6}$$

By the induction hypothesis, we have:

$$\mathsf{PA}_\omega \,\big|\!\frac{\delta^{m_0}\cdot(\omega\cdot(\eta+1))}{\omega}\, \Theta^*, F^*(s,\bar\eta)$$

for every $\eta <_\delta \alpha$ and arbitrary closed term s', yielding

$$\mathsf{PA}_\omega \,\big|\!\frac{\delta^{m_0}\cdot(\omega\cdot\alpha)+1}{\omega}\, \Theta^*, \neg(\bar\eta <_\delta \bar\alpha) \vee F^*(s',\bar\eta)$$

via an inference $(\vee)$. If r is a closed term such that $r^{\mathbf{N}}$ is different from all η preceding α, then $\neg(r <_\delta \bar\alpha)$ is an axiom, and thus, via an inference $(\vee)$, we arrive at $\mathsf{PA}_\omega \,\big|\!\frac{1}{0}\, \Theta^*, \neg(r <_\delta \bar\alpha) \vee F^*(s',r)$. Thus from the above we conclude that

$$\mathsf{PA}_\omega \,\big|\!\frac{\delta^{m_0}\cdot(\omega\cdot\alpha)+1}{\omega}\, \Theta^*, \neg((\bar k)_0 <_\delta \bar\alpha) \vee F^*((\bar k)_1,(\bar k)_0)$$

holds for all k, so that, via an application of the ω-rule, we get:

$$\mathsf{PA}_\omega \,\big|\!\frac{\delta^{m_0}\cdot(\omega\cdot\alpha)+2}{\omega}\, \Theta^*, \forall z\, (\neg((z)_0 <_\delta \bar\alpha) \vee F^*((z)_1,(z)_0)). \tag{4.7}$$

Applying the Reduction Lemma 4.1.3 to (4.5) and (4.7) yields:

$$\mathsf{PA}_\omega \,\big|\!\frac{\delta^{m_0}\#(\delta^{m_0}\cdot(\omega\cdot\alpha))+2}{\omega}\, \Theta^*, \forall x\, F^*(x,\bar\alpha). \tag{4.8}$$

From (4.8) we finally get:

$$\mathsf{PA}_\omega \,\big|\!\frac{\delta^{m_0}\cdot(\omega\cdot(\alpha+1))}{\omega}\, \Theta^*, \forall x\, F^*(x,\bar\alpha)\,,$$

confirming (4.6).
If a term q has the property that $q^{\mathbf{N}}$ is not in the field of $<_\delta$ then one can directly infer from (4.5) that

$$\mathsf{PA}_\omega \,\big|\!\frac{\delta^{m_0}}{\omega}\, \Theta^*, \forall x\, F^*(x,q). \tag{4.9}$$

The reason for this is that if the formula $\exists z\, ((z)_0 <_\delta q \wedge \neg F^*((z)_1,(z)_0))$ figures as the main formula of an inference in this derivation its minor formula is of the form $(p)_0 <_\delta q \wedge \neg F^*((p)_1,(p)_0)$. The latter formula conjunctively contains a false atomic formula. Such a formula can always be erased from the derivation. Formally, of course, this has to be proved by a separate induction on the ordinal of the derivation. (4.6) and (4.9) now yield

$$\mathsf{PA}_\omega \,\big|\!\frac{\delta^{m}}{\omega}\, \Theta^*, F^*(t,s)$$

for all closed terms t and s, since $\omega\cdot(\alpha+1) <_\delta \delta$ on account of $\omega\cdot\delta = \delta$. $\square$

4.1.2 Eliminating cuts with Δ_0-formulas

The next step is to eliminate cuts with Δ_0-formulas that are not atomic.

Lemma 4.1.5. *Let* $0 < n < \omega$ *and suppose* $\mathsf{PA}_\omega \overset{\alpha}{\underset{n+1}{\vdash}} \Gamma$. *Then* $\mathsf{PA}_\omega \overset{\omega\cdot\alpha}{\underset{n}{\vdash}} \Gamma$.

Proof. We proceed by induction on α. The crucial case is when the last inference was a cut of rank n with cut formulas $A, \neg A$. Note that A is not an atomic formula. We then have $\mathsf{PA}_\omega \overset{\alpha_0}{\underset{n+1}{\vdash}} \Gamma, A$ and $\mathsf{PA}_\omega \overset{\alpha_0}{\underset{n+1}{\vdash}} \Gamma, \neg A$ for some $\alpha_0 < \alpha$. The induction hypotheses furnishes us with

$$\mathsf{PA}_\omega \overset{\omega\cdot\alpha_0}{\underset{n}{\vdash}} \Gamma, A \quad \text{and} \quad \mathsf{PA}_\omega \overset{\omega\cdot\alpha_0}{\underset{n}{\vdash}} \Gamma, \neg A \tag{4.10}$$

Let A be of the form $(\exists x \le t)\, F(x)$, the case $(\forall x \le t)\, F(x)$ being symmetric. Then $\neg A$ is the formula $(\forall x \le t)\, \neg F(x)$. From (4.10) we obtain

$$\mathsf{PA}_\omega \overset{\omega\cdot\alpha_0}{\underset{n}{\vdash}} \Gamma, F(\bar{0}), \ldots, F(\bar{p}) \quad \text{and} \quad \mathsf{PA}_\omega \overset{\omega\cdot\alpha_0}{\underset{n}{\vdash}} \Gamma, \neg F(\bar{k}) \tag{4.11}$$

for all $k \le p$, where p is the numerical value of t. As the formulas $F(\bar{k}), \neg F(\bar{k})$ have rank $< n$, we can employ $(p+1)$-many cuts to (4.11) to arrive at $\mathsf{PA}_\omega \overset{\omega\cdot\alpha_0+p+1}{\underset{n}{\vdash}} \Gamma$. Thus we have $\mathsf{PA}_\omega \overset{\omega\cdot\alpha}{\underset{n}{\vdash}} \Gamma$ as $\omega\cdot\alpha_0 + p + 1 < \omega\cdot\alpha$. A similar argument works when A is of either form $A_0 \wedge A_1$ or $A_0 \vee A_1$. □

Corollary 4.1.6. *If* $\mathsf{PA}_\omega \overset{\alpha}{\underset{\omega}{\vdash}} \Gamma$ *then* $\mathsf{PA}_\omega \overset{\omega^{\omega}\cdot\alpha}{\underset{1}{\vdash}} \Gamma$.

