



École doctorale n° 532 MSTIC

Thèse de doctorat en Informatique

présentée par

Quentin LE HOUÉROU

sous la direction de

Julien CERVELLE et Ludovic PATEY

The metamathematics of Ramsey-like theorems

Jury :

Olivier BOURNEZ	École Polytechnique	Président du jury
Julien CERVELLE	Université Paris-Est Créteil	Directeur de thèse
Laura FONTANELLA	Université Paris-Est Créteil	Examinatrice
Denis HIRSCHFELDT	University of Chicago	Rapporteur
Leszek KOŁODZIEJCZYK	University of Warsaw	Rapporteur
Régine LALEAU	Université Paris-Est Créteil	Examinatrice
Ludovic PATEY	Université Paris-Cité	Directeur de thèse

Abstract en français

Cette thèse a pour objectif d'étudier le théorème de Ramsey et plusieurs de ses variantes sous un angle métamathématique. On y met principalement en évidence la nature chaotique de la théorie de Ramsey en prouvant des résultats de séparation, à travers un large éventail de techniques provenant principalement de la calculabilité et de la théorie de la preuve.

Du point de vue de la calculabilité, la force et la faiblesse calculatoire de certains de ces théorèmes sont étudiées, qu'il s'agisse de leur incapacité à coder le problème de l'arrêt ou de l'existence de solutions probabilistes.

Du point de vue de la théorie de la preuve, de nouveaux encadrements des conséquences au premier ordre du théorème de Ramsey pour les paires sont obtenus, notamment à travers l'étude de modèles non standards de l'arithmétique du second ordre ainsi que des liens entre les versions finitaires et infinitaires du théorème.

On y prouve en particulier plusieurs résultats de conservation, et l'on caractérise les motifs dont l'évitement est équivalent au théorème de Ramsey.

Abstract in English

This thesis aims to study Ramsey's theorem and several of its variants from a metamathematical perspective. It primarily highlights the chaotic nature of Ramsey theory by proving separation results, using a wide range of techniques drawn mainly from computability theory and proof theory.

From the standpoint of computability theory, the computational strength and weakness of some of these theorems are investigated, whether in terms of their inability to encode the halting problem or the existence of probabilistic solutions.

From the standpoint of proof theory, new bounds on the first-order consequences of Ramsey's theorem for pairs are obtained, notably through the study of nonstandard models of second-order arithmetic as well as the connections between the finitary and infinitary versions of the theorem.

In particular, several conservation results are proved, and the patterns whose avoidance is equivalent to Ramsey's theorem are characterized.

Remerciements

Tout d'abord, un immense merci à Leszek Kołodziejczyk et Denis Hirschfeldt d'avoir accepté d'être rapporteurs et d'avoir consacré tant de temps à la relecture détaillée de mon manuscrit. Merci aussi aux autres membres du jury, Olivier Bournez, Laura Fontanella et Régine Laleau d'avoir accepté d'en faire partie. Merci à Benoît Barbot et Laurent Bienvenu d'avoir été membres de mon CSI et pour leurs mots encourageants.

Merci à l'école doctorale MSTIC de m'avoir permis d'aller à Leeds et Amsterdam rencontrer les autres chercheurs de ma communauté et merci au LACL de prioriser les doctorants malgré un budget resserré.

Un immense merci à Ludovic et Julien, mes directeurs de thèse. À Ludovic pour avoir réussi, par je ne sais quel miracle, à jongler avec trois doctorants et une vie de famille tout en trouvant malgré tout autant de temps pour travailler avec moi, et ce, malgré la distance qui nous séparait. Certaines personnes choisissent un encadrant de thèse dans le but d'étudier un domaine précis, de travailler avec un grand chercheur ou parce qu'elles savent qu'elles seront bien encadrées ; avec Ludovic, j'ai eu l'immense chance de tout avoir en un. Et merci à Julien pour les passionnantes discussions me rappelant que la logique mathématique ne se résume pas à l'étude du théorème de Ramsey.

Merci à tous les professeurs que j'ai eus au cours de mon parcours académique, chacun ayant apporté sa pierre à la personne que je suis aujourd'hui. Merci tout particulièrement à Pascal Koiran de m'avoir fait découvrir ce merveilleux domaine qu'est la calculabilité en L3 et de m'avoir recommandé le nom de Ludovic pour un premier stage de recherche, qui a été pour moi une éclaircie entre les divers confinements. Merci aussi à Paul-Elliot Anglès d'Auriac d'avoir consacré autant de son temps à me co-encadrer durant ce stage et de m'avoir partagé sa passion pour les mathématiques à rebours.

Merci à Keita Yokoyama qui a tout fait pour que mon séjour à Sendai se passe bien et merci aussi à Daichi, Ethan, Hiroyuki et Yudai ainsi qu'à toutes les autres personnes que j'y ai rencontrées et qui m'ont fait sentir moins seul si loin de chez moi.

Merci à tous mes collègues du LACL qui ont fait de ce laboratoire un endroit où il fait bon travailler. Merci aux doctorants, Abdel, Ada, Adrien, Charles, Enzo, Fatehmeh, Hong-Linh, Manpreet, Noé, Olivier, Rémi, Théo, et Yahia qui ont su apporter de la vie (parfois trop) dans les locaux (étroits) du laboratoire. Merci à Ahmed et William qui m'ont précédé dans l'étude sans fin des variantes du théorème de Ramsey et avec qui j'ai pu en découvrir certains recoins. Merci aux chercheurs du LACL, Antoine, Benoît, Daniele, Emmanuel, Julien, Laura, Luidnel et tous les autres de m'avoir fait découvrir les coulisses du métier de chercheur ainsi que leurs travaux grâce à toutes les discussions que j'ai pu avoir avec eux pendant les pauses. Un immense merci aussi à Flore et Nicolas de toujours être là pour aider les pauvres

petits doctorants en difficulté.

Merci à toutes les merveilleuses personnes que j'ai eu l'occasion de rencontrer au cours de mes études à Lyon et à Paris, Alexis, Arthur, Clémence, Coralie, Guilhem, Olivier, Pauline et tant d'autres dont j'oublie parfois les prénoms, mais jamais le plaisir que j'ai à les fréquenter.

Enfin, un immense merci à toute ma famille. À mes parents, Stéphane et Stéphanie, à qui je dois beaucoup (notamment le fait d'être né, mais pas que) et qui ont toujours été là pour moi. À mon petit frère, Ewen, à qui je souhaite le meilleur pour sa vie et ses études. Merci également à mes oncles, tantes, cousins, cousines et grands-parents pour l'intérêt qu'ils ont porté à mes travaux et leurs encouragements, même si je n'ai pas toujours pu leur expliquer ce que je faisais : Béatrice, Camille, Carole, Didier, Elise, Gabin, Gérard, Guillaume, Julien, Laurent, Henriette, Joseph, Mireille, Sophie, Valérie, Thibaut, Titouan, merci à vous tous. Merci enfin à mes grands-parents, Danièle et Martial, qui ont grandement contribué à l'amour que je porte pour les mathématiques et dont j'aurais tant aimé qu'ils puissent être encore là pour partager ce moment avec eux.

Et merci au contribuable français ainsi qu'à toutes les personnes que j'ai oublié de mentionner.