Proof. We use induction on α. The only interesting case arises when the last inference is a cut with a formula A of rank $k > 0$. Then we have $\mathsf{PA}_\omega \overset{\alpha_0}{\underset{\omega}{\vdash}} \Gamma, A$ and $\mathsf{PA}_\omega \overset{\alpha_0}{\underset{\omega}{\vdash}} \Gamma, \neg A$ for some $\alpha_0 < \alpha$. The induction hypothesis yields $\mathsf{PA}_\omega \overset{\omega^{\omega}\cdot\alpha_0}{\underset{1}{\vdash}} \Gamma, A$ and and $\mathsf{PA}_\omega \overset{\omega^{\omega}\cdot\alpha_0}{\underset{1}{\vdash}} \Gamma, \neg A$. Hence $\mathsf{PA}_\omega \overset{\omega^{\omega}\cdot\alpha_0+1}{\underset{k+1}{\vdash}} \Gamma$. Applying Lemma 4.1.5 k times we arrive at $\mathsf{PA}_\omega \overset{\omega^k\cdot(\omega^{\omega}\cdot\alpha_0+1)}{\underset{1}{\vdash}} \Gamma$. As $\omega^k \cdot (\omega^\omega \cdot \alpha_0 + 1) = \omega^\omega \cdot \alpha_0 + \omega^k \le \omega^\omega \cdot \alpha$ we also have $\mathsf{PA}_\omega \overset{\omega^{\omega}\cdot\alpha}{\underset{1}{\vdash}} \Gamma$ as desired. □

Note that $\delta \ge \omega^\omega$ since $\omega \cdot \delta = \delta$.

Corollary 4.1.7. *Let* $m > 0$. *If* $\mathsf{PA}_\omega \overset{\delta^m}{\underset{\omega}{\vdash}} \Gamma$ *then* $\mathsf{PA}_\omega \overset{\delta^{m+1}}{\underset{1}{\vdash}} \Gamma$.

Proof. Corollary 4.1.6 yields $\mathsf{PA}_\omega \overset{\omega^{\omega}\cdot\delta^m}{\underset{1}{\vdash}} \Gamma$. Thus the desired conclusion follows as $\omega^\omega \cdot \delta^m \le \delta \cdot \delta^m = \delta^{m+1}$. □

4.1.3 Upper bounds for the provable well-orderings of $\mathbf{T}_\delta$

The results of the previous section can be utilized to determine the ordinal rank of provable well-orderings of $\mathbf{T}_\delta$. Let $\prec$ be a primitive recursive ordering. $\prec$ is said to be a *provable well-ordering* of $\mathbf{T}_\delta$ if $\mathbf{T}_\delta$ proves that $\prec$ is a total linear ordering and that $Fund_\prec$ holds.

Assuming $\mathbf{T}_\delta \vdash Fund_\prec$, from Theorem 4.1.1, Theorem 4.1.4 and Corollary 4.1.7 we can derive

$$\mathsf{PA}_\omega \overset{\delta^m}{\underset{1}{\vdash}} \forall v[\forall u(u \prec v \to \mathsf{U}(u)) \;\to\; \mathsf{U}(v)] \;\to\; \forall v \mathsf{U}(v) \tag{4.12}$$

for some $m > 0$, where $F \to F'$ stands for $\neg F \vee F'$ for any formulas F, F'.

There are several ways of obtaining an upper bound for the order type of $\prec$ in terms of the length of a cut-free deduction of $Fund_\prec$ (see e.g. [Sch2, Theorem 23.1], [FS, Theorem 2.27]) which ultimately go back to Gentzen. Schütte [Sch2, Theorem 23.1] obtains particularly sharp bounds. He shows that the length α of a cut-free derivation of transfinite induction along an ordering $\prec$ (of which $Fund_\prec$ is an instance) provides an upper bound for the ordinal rank of $\prec$ if $\omega \cdot \alpha = \alpha$. However, in our case, we need to extract bounds from deductions that still have cuts with formulas $\mathsf{U}(s), \neg\mathsf{U}(s)$.[4] We could first eliminate these remaining cuts, yet we would get bounds of the form 2^{δ^m} (as in [Sch3], p. 882), and these would be too high for our purpose of showing that δ^ω is the proof-theoretic ordinal of $\mathbf{T}_\delta$. To overcome this obstacle, we draw on a technique proposed by Rathjen (e.g., see [Rat2]). We extend PA_ω by yet another infinitary rule $\text{Prog}_\prec$ due to Schütte [Sch1, p. 384], that is:

$$\frac{\Gamma, \mathsf{U}(\bar{m}) \ \text{ for all } m \prec n}{\Gamma, \mathsf{U}(s)} \tag{4.13}$$

whenever s is a closed term with value n.

We denote by PA^*_ω the extension of PA_ω with the rule $\text{Prog}_\prec$. Notice that all the results obtained in the previous sections for PA_ω also hold for PA^*_ω: indeed, all the arguments can be adapted *verbatim* to PA^*_ω, with the only exception of the proof of Lemma 4.1.3, which requires to be slightly extended (see [Sch3], p. 883).

Now, let $\text{PROG}_\prec$ be an abbreviation for $\forall v\,[\forall u\,(u \prec v \to \mathsf{U}(u)) \to \mathsf{U}(v)]$. The rule $\text{Prog}_\prec$ has the effect of making $\text{PROG}_\prec$ provable.

Lemma 4.1.8. $\mathsf{PA}^*_\omega \mid\!\frac{\alpha}{1}\, \Gamma, \neg\text{PROG}_\prec \ \ \Rightarrow\ \ \mathsf{PA}^*_\omega \mid\!\frac{3\cdot\alpha}{1}\, \Gamma\,.$

Proof. We proceed by induction on α. If $\neg\text{PROG}_\prec$ was not the main formula of the last inference then the desired result follows immediately by applying the inductive assumption to its premisses and subsequently reapplying the same inference. Thus suppose that $\neg\text{PROG}_\prec$ was the main formula of the last inference. Then, by the weakening rule, we have

$$\mathsf{PA}^*_\omega \mid\!\frac{\alpha_0}{1}\, \Gamma,\, \neg\text{PROG}_\prec, \forall u(u \prec s \to \mathsf{U}(u)) \,\wedge\, \neg\mathsf{U}(s) \tag{4.14}$$

for some $\alpha_0 <_\delta \alpha$ and some closed term s. The induction hypothesis yields

$$\mathsf{PA}^*_\omega \mid\!\frac{3\cdot\alpha_0}{1}\, \Gamma,\, \forall u(u \prec s \to \mathsf{U}(u)) \,\wedge\, \neg\mathsf{U}(s)\,. \tag{4.15}$$

Using inversion for $(\wedge)$, $(\forall)$ and $(\vee)$ we arrive at

$$\mathsf{PA}^*_\omega \mid\!\frac{3\cdot\alpha_0}{1}\, \Gamma,\, \neg\,(\bar{n} \prec s),\, \mathsf{U}(\bar{n}) \tag{4.16}$$

for all n, and

$$\mathsf{PA}^*_\omega \mid\!\frac{3\cdot\alpha_0}{1}\, \Gamma,\, \neg\mathsf{U}(s)\,. \tag{4.17}$$

[4] They may also contain cuts with formulas $R(t_1, \ldots, t_k), \neg R(t_1, \ldots, t_k)$, where R is a symbol for a primitive recursive predicate. But these are entirely harmless.

Since $\mathsf{PA}^*_\omega \mid\!\frac{0}{0}\, \Gamma, \bar{n} \prec s$ holds for all n with $n \prec s^{\mathbf{N}}$, we can apply cuts and the rule $\mathrm{Prog}_\prec$ to (4.16) to arrive at

$$\mathsf{PA}^*_\omega \mid\!\frac{3\cdot\alpha_0+2}{1}\, \Gamma,\ \mathsf{U}(s). \tag{4.18}$$

Applying Cut to (4.18) and (4.17) yields

$$\mathsf{PA}^*_\omega \mid\!\frac{3\cdot\alpha_0+3}{1}\, \Gamma \tag{4.19}$$

and hence

$$\mathsf{PA}^*_\omega \mid\!\frac{3\cdot\alpha}{1}\, \Gamma. \tag{4.20}$$

□

Corollary 4.1.9. *For all n, $\mathsf{PA}^*_\omega \mid\!\frac{\delta^m}{1}\, \mathsf{U}(\bar{n})$.*

Proof. Follows from (4.12) and Lemma 4.1.8. Note that $m > 0$. □

For a closed numerical term s we denote by $|s|_\prec$ the ordinal $\{|\bar{n}|_\prec \mid \bar{n} \prec s \text{ is true}\}$.

Proposition 4.1.10. *Assume that the sequent $\neg\mathsf{U}(t_1), \ldots, \neg\mathsf{U}(t_r), \mathsf{U}(s_1), \ldots, \mathsf{U}(s_q)$ is not an axiom and $s_1 \preceq \ldots \preceq s_q$ holds. Then*

$$\mathsf{PA}^*_\omega \mid\!\frac{\alpha}{1}\, \neg\mathsf{U}(t_1), \ldots, \neg\mathsf{U}(t_r), \mathsf{U}(s_1), \ldots, \mathsf{U}(s_q)$$

implies

$$|s_1|_\prec \quad < \quad \omega \cdot \alpha. \tag{4.21}$$

Proof. Let $\neg\mathsf{U}(\vec{t})$ be an abbreviation for $\neg\mathsf{U}(t_1), \ldots, \neg\mathsf{U}(t_r)$. In the above, we allow $r = 0$, in which case $\neg\mathsf{U}(\vec{t})$ is the empty sequent.
We proceed by induction on α. As the sequent is not an axiom, it must have been inferred. The only two possibilities are applications of $\mathrm{Prog}_\prec$ or cuts with atomic formulas.