Contents

Remerciements	4
1 Introduction	9
1.1 Introduction en français	9
1.2 Introduction in English	12
I State of the art	17
2 Hilbert's program	19
2.1 Historical background	19
2.2 Primitive recursive arithmetic	22
3 First-order arithmetic and fragments	27
3.1 Arithmetical hierarchy	27
3.2 First-order PRA	28
3.3 Induction	29
3.4 Collection	31
3.5 Bounded comprehension	32
3.6 Well-foundedness	32
4 Computability	35
4.1 Relativized computability and arithmetical sets	36
4.2 PA degrees	39
5 Second-order arithmetic	43
5.1 Relativized arithmetical and analytical hierarchies	43
5.2 Henkin models	45
5.3 Reverse mathematics	47
5.4 Semantic proofs	50
6 Ramsey's theorem	53
6.1 First-order part	55
6.2 Ramsey-like theorems	56
II Computational content of theorems	61
7 Separable permutations	63
7.1 Introduction	63
7.2 Separable permutations and ADS	67
7.3 Basis for separable permutations	69
7.4 The example of $\text{coSRT}_2^2(120)$	69
7.5 General case	73
7.6 Open questions	76

8	Theorems for the Schreier barrier	79
8.1	Introduction	79
8.2	Colorings of all finite sets	81
8.3	Colorings of exactly ω -large sets	83
8.4	FS^{ω} and TS^{ω} code $\mathcal{O}^{(\omega)}$	87
8.5	RRT^{ω} avoids cones	90
8.6	Conclusions and perspectives	95
III	First-order part	99
9	Non-standard models	101
9.1	Computability and definability in non-standard models	102
9.2	Π_1^1 -conservation	105
10	Π_1^1-conservation and well-foundedness	109
10.1	Introduction	109
10.2	Conservation over $\text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$	110
10.3	Conservation over $\text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$	118
10.4	Π_1^1 -conservation of RT_2^2	128
11	Bounded Ramsey theorem	131
11.1	$\text{B}\Sigma_2^0$ does not imply BRT_2^2	132
11.2	ADS does not imply BRT_2^2	134
IV	Partial conservation	143
12	Cuts	145
12.1	Cuts and Π_2 -conservation	146
12.2	Largeness and $\forall\Pi_3^0$ -conservation	147
12.3	Largeness(θ) and $\forall\Pi_4^0$ -conservation	151
13	α-largeness	155
13.1	Largeness below ϵ_0	156
13.2	Largeness below ω^ω	164
14	$\forall\Pi_4^0$-conservation of RT_2^2	173
15	$\forall\Pi_4^0$-conservation of OVW	179
15.1	Introduction	179
15.2	Hales-Jewett and Graham-Rothschild theorems	181
15.3	Ordered variable word theorem	183

Chapter 1

Introduction

1.1 Introduction en français

De nombreux théorèmes mathématiques peuvent être vus comme des problèmes, affirmant que toute instance admet une solution (tout corps possède une clôture algébrique, toute suite bornée de réels admet une sous-suite convergente, etc.). La difficulté à décrire explicitement cette solution à partir de l'instance peut alors être vue comme une mesure de la puissance du théorème.

Lorsque le théorème concerne des objets finis, cette question relève de la théorie de la complexité : pour une instance de taille n , combien d'étapes de calcul sont nécessaires pour produire une solution ?

Lorsqu'il s'agit, en revanche, d'objets qui ne peuvent pas nécessairement être décrits par un nombre fini de symboles (nombres réels, espaces topologiques, etc.), on peut alors souvent se restreindre au cas dénombrable (où ils peuvent alors être décrits par une suite infinie de zéros et de uns, ou, de manière équivalente, par un ensemble d'entiers). On entre ainsi dans le domaine de la calculabilité.

Tous les ensembles d'entiers ne sont pas aussi complexes. Il y a les ensembles finis et plus généralement les ensembles dits calculables, qui, même s'ils peuvent être infinis, sont néanmoins descriptibles par un algorithme, donc par une suite finie de symboles. C'est le cas de la plupart des ensembles rencontrés en mathématiques, comme l'ensemble des nombres premiers ou encore la suite des décimales de π . Toutefois, il existe aussi des ensembles non calculables, qui constituent en réalité l'immense majorité des ensembles d'entiers.

Une tentation possible est alors de restreindre sa pratique mathématique uniquement aux objets calculables, pour des raisons philosophiques ou pratiques. Une telle approche reviendrait cependant à se priver d'une grande partie des mathématiques modernes, puisque de nombreux théorèmes impliquent nécessairement l'existence d'objets non calculables. Par exemple, il existe une suite calculable de réels dans $[0, 1]$ telle qu'aucune de ses sous-suites convergentes ne soit calculable, donc le théorème de Bolzano–Weierstrass requiert l'existence de tels objets pour être vrai.

Il est possible, grâce aux outils de la calculabilité, de hiérarchiser ces ensembles non calculables (certains ensembles peuvent être vus comme plus incalculables que d'autres) et, par là même, de classer les théorèmes mathématiques selon la complexité calculatoire des objets qu'ils mettent en jeu.

La question de la puissance calculatoire des théorèmes mathématiques est étroitement liée à celle de leur puissance axiomatique. Les mathématiques à rebours (en anglais *reverse mathematics*) constituent une branche de la logique mathématique, initiée par Harvey Friedman, dont l'objectif est de déterminer les axiomes minimaux nécessaires pour démontrer un théorème donné.

L'idée des mathématiques à rebours consiste à se placer dans un système axiomatique volontairement restreint, dans lequel seule l'existence des ensembles

calculables peut être démontrée. Dans un tel cadre, les théorèmes faisant intervenir des ensembles non calculables ne sont plus démontrables (sans pour autant être réfutables, puisque l'on ne suppose pas l'inexistence des ensembles non calculables).

On peut ensuite étudier quelles hypothèses supplémentaires sont nécessaires pour établir ces théorèmes. La méthode consiste alors, après avoir démontré un théorème à partir d'axiomes plus forts, à inverser le raisonnement : on cherche à montrer que le théorème lui-même permet de retrouver ces axiomes, montrant par là qu'ils étaient nécessaires à la preuve originale.

Émerge alors un étonnant phénomène de structure : la quasi-totalité des théorèmes classiques se répartissent en cinq grandes classes d'équivalence modulo ces mathématiques calculables. La puissance axiomatique de chacune de ces classes correspond à une certaine conception philosophique des mathématiques et à un niveau de puissance calculatoire.

Le théorème de Ramsey pour les paires (dans sa version infinie) constitue l'un des premiers contre-exemples naturels à ce phénomène de structuration. Se pose alors la question du positionnement des autres résultats de la théorie de Ramsey, notamment en ce qui concerne les généralisations, les affaiblissements ou les variations du théorème de Ramsey.

Plutôt que de s'organiser en une sixième grande classe d'équivalence, ces différents énoncés forment au contraire une sorte de « zoo » de classes nouvelles : certaines plus fortes, d'autres plus faibles, et d'autres encore incomparables entre elles.

L'un des objectifs de cette thèse est l'étude de ces énoncés (notamment du point de vue calculatoire ou des mathématiques à rebours), en particulier en ce qui concerne les méthodes permettant de séparer deux théorèmes, c'est-à-dire de montrer que l'un ne découle pas de l'autre. S'il est souvent « facile » de montrer qu'un résultat est conséquence d'un autre en donnant une preuve, démontrer l'absence de preuve est en revanche généralement bien plus délicat.

L'un des principaux outils mobilisés pour obtenir ce type de résultats repose sur l'utilisation de modèles non standards de l'arithmétique. Il s'agit de structures qui ressemblent, à bien des égards, aux entiers naturels standards $\{0, 1, 2, \dots\}$, tout en s'en écartant sur certains points spécifiques. On peut les rapprocher, par analogie, des modèles de géométries non euclidiennes, qui ont permis de montrer que le postulat des parallèles ne découle pas des autres axiomes de la géométrie euclidienne, ou encore de l'exemple du coq déplumé, utilisé par Diogène pour mettre en évidence les limites de la définition par Platon de l'Homme comme « bipède sans plumes ».

1.1.1 Structure de la thèse et contributions

Cette thèse se divise en quatre parties.

Partie I : Cette partie est un état de l'art. Sous le fil directeur du programme de Hilbert introduit au Chapitre 2, y sont présentés divers fragments de l'arithmétique du premier ordre (au Chapitre 3) ainsi que des théories de l'arithmétique du second ordre et les mathématiques à rebours (au Chapitre 5). Le Chapitre 4 introduit diverses notions de calculabilité ainsi que leurs liens avec la définissabilité en arithmétique. Enfin, le théorème de Ramsey, ses propriétés calculatoires et vis-à-vis des mathématiques à rebours sont introduits au Chapitre 6.

Partie II : Cette partie porte sur l'étude de théorèmes d'un point de vue calculatoire. Le Chapitre 7 introduit de nouveaux outils en théorie algorithmique de l'aléatoire et s'intéresse à la caractérisation d'une classe particulière de permutations, les permutations séparables, comme motifs les plus difficiles à éviter dans certaines variantes du théorème de Ramsey. Les résultats présentés sont issus de l'article :

Q. Le Houérou and L. Patey “Ramsey-like theorems for separable permutations.” preprint (2025).

Le Chapitre 8 étudie la généralisation à des coloriage de tuples de taille non bornée de trois variantes du théorème de Ramsey possédant une forme de faiblesse combinatoire. Il y est montré que cette faiblesse ne subsiste que pour l'une d'entre elles, le théorème de Ramsey arc-en-ciel. Les résultats présentés sont issus de l'article :

L. Carlucci, O. Gjetaj, Q. Le Houérou, and L. Levy Patey “Ramsey-like theorems for the Schreier barrier.” *The Journal of Symbolic Logic*, page 1–29, 2025.

Partie III : Cette partie regroupe divers résultats obtenus par construction de nouveaux modèles par la méthode des ω -extensions. Le Chapitre 9 propose un état de l'art de l'utilisation d'outils issus de la théorie de la calculabilité au sein de modèles non standards de l'arithmétique. On y présente le schéma général de construction de nouveaux modèles par ω -extension, ainsi que son utilisation pour établir des résultats de conservation pour les formules du premier ordre.

Le Chapitre 10 fournit une nouvelle borne supérieure à l'ensemble des conséquences au premier ordre du théorème de Ramsey pour les paires. Les résultats présentés sont issus de l'article :

Q. Le Houérou, L. Levy Patey, and K. Yokoyama “Conservation of Ramsey's theorem for pairs and well-foundedness.” *Trans. Amer. Math. Soc.*, 378(3):2157–2186, 2025.

Le Chapitre 11 étudie une version déséquilibrée du théorème de Ramsey, dans laquelle les cliques monochromatiques d'une certaine couleur sont de taille bornée. Les résultats présentés sont issus de l'article :

Q. Le Houérou and L. Patey. “The reverse mathematics of bounded Ramsey's theorem for pairs.” *The Journal of Symbolic Logic*, pages 1–26, 2026.

Partie IV : Cette dernière partie explore les liens entre des versions finitaires et infinitaires de théorèmes de la théorie de Ramsey, en particulier en vue de l'obtention de résultats de conservation partiels. Le Chapitre 12 présente la notion d'ensemble α -large due à Ketonen et Solovay [72], et présente de nouvelles techniques de construction de modèles non standard par coupure basées sur ces ensembles. Ces techniques sont issues de l'article :

Q. Le Houérou, L. Levy Patey, and K. Yokoyama “ Π_4^0 conservation of Ramsey's theorem for pairs.” *J. London Math. Soc.*, 113: e70419. (2026)

Le Chapitre 13 présente des résultats d'intérêt essentiellement combinatoires en donnant des bornes optimales à des versions finitaires du théorème de Ramsey et du principe des tiroirs portant sur les ensembles α -larges. Les résultats présentés sont principalement issus de l'article :

Q. Le Houérou and L. Patey “Partition theorems for Ketonen-Solovay largeness.” preprint (2026).

Le Chapitre 14 exploite ce formalisme de construction de coupures, ainsi que les bornes obtenues au chapitre précédent, afin d'établir la caractérisation des conséquences au premier ordre du théorème de Ramsey pour les paires, pour une large classe de formules, obtenue dans l'article « Π_4^0 conservation of Ramsey's theorem for pairs ». La construction présentée y simplifie celle de l'article original.

Le Chapitre 15 applique ce même formalisme à des théorèmes portant sur des coloriage d'arbres afin d'en obtenir des résultats de conservation partiels. Les résultats présentés sont issus de l'article :

Q. Le Hou  rou and L. Levy Patey “ Π_1^0 conservation of the Ordered Variable Word theorem.” *The Journal of Symbolic Logic* 2025:1–16.

Par souci de bri  vet  , ainsi que dans la volont   de proposer un document formant un ensemble coh  rent, certains r  sultats obtenus au cours de ces trois ann  es n’ont pu   tre mentionn  s que tr  s bri  vement, voire ne figurent pas dans cette th  se. C’est notamment le cas de r  sultats portant sur la comparaison de la taille des preuves entre diverses th  ories, obtenus dans l’article :

Q. Le Hou  rou and L. Patey. “Largeness notions and polytime translation for $\forall\Sigma_3^0$ -consequences of RT_2^2 .” preprint (2026).

Le formalisme des interpr  tations de forcing introduit par Avigad [4] et tel qu’utilis   par Ko  łodziejczyk et al. [78], de m  me que la combinatoire mobilis  e dans cet article,   taient trop lourds    introduire dans le cadre de cette th  se et le r  sultat y est simplement mentionn  .

Il en va   galement ainsi de la comparaison des diff  rentes versions du principe des tiroirs le long de la hi  rarchie arithm  tique obtenue dans :

Q. Le Hou  rou, A. Mimouni and L. Levy Patey. “The reverse mathematics of the pigeonhole hierarchy.” preprint (2024).

Ces r  sultats ne s’ins  raient pas naturellement au sein des parties de cette th  se et reposaient aussi sur un formalisme lourd.

1.2 Introduction in English

Many mathematical theorems can be viewed as problems asserting that every instance admits a solution (every field has an algebraic closure, every bounded sequence of real numbers has a convergent subsequence, etc.). The difficulty of explicitly describing such a solution from the instance may then be regarded as a measure of the strength of the theorem.

When the theorem concerns finite objects, this question belongs to complexity theory: for an instance of size n , how many computational steps are required to produce a solution?

When dealing instead with objects that cannot necessarily be described by a finite number of symbols (real numbers, topological spaces, etc.), it is often possible to restrict our attention to the countable case (where such objects can then be described by an infinite sequence of zeros and ones, or equivalently by a set of integers). This leads into the realm of computability theory.