Case 1: The last inference was $\mathrm{Prog}_\prec$. Then there is a term s_j and $\alpha_0 <_\delta \alpha$ such that $\mathsf{PA}^*_\omega \mid\!\frac{\alpha_0}{1}\, \neg\mathsf{U}(\vec{t}), \mathsf{U}(s_1), \ldots, \mathsf{U}(s_q), \mathsf{U}(\bar{n})$ for all $\bar{n} \prec s_j$. As $s_1 \preceq s_j$ this also holds for all $\bar{n} \prec s_1$. The induction hypothesis yields that

$$|\bar{n}|_\prec < \omega \cdot \alpha_0$$

holds for those $\bar{n} \prec s_1$ for which the sequent is not an axiom. By definition of PA^*_ω,

$$\neg\mathsf{U}(\vec{t}), \mathsf{U}(s_1), \ldots, \mathsf{U}(s_q), \mathsf{U}(\bar{n}) \tag{4.22}$$

is an axiom only if $\bar{n}$ has the same value as some $t_1, \ldots, t_r$; then there are only finitely many n for which (4.22) is an axiom. Thus $|s_1|_\prec < \omega \cdot \alpha_0 + \omega$, whence $|s_1|_\prec < \omega \cdot \alpha$.

Case 2: The last inference was a cut with cut formulas $\mathsf{U}(p), \neg\mathsf{U}(p)$, i.e., we have

$$\mathsf{PA}_\omega^* \mid\frac{\alpha_0}{1} \neg\mathsf{U}(\vec{t}), \mathsf{U}(s_1), \ldots, \mathsf{U}(s_q), \mathsf{U}(p) \tag{4.23}$$

$$\mathsf{PA}_\omega^* \mid\frac{\alpha_0}{1} \neg\mathsf{U}(\vec{t}), \mathsf{U}(s_1), \ldots, \mathsf{U}(s_q), \neg\mathsf{U}(p) \tag{4.24}$$

for some $\alpha_0 < \alpha$ and closed term p. If the sequent from (4.24) is not an axiom, the induction hypothesis applied to that derivation yields $|s_1|_\prec < \omega \cdot \alpha_0$. If it is an axiom, there is an s_j such that p and s_j evaluate to the same numeral, and hence $s_1 \preceq p$. In that case, the induction hypothesis applied to (4.23) yields $|s_1|_\prec < \omega \cdot \alpha_0$.

Case 3: The last inference was a cut with cut formulas $R(\vec{q}), \neg R(\vec{q})$ for a symbol R for a primitive recursive relation. Then we have

$$\mathsf{PA}_\omega^* \mid\frac{\alpha_0}{1} \neg\mathsf{U}(\vec{t}), \mathsf{U}(s_1), \ldots, \mathsf{U}(s_q), R(\vec{q}) \tag{4.25}$$

$$\mathsf{PA}_\omega^* \mid\frac{\alpha_0}{1} \neg\mathsf{U}(\vec{t}), \mathsf{U}(s_1), \ldots, \mathsf{U}(s_q), \neg R(\vec{q}) \tag{4.26}$$

for some $\alpha_0 < \alpha$. If $R(\vec{q})$ is true, it follows from (4.26) that we also have

$$\mathsf{PA}_\omega^* \mid\frac{\alpha_0}{1} \neg\mathsf{U}(\vec{t}), \mathsf{U}(s_1), \ldots, \mathsf{U}(s_q)$$

since $\neg R(\vec{q})$ is false and thus it can be ditched from the derivation of (4.26). Hence the induction hypothesis yields $|s_1|_\prec < \omega \cdot \alpha_0$.
Likewise, if $R(\vec{q})$ is false, it follows from (4.25) that we also have

$$\mathsf{PA}_\omega^* \mid\frac{\alpha_0}{1} \neg\mathsf{U}(\vec{t}), \mathsf{U}(s_1), \ldots, \mathsf{U}(s_q)$$

and hence the induction hypothesis yields $|s_1|_\prec < \omega \cdot \alpha_0$.
To formally prove both these cases, one should apply a further induction on the length of the derivation. □

Corollary 4.1.11. *The following implications hold:*

(i) $\mathsf{PA}_\omega^* \mid\frac{\alpha}{1} \mathsf{U}(s) \Rightarrow |s|_\prec < \omega \cdot \alpha$

(ii) $\mathsf{PA}_\omega \mid\frac{\beta}{1} \mathit{Fund}_\prec \Rightarrow |\prec| \leq \omega \cdot 3 \cdot \beta$, *where* $|\prec|$ *is the ordinal rank of* $\prec$.

Proof. (i) is an immediate consequence of Proposition 4.1.10, while (ii) follows from (i) and Lemma 4.1.8. □

In sum, it follows that the ordinal rank of $\prec$ is not larger than δ^m for $m \in \omega$, and hence δ^ω is an upper bound for the proof-theoretic ordinal of $\mathbf{T}_\delta$.
Proposition 4.1.10 can also be proved via techniques in A. Beckmann's dissertation, notably his [Bec, 5.2.5 Boundedness Theorem] that also features in [BP].
Turning to lower bounds, one can easily show, using external induction on n, that $\mathbf{T}_\delta \vdash \mathit{Fund}_{\prec_{\delta^n}}$. This is a folklore result; details can be found in [Som, Lemma 4.3] Finally, as a consequence of the results gathered so far, we have the following theorem.

Theorem 4.1.12. *The proof-theoretic ordinal of* $\mathbf{T}_\delta$ *is* δ^ω.

It remains to transfer this result to our target theory $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$.

4.2 Proof-theoretic ordinal of $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$

In this last section, we show that we can apply our results from the previous section to conclude that the ordinal of $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$ is δ^ω.
To prove such conservativity result, we proceed as follows. We start by showing that any $\mathcal{L}_2$-structure whose first order part is a model of $\mathbf{T}_\delta$ can be extended to a model of $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$. We follow the scheme of Simpson ([Sim2], IX.1). By writing that $\mathbf{M}_1$ is an ω-submodel of $\mathbf{M}_2$ we mean that $\mathbf{M}_1 = (M_1, \mathcal{S}_1)$ and $\mathbf{M}_2 = (M_1, \mathcal{S}_2)$ where $\mathcal{S}_1 \subseteq \mathcal{S}_2$. In other words, the two models share the same first order part M_1.