Not all sets of integers are equally complex. There are finite sets and, more generally, the so-called computable sets, which, although they may be infinite, can nevertheless be described by an algorithm, and hence by a finite string of symbols. This is the case for most sets encountered in mathematics, such as the set of prime numbers or the sequence of decimal digits of π . However, there also exist non-computable sets, which in fact constitute the overwhelming majority of sets of integers.

One possible temptation is therefore to restrict mathematical practice solely to computable objects, for philosophical or practical reasons. Such an approach would, however, amount to giving up a large part of modern mathematics, since many theorems necessarily imply the existence of non-computable objects. For example, there exists a computable sequence of reals in $[0, 1]$ such that none of its convergent subsequences is computable, and therefore the Bolzano-Weierstrass theorem requires the existence of such objects in order to hold.

Using the tools of computability theory, it is possible to organize these non-computable sets into a hierarchy (some sets may be regarded as more non-computable than others) and thereby classify mathematical theorems according to the computational complexity of the objects they involve.

The question of the computational strength of mathematical theorems is closely related to that of their axiomatic strength. Reverse mathematics (les mathématiques à rebours in French) is a branch of mathematical logic, initiated by Harvey Friedman, whose aim is to determine the minimal axioms required to prove a given theorem.

The idea behind reverse mathematics is to work within a deliberately restricted axiomatic system, in which only the existence of computable sets can be proved. In such a framework, theorems involving non-computable sets are no longer provable (without, however, becoming refutable, since one does not assume the nonexistence of non-computable sets).

One may then investigate which additional assumptions are required to establish these theorems. The method consists, after proving a theorem from stronger axioms, in reversing the argument: one seeks to show that the theorem itself suffices to recover those axioms, thereby demonstrating that they were necessary for the original proof.

A surprising structural phenomenon then emerges: the overwhelming majority of classical theorems fall into five major equivalence classes modulo computable mathematics. The axiomatic strength of each of these classes corresponds to a particular philosophical conception of mathematics and to a certain level of computational power.

Ramsey's theorem for pairs (in its infinite form) constitutes one of the first natural counterexamples to this structuring phenomenon. This raises the question of the position of other results from Ramsey theory, especially regarding generalizations, weakenings, or variations of Ramsey's theorem.

Rather than organizing themselves into a sixth major equivalence class, these various statements instead form a kind of "zoo" of new classes: some stronger, some weaker, and others incomparable with one another.

One of the aims of this thesis is the study of these statements (notably from the viewpoint of computability theory and reverse mathematics), in particular concerning methods for separating two theorems, that is, for showing that one does not follow from the other. While it is often "easy" to show that one result is a consequence of another by giving a proof, demonstrating the absence of such a proof is generally far more delicate.

One of the main tools used to obtain such results relies on the use of nonstandard models of arithmetic. These are structures that resemble the standard natural numbers $\{0, 1, 2, \dots\}$ in many respects, while differing from them at certain specific points. By analogy, they may be compared to models of non-Euclidean geometry, which made it possible to show that the parallel postulate does not follow from the other axioms of Euclidean geometry, or to the example of the plucked chicken used by Diogenes to expose the limitations of Plato's definition of Man as a "featherless biped".

1.2.1 Structure of the thesis and contributions

This thesis is divided into four parts.

Part I: This part is a survey of the state of the art. Following the guiding thread of Hilbert's program introduced in Chapter 2, it presents various fragments of first-order arithmetic (in Chapter 3), as well as theories of second-order arithmetic and reverse mathematics (in Chapter 5). Chapter 4 introduces computability-theoretic notions and their connections with arithmetical definability. Finally, Ramsey's theorem, together with its computability-theoretic and reverse-mathematical properties, is

introduced in Chapter 6.

Part II: This part focuses on the study of theorems from a computability-theoretic perspective. Chapter 7 introduces new tools from algorithmic randomness theory and studies the characterization of a particular class of permutations, namely separable permutations, as the hardest patterns to avoid in certain variants of Ramsey’s theorem. The results presented come from the article:

Q. Le Hou  rou and L. Patey, “Ramsey-like theorems for separable permutations.” Preprint (2025).

Chapter 8 studies the generalization to colorings of tuples of unbounded size of three variants of Ramsey’s theorem exhibiting a form of combinatorial weakness. It is shown that this weakness persists for only one of them, namely the rainbow Ramsey theorem. The results presented come from the article:

L. Carlucci, O. Gjetaj, Q. Le Hou  rou, and L. Levy Patey, “Ramsey-like theorems for the Schreier barrier.” *The Journal of Symbolic Logic*, pages 1–29, 2025.

Part III: This part gathers several results obtained through the construction of new models by the method of ω -extensions. Chapter 9 provides a survey of the use of computability-theoretic tools within nonstandard models of arithmetic. It presents the general scheme for constructing new models via ω -extensions, together with its use in establishing conservation results for first-order formulas.

Chapter 10 provides a new upper bound on the set of first-order consequences of Ramsey’s theorem for pairs. The results presented come from the article:

Q. Le Hou  rou, L. Levy Patey, and K. Yokoyama, “Conservation of Ramsey’s theorem for pairs and well-foundedness.” *Transactions of the American Mathematical Society*, 378(3):2157–2186, 2025.

Chapter 11 studies an unbalanced version of Ramsey’s theorem in which monochromatic cliques of a given color are required to have bounded size. The results presented come from the article:

Q. Le Hou  rou and L. Patey. “The reverse mathematics of bounded Ramsey’s theorem for pairs.” *The Journal of Symbolic Logic*, pages 1–26, 2026.

Part IV: This final part explores the connections between finitary and infinitary versions of theorems from Ramsey theory, in particular with a view toward obtaining partial conservation results. Chapter 12 introduces the notion of α -large sets due to Ketonen and Solovay [72], and presents new techniques for constructing nonstandard models by cuts based on such sets. These techniques arise from the article:

Q. Le Hou  rou, L. Levy Patey, and K. Yokoyama, “ Π^0_4 conservation of Ramsey’s theorem for pairs.” *Journal of the London Mathematical Society*, 113:e70419 (2026).

Chapter 13 presents results of primarily combinatorial interest by giving optimal bounds for finitary versions of Ramsey’s theorem and the pigeonhole principle over α -large sets. The results presented mainly come from the article:

Q. Le Hou  rou and L. Patey, “Partition theorems for Ketonen–Solovay largeness.” Preprint (2026).

Chapter 14 exploits this formalism for constructing cuts, together with the bounds obtained in the previous chapter, in order to establish the characterization of the first-order consequences of Ramsey’s theorem for pairs for a broad class of formulas, obtained in the article “ Π_4^0 conservation of Ramsey’s theorem for pairs.” The construction presented there simplifies that of the original article.

Chapter 15 applies the same formalism to theorems concerning colorings of trees in order to obtain partial conservation results. The results presented come from the article:

Q. Le Hou  rou and L. Levy Patey, “ Π_4^0 conservation of the Ordered Variable Word theorem.” *The Journal of Symbolic Logic* 2025:1–16.

For the sake of brevity, and in the interest of providing a document forming a coherent whole, some results obtained during these three years could only be mentioned very briefly, or do not appear in this thesis at all. This is notably the case for results concerning comparisons of proof lengths between various theories, obtained in the article:

Q. Le Hou  rou and L. Patey, “Largeness notions and polytime translation for $\forall\Sigma_3^0$ -consequences of RT_2^2 .” Preprint (2026).