Lemma 4.2.1. *Let $\mathbf{M}$ be an $\mathcal{L}_2$-structure which satisfies the axioms of $\mathbf{T}_\delta$. Then $\mathbf{M}$ is an ω-submodel of some model of $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$.*

Proof. We first show that $\mathbf{M}$ can be extended to a model $\mathbf{M}'$ satisfying RCA_0 and $\mathrm{TI}(\delta, \Delta^0_0)$ with the same first-order domain as $\mathbf{M}$. Then we show that such an extension also satisfies $\mathrm{WO}(\delta)$.
The ω-extension $\mathbf{M}'$ is defined exactly as in Simpson, Lemma IX.1.8, i.e., the second-order part is given by the Δ^0_1-definable sets of the base model $\mathbf{M}$. By Lemma IX.1.8 of [Sim2] we have that $\mathbf{M}'$ satisfies RCA_0.
Then, in order to check that $\mathrm{TI}(\delta, \Delta^0_0)$ is also satisfied, we use the first claim in Simpson's Lemma IX.1.8. Let φ be a Σ^0_0 formula with no free set variables and parameters in M'. Then, there exists a Π^0_1-formula φ_Π with the same free variables and parameters only in M such that φ and φ_Π are equivalent over M'. Thus, $\mathrm{TI}(\delta, \Pi^0_1)$ in $\mathbf{M}$ implies $\mathrm{TI}(\delta, \Delta^0_0)$ in $\mathbf{M}'$.
Finally, we show that M' also satisfies $\mathrm{WO}(\delta)$. Since δ is a linear ordering by definition and $\mathrm{WF}(\delta)$ is equivalent to $\mathrm{ME}(\delta)$ by Theorem 1.1.12, we can just show that $\mathrm{ME}(\delta)$ holds. Assume otherwise, letting S be a set witnessing $\neg\mathrm{ME}(\delta)$: then, we would have that $\bar{S}$, i.e. the complement of S, witnesses the failure of an instance of $\mathrm{TI}(\delta, \Delta^0_0)$. More precisely: suppose that S is non-empty and has no $<_\delta$-minimal element. Then $\exists x(x \in S)$. On the other hand, $\bar{S}$ is in M' (since any model of RCA_0 is closed under Turing reducibility hence under complement) and $\forall x(\forall y(y <_\delta x \to y \in \bar{S}) \to x \in \bar{S})$. Suppose in fact that for some x, $\forall y(y <_\delta x \to y \in \bar{S})$ but $x \in S$. Then all $y <_\delta x$ are not in S but x is in S and thus x is the minimum of S, contra our hypothesis. $\square$

Then we can prove the following conservativity result.

Lemma 4.2.2. *If $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$ proves $\forall X F(X)$ with $F(X)$ arithmetic, then $\mathbf{T}_\delta$ proves the formula F' obtained from $F(X)$ by replacing expressions of the form $t \in X$ with $\mathsf{U}(t)$.*

Proof. By way of contradiction, assume $\neg F'$ holds in some $\mathcal{L}_2$-model $\mathbf{M}$ of $\mathbf{T}_\delta$. Then we can extend $\mathbf{M}$ to an $\mathcal{L}_2$-model $\mathbf{M}'$ for $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$ as in the proof of Lemma 4.2.1. Since the second order part of $\mathbf{M}'$ contains all the Δ^0_1-definable sets, there exists X in $\mathbf{M}'$ such that $x \in X$ if and only if $\mathsf{U}(x)$ for all x. Hence $\neg F(X)$ holds, contradicting our hypothesis. $\square$

Now we show that from the proof-theoretic ordinal of $\mathbf{T}_\delta$, obtained in the previous section and based on the concept of $Fund_\prec$, we can derive the Π^1_1-ordinal of the theory $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$. First, we translate $Fund_\prec$ accordingly to the statement of Lemma 4.2.2.

Definition 4.2.3. *For any ordering $\mathcal{X}$, we define $Fund(\mathcal{X})$ as follows:*

$$\forall Y[\forall y(\forall z(z <_{\mathcal{X}} y \rightarrow z \in Y) \rightarrow y \in Y) \wedge \forall y(y \in Y)]$$

Then, we have the following result.

Lemma 4.2.4. $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$ *does not prove $Fund(\delta^\omega)$.*

Proof. From Theorem 4.1.12 and Lemma 4.2.2. ☐

Now, it only remains to show that $\mathrm{WF}(\mathcal{X})$ implies $Fund(<_{\mathcal{X}})$ over RCA_0, for any ordering $\mathcal{X} = (X, \mathcal{X})$.