The formalism of forcing interpretations introduced by Avigad [4] and used by Ko  odziejczyk et al. [78], together with the combinatorics involved in this article, was too heavy to introduce within the framework of this thesis, and the result is therefore only mentioned.

The same applies to the comparison of the different versions of the pigeonhole principle along the arithmetical hierarchy obtained in:

Q. Le Hou  rou, A. Mimouni, and L. Levy Patey, “The reverse mathematics of the pigeonhole hierarchy.” Preprint (2024).

These results did not fit naturally within the different parts of this thesis and also relied on a rather heavy formalism.

Part I

State of the art

Chapter 2

Hilbert's program

2.1 Historical background

The motivation behind this thesis can be traced back to a program initiated by David Hilbert at the beginning of the 20th century, aimed at establishing a solid foundation for all mathematics. At that time, the discipline was facing a profound crisis. For centuries, the concept of infinity used by mathematicians was only infinite in potential: a line could always be extended further, and there was always a larger integer. The idea of treating these potential infinities as completed or actual objects that can be manipulated was dismissed, on the grounds that it was both unnecessary for mathematical practice and logically problematic. Such objections could be found in writings by Aristotle,

Our account does not rob the mathematicians of their science, by disproving the actual existence of the infinite in the direction of increase, in the sense of the untraversable. In point of fact, they do not need the infinite and do not use it. They postulate only that the finite straight line may be produced as far as they wish.

(Aristotle. (circa 360 B.C), *Physics, Book III, Part 7* (R. P. Hardie & R. K. Gaye, Trans.))

and Thomas Aquinas,

The same applies to a mathematical body. For if we imagine a mathematical body actually existing, we must imagine it under some form, because nothing is actual except by its form; hence, since the form of quantity as such is figure, such a body must have some figure, and so would be finite; for figure is confined by a term or boundary.

(Aquinas, T. (circa 1270), *Summa Theologica I, q. 7, A3.C5* (C. Decaen, Trans.))

Nevertheless, this did not stop Georg Cantor from making use of the actual infinity in his development of set theory. His work sparked further philosophical objections, with the examples of L. E. J. Brouwer's intuitionism, or the finitism of Leopold Kronecker or Ludwig Wittgenstein.

Let's imagine a man whose life goes back for an infinite time and who says to us: 'I'm just writing down the last digit to Pi and it's 2'. Every day of his life he has written down a digit, without ever having begun; he has just finished. This seems utter nonsense, and a reductio ad absurdum of the concept of infinite totality.

(Wittgenstein, L. (1975), *Philosophical Remarks*, p. 166)

His theory was also under attack mathematically, for the many paradoxes appearing from the careless manipulation of sets, such as the Russell and Burali-Forti paradoxes.

Seeing the potential of set theory as a unifying theory for all mathematics, and not wanting to give up on the powerful proof techniques resulting from it, Hilbert proposed a program that, if successful, would address those philosophical objections and provide a strong guarantee that no new contradictions would arise.

It is not easy to provide a precise formulation of what the goals of this program were, as it was formulated over a span of 30 years through a series of articles, conferences, and books, each emphasizing different aspects. Nevertheless, a general formulation of it can be stated as follows:

The first step of the program was to give a precise formalization of all the mathematics developed up to that point:

Everything that previously made up mathematics is to be rigorously formalized, so that mathematics proper or mathematics in the strict sense becomes a stock of formulae.
(Hilbert, D. (1923), *Die logischen Grundlagen der Mathematik*. (Ewald, Trans.))

This would make it possible to see mathematical propositions and proofs as finite sequences of symbols obeying some syntactic rules. With that, verifying the correctness of a proof became a mechanical process, independent of the intended meaning of the symbols involved. This part of the program has been rather successful, with the examples of Zermelo-Fraenkel set theory (ZFC), or second-order arithmetic (Z_2) which can both be used to formalize a large part of mathematics, in particular these infinitary arguments. For practical reasons, which will be made clear later, we will work in the framework of second-order arithmetic, which we define in Chapter 5.

The next step is to isolate a fragment of mathematics corresponding to the uncontroversial “finitary reasoning” that operates only with finite objects in finite ways. Several possible candidates for such a theory have been proposed. Here we will follow Tait [137] which identified the system PRA of Primitive Recursive Arithmetic as a good formalization of finitary mathematics. The system PRA will be described in detail in Section 2.2.

Finally, now that all mathematical reasoning has been reduced to a finite manipulation of symbols, it is possible to discuss these strong theories dealing with infinite totalities from a finitist viewpoint. The final step is then to prove, using only finitary reasoning, that those strong theories are *consistent*, that is, that they do not prove any contradiction and can therefore be safely used in the development of new branches of mathematics. Thus, a proof in PRA of the consistency of Z_2 (which we write $\text{PRA} \vdash \text{Con}(Z_2)$) would be seen as a realization of Hilbert’s program.

Several other goals in Hilbert’s program can be identified:

Independence: Before adding a new statement as an axiom to a theory, we must ensure not only that the resulting theory remains consistent, but also that the statement is *independent* of the theory, meaning it cannot be derived from the other axioms, and thus its addition does not introduce any redundancy.

Completeness: A theory is said to be *complete* if it is consistent and if, for every statement expressible in its language, the theory can either prove the statement or prove its negation. Or, to put it another way, that no new axiom can be added without either introducing a contradiction, or being redundant.

Decidability: Another desirable property of a theory would be its *decidability*, that is, the existence of an algorithmic procedure that can determine whether any given statement is provable within the theory. If a theory is complete, then we can consider the algorithm that, given a statement, systematically lists all possible proofs in that theory until it finds either a proof of the statement or of its negation. This algorithm will always provide an answer, thereby proving the decidability of the theory. This will not be the last link between the theory of proofs and the theory of computability, as we will see in Chapter 4.

Conservativity: An even stronger justification for the use of the actual infinity in mathematical practice would be to establish the *conservativity* of the stronger infinitary theory over the weak finitary one, that is, to show that every proof of a finitary statement made using infinitary methods could also be obtained without them. In this way, the use of actual infinity becomes merely a convenient shortcut, one that can always be avoided if desired. Hilbert did not explicitly set conservativity as a goal, but it emerged as a byproduct of several consistency proofs developed by Hilbert, Bernays, Ackermann, and others. Conservation proofs will also play a key role in this thesis.

Once it has been established that one theory is conservative over another, a natural question arises: does the stronger theory allow for shorter proofs than the weaker one? These kinds of questions were first investigated in 1936 by Kurt Gödel in a short paper [52], and has since gained increasing importance, particularly due to its connections with complexity theory.

One way to establish the consistency of a given theory is to assign meanings to the symbols in its language and construct a model for it; that is, to exhibit a structure that satisfies all the axioms of the theory. We will refer to arguments of this kind as semantic proofs. Such proofs indeed demonstrate the theory's consistency, since if the theory were to prove a contradiction, then every model of the theory would also satisfy that contradiction, which is impossible once we have constructed even a single model. More generally, to show that a theory does not prove a particular statement, it is sufficient to build a model of that theory in which the statement is false. One famous example of this method is the models of non-Euclidean geometries, such as the Poincaré disk model. These models satisfy the first four axioms of Euclidean geometry but not the parallel postulate, thereby showing that it cannot be derived from the other axioms.

Unfortunately, except in a few very simple cases such as the consistency proof for propositional logic, where the intended model is the finite set $\{\text{true}, \text{false}\}$, proofs of this kind are generally not finitist, since the theories in question usually admit only infinite models, and constructing such models requires some form of set-existence axioms:

Das Problem der Erfüllbarkeit eines Axiomensystems (bzw. einer logischen Formel), welches im Falle endlicher Individuenbereiche positiv durch Aufweisung gelöst werden kann, ist im Falle, wo zur Erfüllung der Axiome nur ein unendlicher Individuenbereich in Betracht kommt, nicht nach dieser Methode lösbar, weil die Existenz unendlicher Individuenbereiche nicht als ausgemacht gelten kann.¹

(Hilbert, D., Bernays, P. (1934), Grundlagen der Mathematik.)

Another way to establish the consistency of a theory is through syntactic manipulations. There are several possible approaches to such proofs, but the general strategy is to assume the existence of a proof of a contradiction within the theory and then transform that proof into a simpler one, using techniques such as cut elimination. This transformation may eliminate certain axioms, thereby reducing the consistency of the theory to that of a weaker one, or it may simplify proofs to a form where an induction on their structure shows that no proof of a contradiction can exist.

Unfortunately, Hilbert's program came crashing down due to Gödel's incompleteness theorems. By the first theorem, no complete theory of even

¹There is currently no official translation of this passage, so here is the author's own version: "The problem of the satisfiability of an axiom system (or a logical formula), which can be solved positively by demonstration in the case of finite individual domains, cannot be solved using this method in cases where only an infinite individual domain can be considered for the satisfaction of the axioms, because the existence of infinite individual domains cannot be taken for granted."

basic arithmetic can be given², and, even more damning, the second theorem showed that all the theories in question cannot prove their own consistency, let alone that of any stronger theory.

Nevertheless, although the program in its full generality has failed, certain partial results remain possible. For example, rather than directly proving the consistency of a theory using finitary means, one may instead show that its consistency can be reduced to that of a weaker theory. A notable example of such a result is due to Cohen [28], who, through his method of forcing, showed that the consistency of ZF is equivalent to the consistency of ZF plus the negation of the continuum hypothesis. Moreover, his proof could be reformulated to be finitistically acceptable, that is, $\text{PRA} \vdash \text{Con}(\text{ZF}) \rightarrow \text{Con}(\text{ZF} + \neg\text{CH})$ (see, for instance, Weaver [141] for a discussion on how set-theoretic forcing can be formalized within weak theories such as Peano arithmetic).

We will say that a theory $\mathcal{T}$ is *finitistically reducible* if it can be shown, using finitary methods, that the consistency of $\mathcal{T}$ can be reduced to that of finitary mathematics (formalized here by PRA), that is, if $\text{PRA} \vdash \text{Con}(\text{PRA}) \rightarrow \text{Con}(\mathcal{T})$ (rather than a direct proof $\text{PRA} \vdash \text{Con}(\mathcal{T})$). Such a result may be regarded as a partial realization of Hilbert's program, since the consistency of finitary mathematics itself was never in question. And, as noted by Simpson [127], a large part of modern mathematical practice is finitistically reducible.

The first part of the thesis presents the foundational material necessary to understand the contributions outlined in the other three. This exposition proceeds gradually, beginning with the introduction of the weak finitistic theory PRA and progressively introducing stronger theories, while also introducing the necessary background material. Along the way, we examine the extent to which each of these theories is finitistically reducible.

In this first part, all the statements of relative consistency between theories have been proven by syntactic proofs: they proceed by showing that a contradiction derived in a stronger theory can be transformed into a contradiction in a weaker one (such as PRA in the case of proofs of finitistic reducibility). Moreover, such proofs often yield a stronger result: not only can proofs of contradictions be transformed, but also proofs of a broader class of statements, thereby proving the conservation of the strong theory over the weak one for these statements.

The culminating result of this first part is the fact that a rather strong theory of second-order arithmetic, called WKL_0 , is finitistically reducible. This theory corresponds exactly to the strength needed to do most semantic proofs of relative consistency: if a weak theory is consistent, then it has a model (by Gödel's completeness theorem, provable in WKL_0), and we can then transform this model into one of a stronger theory, thereby proving its consistency. Similarly, such semantic proofs can also yield partial conservation results, when some properties of the old model are preserved in the new one.

For that reason, in the remaining parts of this thesis, all the proofs of relative consistency, partial conservation, and non-implications will be done using semantic proof techniques, which explains why we will not go into the details of the syntactic proofs of the first part.

2.2 Primitive recursive arithmetic

Hilbert drew the line between finitary and infinitary mathematics at the use of quantifiers on an unrestricted domain:

²More precisely, no such computable theory exists. This computability requirement is typically implicit when seeking a workable formal theory; examples of non-computable theories will appear in Chapter 4.

*Now, where do we first depart from the concrete intuitive and from the finite?
Obviously, already with the application of the concepts 'all' and 'there exists'.
(Hilbert, D. (1923), Die logischen Grundlagen der Mathematik. (Ewald, Trans.))*

This restriction may seem too harsh, but, surprisingly, Skolem's theory of Primitive Recursive Arithmetic (PRA) [131] demonstrated that a significant portion of what is known about the natural numbers can be formalized without the use of the existential and universal quantifiers.

To still be able to make non-trivial statements about the integers (i.e., statements whose truth cannot just be checked by a computation, unlike simple identities such as $2 + 3 = 3 + 2$), PRA treats formulas like $x + y = y + x$ as implicitly universally quantified. Philosophically, the formula $x + y = y + x$ can be seen as a statement about two arbitrary numbers x and y , whereas the universally quantified formula $\forall x \forall y (x + y = y + x)$ is seen as an infinite conjunction over all possible values of x and y . This distinction makes the former acceptable within a finitistic framework, but not necessarily the latter. However, in first-order logic, due to the rules of universal generalization and instantiation, both formulas effectively express the same proposition (in an empty context where nothing is assumed about x and y , which makes them truly arbitrary).

2.2.1 Primitive recursive functions

The theory PRA is based on a subclass of computable functions known as the primitive recursive functions. These are the functions whose inductive definitions make them obviously computable and total (roughly speaking, another way to view them is as functions computed by programs that use only “for” loops for iteration). For a detailed explanation of why these functions are considered the “finitistic” ones, see Tait [137].

Definition 2.2.1 (Primitive recursive function). A *primitive recursive function* is a function from $\mathbb{N}^k$ to $\mathbb{N}$ for some $k \in \mathbb{N}$ that belongs to the class defined inductively as follows:

- The constant 0 is primitive recursive.
- The successor function, $\text{succ} : \mathbb{N} \rightarrow \mathbb{N}$, is primitive recursive.
- The projections $p_i^n : \mathbb{N}^n \rightarrow \mathbb{N}, (x_0, \dots, x_{n-1}) \mapsto x_i$ are primitive recursive for every $i, n \in \mathbb{N}$ with $i < n$.
- If f is an n -ary primitive recursive function and $g_0, \dots, g_{n-1}$ are m -ary primitive recursive functions, then $\vec{x} \mapsto f(g_0(\vec{x}), \dots, g_{n-1}(\vec{x}))$ is an m -ary primitive recursive function.
- If f and g are respectively n -ary and $n + 2$ -ary primitive recursive functions, then $h : \mathbb{N}^{n+1} \rightarrow \mathbb{N}$, defined by:

$$\begin{aligned} - & h(0, x_0, \dots, x_{n-1}) = f(x_0, \dots, x_{n-1}) \\ - & h(y + 1, x_0, \dots, x_{n-1}) = g(y, h(y, x_0, \dots, x_{n-1}), x_0, \dots, x_{n-1}) \end{aligned}$$

is primitive recursive.