Lemma 4.2.5. *Over* RCA_0*,* $\mathrm{WF}(\mathcal{X})$ *implies $Fund(\mathcal{X})$ for any ordering $\mathcal{X}$.*

Proof. Suppose $\mathrm{WF}(\mathcal{X})$ but $\neg Fund(\mathcal{X})$. Then, we have:

$$\exists Y[\forall y(\forall z(z <_{\mathcal{X}} y \rightarrow z \in Y) \rightarrow y \in Y) \wedge \exists y(y \notin Y)]$$

Therefore $\bar{Y}$, i.e., the complement of Y, is non-empty. Since

$$\forall y(\forall z(z <_{\mathcal{X}} y \rightarrow z \in Y) \rightarrow y \in Y)$$

holds by assumption, then for any $y \in \bar{Y}$ there exist $z \in \bar{Y}$ s.t. $z <_{\mathcal{X}} y$. Hence, we can use $\bar{Y}$ to recursively define an infinite descending sequence in $\mathcal{X}$, thus contradicting $\mathrm{WF}(\mathcal{X})$. ☐

Lemma 4.2.6. $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$ *does not prove* $\mathrm{WF}(\delta^\omega)$*.*

Proof. By way of contradiction, assume that we can derive $\mathrm{WF}(\delta^\omega)$ in $\mathsf{RCA}_0+\mathrm{WO}(\delta)$. Then, by Lemma 4.2.5, we have $\mathsf{RCA}_0 + \mathrm{WO}(\delta) \vdash Fund(\delta^\omega)$, hence contradicting Lemma 4.2.4. ☐

Theorem 4.2.7. *The proof-theoretic ordinal of* $\mathsf{RCA}_0 + \mathrm{WO}(\delta)$ *is δ^ω.*

Proof. The upper bound is given by Lemma 4.2.6. The lower bound follows from the observation that, for each n, $\mathsf{RCA}_0 + \mathrm{WO}(\delta) \vdash \mathrm{WO}(\delta^n)$. The proof, which we omit, is analogous to the proof that $\mathsf{RCA}_0 \vdash \mathrm{WO}(\omega^n)$ for each n. ☐

Conclusion

We investigated different topics concerning Ramsey-theoretic principles and well orderings, providing several original results. After giving an introductory overview (Chapter 1) about Proof Theory, Reverse Mathematics and Computability Theory, in Chapter 2 we studied the relations between Ramsey's Theorem and the so-called *well-ordering principles*: first, we better formalized former results due to Carlucci and Zdanowski, then we extended the argument to Ramsey's Theorem for large sets, hence answering a question left open in [CZ1]. We provided both implications over RCA_0 and Weihrauch reductions, giving a method that is likely applicable to stronger well-ordering principles as well. We also showed how the weakest restriction of Hindman's Theorem known to be equivalent to ACA_0 – i.e., $\mathsf{HT}^{=3}_2[\mathsf{ap}]$ – Weihrauch-reduces $\mathsf{WOP}(\mathcal{X} \to \boldsymbol{\omega}^{\mathcal{X}})$ – that is, the well-ordering principle for base-ω exponentiation, which is known to be equivalent to ACA_0.

Then, in Chapter 3, we analyzed Ramsey-like principles for infinitely many colours. Starting from a corollary of the Canonical Ramsey's Theorem (namely, its application to regressive colourings), we formulated and studied an analogous variation of Hindman's Theorem: in particular, we defined the class of λ-regressive colourings and we applied Taylor's Theorem (that is, the "canonical" version of Hindman's Theorem) to this class of functions, hence obtaining a novel principle – that we called Regressive Hindman's Theorem – with some interesting features. First of all, it preserves various properties of Hindman's Theorem and Taylor's Theorem: it self-improves to its own version with apartness, it implies the standard Hindman's Theorem over RCA_0 and it can be parameterized by limiting the length of the sums in the solution, so as to produce a hierarchy of bounded versions of the full principle. These versions, in turn, were analyzed, highlighting how their (apparently) weakest version – i.e., the Regressive Hindman's Theorem for sums of exactly 2 elements – endowed with apartness is already equivalent to ACA_0. This fact results in a nice symmetry with Ramsey's Theorem: the regressive version of the latter, indeed, is equivalent to ACA_0 even when applied to colourings of pairs, while the standard Ramsey's Theorem requires triples, which is exactly what happens with Hindman's Theorem and its regressive version (both with apartness), whose equivalence with ACA_0 is obtained, respectively, for sums of 3 elements and sums of 2 elements. Finally, we mentioned that a Weihrauch reduction from such restriction of Regressive Hindman's Theorem and $\mathsf{WOP}(\mathcal{X} \to \boldsymbol{\omega}^{\mathcal{X}})$ can be found: this last result highlights how the use of regressive functions can help in simplifying arguments about Hindman's Theorem, hence giving an additional tool in the research for the optimal bounds on this kind of principles.

In the last Chapter, then, we answered a natural question that can arise when dealing with well-orderings, especially while using them in ordinal analysis: namely, we measured the strength of the family of systems $\mathsf{RCA}_0 + \mathrm{WO}(\alpha)$, where $\alpha = \omega \cdot \alpha$. Several theorems were proved equivalent to some of these systems: therefore, besides answering the original question, we also gave an ordinal analysis of all such principles.

Finally, different lines of research can be derived from the results presented here: first, one could try to extend the Weihrauch reductions from Ramsey's principles to stronger well-ordering principles. Moreover, various open questions about Hindman's Theorem can be shifted to its regressive version, so as to investigate (for instance) its optimal bounds or the strength of the restricted versions without apartness, hence possibly giving a better insight even into the original Hindman's principle as well. Lastly, it may be interesting to relate the Regressive Hindman's Theorem with Hindman-type variants of the Thin Set Theorem (recently studied by Hirschfeldt and Reitzes in [HR]), since they all deal with infinitely many colours.

Bibliography

[AR] B. Afshari, M. Rathjen. Reverse mathematics and well-ordering principles: A pilot study. *Annals of Pure and Applied Logic*, 160(3):231–237, 2009.

[Bec] A. Beckmann. *Separating Fragments of Bounded Arithmetic*. PhD thesis, Universität Münster, 1996.

[BP] A. Beckmann, W. Pohlers. Applications of cut-free infinitary derivations to generalized recursion theory. *Annals of Pure and Applied Logic*, 94:7–19, 1995.

[BHS] A. R. Blass, J. L. Hirst, S. G. Simpson. Logical analysis of some theorems of combinatorics and topological dynamics. In *Logic and combinatorics*, volume 65 of *Contemporary Mathematics*, pages 125–156. American Mathematical Society, 1987.

[BR] V. Brattka, T. Rakotoniaina. On the uniform computational content of Ramsey's Theorem. *The Journal of Symbolic Logic*, 4:1278–1316, 2017.

[Car1] L. Carlucci. Weak Yet Strong restrictions of Hindman's Finite Sums Theorem. *Proceedings of the American Mathematical Society*, pages 819–829, 2018.

[Car2] L. Carlucci. Restrictions of Hindman's Theorem: an overview. In *Connecting with Computability: CiE 2021*, volume 12813 of *Lecture Notes in Computer Science*, pages 94–105. Springer, 2021.

[CKLZ] L. Carlucci, L. A. Kołodziejczyk, F. Lepore, K. Zdanowski. New bounds on the strength of some restrictions of Hindman's Theorem. *Computability*, 9:139–153, 2020.

[CM] L. Carlucci, L. Mainardi. Regressive Versions of Hindman's Theorem. *Archive for Mathematical Logic*, 2024.

[CMR] L. Carlucci, L. Mainardi, M. Rathjen. A Note on the Ordinal Analysis of $\mathsf{RCA}_0 + \mathrm{WO}(\sigma)$. In *Computability in Europe 2019: Computing with Foresight and Industry, 15-19 Jul 2019, Durham, UK*, Lecture Notes in Computer Science, pages 144–155. Springer Cham, 2019.

[CMZ] L. Carlucci, L. Mainardi, K. Zdanowski. Reductions of well-ordering principles to combinatorial theorems. Submitted. Available at `https://arxiv.org/abs/2401.04451`.

[CZ1] L. Carlucci, K. Zdanowski. A note on Ramsey Theorems and Turing Jumps. *How the World Computes. CiE 2012*, 7318:89–95, 2012.

[CZ2] L. Carlucci, K. Zdanowski. The strength of Ramsey's Theorem for coloring relatively large sets. *Journal of Symbolic Logic*, 79:89–102, 2014.

[CDH$^+$] B. F. Csima, D. D. Dzhafarov, D. R. Hirschfeldt, C. G. Jockusch, R. Solomon, L. B. Westrick. The Reverse Mathematics of Hindman's Theorem for sums of exactly two elements. *Computability*, 8:253–263, 2019.

[DDH$^+$] F. G. Dorais, D. Dzhafarov, J. L. Hirst, J. P. Mileti, P. Shafer. On uniform relationships between combinatorial problems. *Transactions of the American Mathematical Society*, 368(2):1321–1359, 2016.

[DH] D. D. Dzhafarov, J. L. Hirst. The polarized Ramsey's theorem. *Archive for Mathematical Logic*, 48(2):141–157, 2011.

[DJSW] D. D. Dzhafarov, C. G. Jockusch, R. Solomon, L. B. Westrick. Effectiveness of Hindman's Theorem for bounded sums. In *Proceedings of the International Symposium on Computability and Complexity (in honour of Rod Downey's 60th birthday)*, volume 10010 of *Lecture Notes in Computer Science*, pages 134–142. Springer, 2016.

[DM] D. D. Dzhafarov, C. Mummert. *Reverse Mathematics. Problems, Reductions, and Proofs*. Theory and Applications of Computability. Springer Cham, 2022.

[ER] P. Erdős, R. Rado. A combinatorial theorem. *Journal of the London Mathematical Society*, 25:249–255, 1950.

[FN] V. Farmaki, S. Negrepontis. Schreier Sets in Ramsey Theory. *Transactions of the American Mathematical Society*, 360(2):849–880, 2008.

[Fri1] H. Friedman. Adjacent Ramsey Theory. Unpublished draft, August 2010, available online on Harvey Friedman's website `https://u.osu.edu/friedman.8/`.

[Fri2] H. Friedman. Some systems of second order arithmetic and their use. In *Proceedings of the International Congress of Mathematicians (Vancouver, 1974)*, volume 1, pages 235—-242, 1975.

[Fri3] H. Friedman. Systems of second-order arithmetic with restricted induction (1-2). *Journal of Symbolic Logic*, 2:551—-560, 1976.

[FMW] H. Friedman, A. Montalbán, A. Weiermann. A characterization of ATR_0 in terms of a Kruskal-like tree theorem. Unpublished draft.

[FP] H. Friedman, F. Pelupessy. Independence of Ramsey theorem variants using ε_0. *Proceedings of the American Mathematical Society*, 144:853–860, 2016.

[FS] H. Friedman, S. Sheard. Elementary descent recursion and proof theory. *Annals of Pure and Applied Logic*, 71:1–45, 1995.

[Gir] J. Girard. *Proof Theory and Logical Complexity.* Biblipolis, 1987.

[GR] R. L. Graham, B. L. Rothschild. Ramsey's theorem for n-parameter sets. *Transactions of the American Mathematical Society*, 159:257–292, 1971.

[HS] K. Hatzikiriakou, S. G. Simpson. Reverse mathematics, Young diagrams, and the ascending chain condition. *Journal of Symbolic Logic*, 82:576–589, 2017.

[Hin1] N. Hindman. The existence of certain ultrafilters on N and a conjecture of Graham and Rothschild. *Proceedings of the American Mathematical Society*, 36(2):341–346, 1972.

[Hin2] N. Hindman. Finite sums from sequences within cells of a partition of N. *Journal of Combinatorial Theory Series A*, 17:1–11, 1974.