Definition 2.2.2. A relation on $\mathbb{N}^n$ is said to be primitive recursive, if the function $\mathbb{N}^n \rightarrow \{0, 1\}$ corresponding to its graph is primitive recursive. Similarly, a subset of $\mathbb{N}$ is primitive recursive if its characteristic function is primitive recursive.

Proposition 2.2.3. The following functions are primitive recursive:

- addition, multiplication, exponentiation, tetration, ...
- The function that, given an index e of a program in some formalism (such as

Turing machine) and an integer s , outputs 1 if the program halts within s steps of computation and 0 otherwise.

One way to obtain a computable total function that is not primitive recursive is by diagonalization: we can computably obtain a list of all primitive recursive functions $p_e : \mathbb{N} \rightarrow \mathbb{N}$ and consider a function f by letting $f(e) = \max_{e' \leq e} p_{e'}(e) + 1$. By definition, f eventually outgrows every primitive recursive function and thus cannot be itself primitive recursive. This diagonalization method is essentially the way the Ackermann function is defined, so that it outgrows additions, multiplications, exponentiations, tetrations, ...

Why are the functions obtained by this diagonalization method not considered “finitistic”? After all, if the primitive recursive functions are obviously total and computable, then diagonalizing over them should also yield a function that is again obviously total and computable. Tait addresses this issue by observing that “the notion of a finitist function is not finitist”, since the notion of a function is itself not finitist (a finitist cannot talk about functions, he can only manipulate some expressions involving some unbound parameters that we view as functions from a nonfinitist viewpoint). Rather, what Tait is trying to accomplish is to “characterize the finitist functions and finitist proofs from a nonfinitist standpoint”: while a finitist may accept any given inductive definition of a primitive recursive function as meaningful, the collection of all such functions (or an effective enumeration of them) is not something the finitist can meaningfully grasp, hence a function whose definition uses this collection is not considered “finitistic”.

2.2.2 PRA

Similarly, the theory PRA described below is a formalization of finitary mathematics, but it should be noted that this formalization is not itself carried out from a finitistic perspective: a finitist would accept every proof made within PRA, yet they would not view the theory PRA itself as something finitistically meaningful. This justifies why the theory PRA is allowed to have a language with infinitely many symbols (one for every primitive recursive function) and to have infinitely many axioms: all that matters is that every proof made in PRA uses finitely many of them.

Definition 2.2.4 ((quantifier-free) PRA (see Avigad [5, Chapter 9])). The language of PRA includes one function symbol for every description of a primitive recursive function (there may be several ways to define the same primitive recursive function starting from the basic functions and using composition and primitive recursion; we pick a symbol for every such way), the propositional connectives ($\wedge, \vee, \neg, \rightarrow$), infinitely many variables, and the equality symbol $=$. We will use symbols like $+$ and $\times$ to denote the function symbols corresponding to their standard descriptions as primitive recursive functions.

The *terms* are obtained inductively using the variables and function symbols (e.g., x , $\text{succ}(0)$, $p_2^3(x, y, 0)$, $x + y$, ...). For every $n \in \mathbb{N}$, we write n as a shorthand for the term $\text{succ}(\text{succ}(\dots \text{succ}(0) \dots))$ with n iterations of succ .

The axioms of PRA are given by:

1. $\neg(\text{succ}(x) = 0)$
2. $p_i^n(x_0, \dots, x_{n-1}) = x_i$ for every n and every $i < n$
3. $h(x_0, \dots, x_{m-1}) = f(g_0(x_0, \dots, x_{m-1}), \dots, g_{n-1}(x_0, \dots, x_{m-1}))$ if h was defined as the composition of f and the g_i .
4. $h(0, x_0, \dots, x_{n-1}) = f(x_0, \dots, x_{n-1})$ and $h(\text{succ}(y), x_0, \dots, x_{n-1}) = g(y, h(y, x_0, \dots, x_{n-1}), x_0, \dots, x_{n-1})$ if h was defined from f and g by primitive recursion.
5. The tautologies of intuitionistic propositional calculus.

We assume that the axioms are closed under substitution, i.e., that the free variables in their formulation can be replaced by any terms.

Finally, we also add the quantifier-free induction rule stating that, from $A(0)$ and $A(x) \rightarrow A(\text{succ}(x))$ we can deduce $A(t)$ for any term t . Using a small trick, this deduction rule can in fact be replaced by axioms expressing the fact that we have induction in a quantifier-free way (see, e.g., Avigad [5, Section 9.4]).

The theory PRA is sufficiently strong to formalize all the basic syntactic proof-theoretic arguments. Indeed, all the standard encodings of propositions and proofs (e.g., Gödel numbering $\ulcorner \cdot \urcorner$), and the various operations on them (renaming of variables, composition of proofs...) are given by primitive recursive functions. Thus, PRA is able to formulate statements about itself. For example, there exists a primitive recursive function $Pr_{\text{PRA}}(x, y)$ that is equal to 1 when x is the encoding of a proof of the formula encoded by y in PRA and 0 otherwise. The consistency of PRA can then be expressed by the formula (with one free parameter x) $\text{Con}(\text{PRA}) := Pr_{\text{PRA}}(x, \ulcorner 0 = 1 \urcorner) = 0$. By Gödel's second incompleteness theorem, it follows that PRA does not prove its own consistency: $\text{PRA} \not\vdash \text{Con}(\text{PRA})$.

Whether PRA is based on classical or intuitionistic propositional calculus is inconsequential, since the resulting theories coincide, as demonstrated by the following result:

Proposition 2.2.5 (see, e.g., Avigad [5, Proposition 9.1.4]). *For every formula A in the language of PRA, $\text{PRA} \vdash A \vee \neg A$.*

Therefore, from this point onward, we will freely use the excluded middle in our proofs. Including it as axioms allows for shorter proofs. Indeed, $A \vee \neg A$ can then be derived in a single line for every formula A , whereas without it, several steps are required. That said, the gap between those two systems is not that significant: by analyzing the proof of the preceding proposition, we obtain proofs in PRA of $A \vee \neg A$ whose size (i.e., total number of symbols used) is bounded by a polynomial in the size of A .

Now that we have identified a theory corresponding to finitary mathematics, we can begin to compare its strength with more powerful infinitary theories. As a first step, we can extend the language of PRA by introducing quantifiers, thereby crossing into the realm of infinitary mathematics. This will be done in the next chapter, which introduces first-order arithmetic and examines the relative strength of several of its theories, both among themselves and in comparison to PRA.