[HLS] N. Hindman, I. Leader, D. Strauss. Open problems in partition regularity. *Combinatorics Probability and Computing*, 12:571–583, 2003.

[Hir] D. R. Hirschfeldt. *Slicing the Truth (On the Computable and Reverse Mathematics of Combinatorial Principles)*, volume 28 of *Lecture Notes Series.* Institute for Mathematical Sciences, National University of Singapore, 2014.

[HJ] D. R. Hirschfeldt, C. G. Jockusch. On notions of computability-theoretic reduction between Π^1_2 principles. *Journal of Mathematical Logic*, 16(1):1278–1316, 2016.

[HR] D. R. Hirschfeldt, S. C. Reitzes. Thin Set Versions of Hindman's Theorem. *Notre Dame Journal of Formal Logic*, 63(4):481–491, 2022.

[Hir1] J. L. Hirst. *Combinatorics in Subsystems of Second Order Arithmetic.* PhD thesis, The Pennsylvania State University, 1987.

[Hir2] J. L. Hirst. Reverse mathematics and ordinal exponentiation. *Annals of Pure and Applied Logic*, 66(1):1–18, 1994.

[Hir3] J. L. Hirst. Hilbert vs. Hindman. *Archive for Mathematical Logic*, 51(1-2):123–125, 2012.

[Joc] C. G. Jockusch. Ramsey's theorem and recursion theory. *Journal of Symbolic Logic*, 37:268–280, 1972.

[KM] A. Kanamori, K. McAloon. On Gödel incompleteness and finite combinatorics. *Annals of Pure and Applied Logic*, 33(1):23–41, 1987.

[KY] A. Kreuzer, K. Yokoyama. On principles between Σ_1- and Σ_2-induction, and monotone enumerations. *Journal of Mathematical Logic*, 16, 2016.

[Liu] J. Liu. RT^2_2 does not imply WKL_0. *The Journal of Symbolic Logic*, 77:609–620, 2012.

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

[MM] A. Marcone, A. Montalbán. The Veblen functions for computability theorists. *The Journal of Symbolic Logic*, 76(2):575–602, 2011.

[McA] K. McAloon. Paris-Harrington incompleteness and progressions of theories. In *Recursion Theory*, volume 42 of *Proceedings of Symposia in Pure Mathematics*, pages 447–460. American Mathematical Society, 1985.

[Mil] J. Mileti. The canonical Ramsey's Theorem and computability theory. *Transactions of the American Mathematical Society*, 160:1309–1340, 2008.

[Pat] L. Patey. The weakness of being cohesive, thin or free in reverse mathematics. *Israel Journal of Mathematics*, 216:905–955, 2016.

[Poh] W. Pohlers. *Proof Theory: The First Step into Impredicativity.* Universitext. Springer Berlin Heidelberg, 2008.

[PR] P. Pudlák, V. Rödl. Partition theorems for systems of finite subsets of integers. *Discrete Mathematics*, 39(1):67–73, 1982.

[Rat1] M. Rathjen. The art of ordinal analysis. In *Proceedings of the International Congress of Mathematicians, Madrid, August 22–30, 2006*, pages 45–69. European Mathematical Society, 2006.

[Rat2] M. Rathjen. The Higher Infinite in Proof Theory. In *Logic Colloquium '95: Proceedings of the Annual European Summer Meeting of the Association of Symbolic Logic, held in Haifa, Israel, August 9–18, 1995*, Lecture Notes in Logic, pages 275–304. Cambridge University Press, 2017.

[RW] M. Rathjen, A. Weiermann. *Reverse Mathematics and Well-ordering Principles*, pages 351–370. World Scientific, 2011.

[Sch1] K. Schütte. Beweistheoretische Erfassung der unendlichen Induktion in der Zahlentheorie. *Mathematische Annalen*, 122:369–389, 1951.

[Sch2] K. Schütte. *Proof Theory.* Grundlehren der mathematischen Wissenschaften. Springer, 1977.

[Sch3] H. Schwichtenberg. Proof Theory: Some Applications of Cut-Elimination. In J. Barwise, editor, *Handbook of Mathematical Logic*, volume 90 of *Studies in Logic and the Foundations of Mathematics*, pages 867–895. Elsevier, 1977.

[SS] D. Seetapun, T. A. Slaman. On the Strength of Ramsey's Theorem. *Notre Dame Journal of Formal Logic*, 36(4):570–582, 1995.

[Sim1] S. Simpson. Ordinal numbers and the Hilbert basis theorem. *Journal of Symbolic Logic*, 53:961–974, 1988.

[Sim2] S. Simpson. *Subsystems of Second Order Arithmetic*. Perspectives in Logic. Cambridge University Press, second edition, 2009.

[Som] R. Sommer. Transfinite induction within Peano Arithmetic. *Annals of Pure and Applied Logic*, 76(3):231–289, 1995.

[Tai] W. W. Tait. Normal derivability in classical logic. In J. Barwise, editor, *The semantics and syntax of infinitary languages*, pages 204–236. Springer, 1968.

[Tay] A. D. Taylor. A canonical partition relation for finite subsets of ω. *Journal of Combinatorial Theory (A)*, 21(2):137–146, 1976.

[Tur] A. Turing. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42(1):230–265, 1936.

[Wei1] K. Weihrauch. *The degrees of discontinuity of some translators between representations of the real numbers*. Informatik-Berichte. FernUniversität in Hagen, 1992.

[Wei2] K. Weihrauch. *The TTE-Interpretation of Three Hierarchies of Omniscience Principles*. Informatik-Berichte. FernUniversität in Hagen, 1992.