Chapter 3

First-order arithmetic and fragments

Bounded quantifiers are quantifiers of the form $\forall x < b \varphi(x)$ or $\exists x < b \varphi(x)$ where b is a term not containing the variable x . These are often defined using unbounded quantifiers by letting $\forall x < b \varphi(x)$ be shorthand for $\forall x(x < b \rightarrow \varphi(x))$ and $\exists x < b \varphi(x)$ be shorthand for $\exists x(x < b \wedge \varphi(x))$. Despite this, their use was not seen by Hilbert as problematic from a finitistic viewpoint, as they can be seen as finite conjunctions or disjunctions:

The assertion that all the objects of a finite existing surveyable totality possess a particular property is logically equivalent to a conjunction of several individual assertions,
Hilbert, D. (1923), Die logischen Grundlagen der Mathematik. (Ewald, Trans.)

And indeed, PRA can simulate the use of bounded quantifiers: if $\varphi(x, \vec{y})$ defines a primitive recursive relation and if b is a term not containing the variable x , then $\exists x < b \varphi(x, \vec{y})$ and $\forall x < b \varphi(x, \vec{y})$ are primitive recursive relations, and there exist formulas in the language of PRA that serve the same purpose. Thus, from now on, we allow the use of bounded quantifiers in PRA proofs. A formula where all the quantifiers are bounded is said to be a *bounded formula*.

For unbounded quantifiers, we have to extend the language of PRA. The resulting theory, which we will call first-order PRA, has the same axioms as the quantifier-free version of PRA (except that we can now take their universal closure: $\forall x \neg(\text{succ}(x) = 0)$ instead of $\neg(\text{succ}(x) = 0)$...), which dispenses with the need to close them under substitution) but in an extended language with quantifiers (and the corresponding rules of first-order logic to handle them).

First-order PRA is still a weak theory, as even in this extended language, we only have induction for quantifier-free formulas (or bounded formulas, by our above remark). The theory of Peano arithmetic (PA) can be seen as the extension of first-order PRA obtained by adding the induction scheme for all first-order formulas (also called arithmetical formulas). Below PA, we can consider several subtheories, called fragments, forming hierarchies measuring the amount of induction at our disposal.

3.1 Arithmetical hierarchy

The standard way to measure the complexity of a formula in first-order arithmetic is via the arithmetical hierarchy. At its base are the simplest formulas, which contain only bounded quantifiers. Then, the more we alternate between existential and universal quantifiers to express the formula, the more complicated it is. In Chapter 4, we will see how the arithmetical hierarchy is closely connected to concepts from computability theory.

Definition 3.1.1 (Σ_n - and Π_n -formulas). We define inductively the classes Σ_n and Π_n of formulas for n an integer:

- The classes Σ_0 and Π_0 are equal and correspond to first-order formulas where all the quantifiers are bounded (e.g., $\forall x < 7 \exists y < x \ 2 \times y = x \vee 2 \times y + 1 = x$)
- If φ is Σ_n , then $\forall x_0 \forall x_1 \dots \forall x_{k-1} \varphi(\vec{x})$ is a Π_{n+1} -formula for every $k \in \mathbb{N}$. If ψ is Π_n , then $\exists x_0 \exists x_1 \dots \exists x_{k-1} \psi(\vec{x})$ is a Σ_{n+1} -formula for every $k \in \mathbb{N}$.

Example 3.1.2. • The property of being a prime number is expressible by a Σ_0 -formula: $\text{prime}(x) := x > 1 \wedge (\forall y < x \forall z < x \ (y+2)(z+2) \neq x)$

- The primitive recursive relations are exactly those definable by a Σ_0 -formula.
- The Goldbach conjecture is a Π_1 -statement: $\forall x \exists y \leq 2x+4 \ \text{prime}(y) \wedge \text{prime}(2x+4-y)$.
- The twin prime conjecture is a Π_2 -statement: $\forall x \exists y \ y > x \wedge \text{prime}(y) \wedge \text{prime}(y+2)$.
- For every formula φ , the statement that there exists a proof of φ in PRA, PA, ZFC..., is Σ_1 .

With this definition, not every first-order formula falls into the Σ_n or Π_n classes for some n , as only the formulas in prenex normal form (with the quantifiers moved to the front) are classified. Nevertheless, classical first-order logic is sufficient to show that any formula is equivalent to one in prenex normal form. For convenience, we will sometimes abuse notation and refer to a formula as being Σ_n or Π_n when it is clear how to rewrite it in prenex normal form (e.g., given a Σ_n -formula φ , we may refer to $\neg\varphi$ as being Π_n , as it is clear how to push the negation inside the quantifiers). Note that we do not close the classes under logical equivalence, as is sometimes done. In doing so, we would sacrifice the desirable feature of these classes having a purely syntactic (and decidable) definition.

Using the existence of a bijection between $\mathbb{N}^k$ and $\mathbb{N}$ for every $k \in \mathbb{N}$, which we denote by $\langle a_0, \dots, a_{k-1} \rangle$, we can group multiple quantifiers into one, writing $\forall \langle x_0, \dots, x_{k-1} \rangle \varphi(x_0, \dots, x_{k-1})$ in place of $\forall x_0 \dots \forall x_{k-1} \varphi(x_0, \dots, x_{k-1})$ (and similarly for existential quantifiers). Therefore, without loss of generality, we will often assume that Σ_n - and Π_n -formulas consist of exactly n strictly alternating single quantifiers.

3.2 First-order PRA

The following proposition shows that the first-order version of PRA is conservative over its quantifier-free counterpart. That is, adding quantifiers does not allow us to prove any new statements in the original, quantifier-free language: every quantifier-free statement provable in first-order PRA is already provable in quantifier-free PRA. In particular, if quantifier-free PRA is consistent (i.e., does not prove $0 = 1$), then its first-order extension is also consistent.

In fact, we have a stronger result: for every Π_2 -statement provable in first-order PRA (e.g., of the form $\forall x \exists y \varphi(x, y)$ with $\varphi(x, y)$ a Σ_0 -formula), there is a statement provable in quantifier-free PRA that implies it (of the form $\varphi(x, f(x))$ for f a primitive recursive function symbol).

Proposition 3.2.1 (see, e.g., [5, Section 9.4]). *The first-order version of PRA is Π_2 -conservative over the quantifier-free version of PRA in the following sense: If first-order PRA proves a formula of the form $\forall \vec{x} \exists y_0, \dots, y_{k-1} \varphi(\vec{x}, y_0, \dots, y_{k-1})$, with φ a bounded formula, then there exists primitive recursive function symbols $f_0, \dots, f_{k-1}$ such that quantifier-free PRA proves $\varphi(\vec{x}, f_0(\vec{x}), \dots, f_{k-1}(\vec{x}))$.*

The original techniques introduced by Hilbert to obtain such “quantifier elimination” results were based on a formalism known as the epsilon calculus. Today, Gentzen’s method of cut elimination is more commonly used. The preceding proposition, for instance, can be proved using Herbrand’s theorem, which in turn can be derived either via cut elimination or from the epsilon theorems of Hilbert and Bernays. These proof techniques can all be formalized in quantifier-free PRA, which gives us our first partial realization of Hilbert’s program: the addition of first-order quantifiers has been finitistically justified, and from now on, we will freely use first-order PRA in our finitary proofs of Π_2 -statements.

Nevertheless, transforming proofs via cut elimination and Herbrand’s theorem can lead to a substantial increase in proof size (up to a tower of exponentials in the size of the original proof, as shown by Statman [136]). And it was conjectured by Ignjatovic [66] that the first-order version of PRA yields a non-elementary speedup over the quantifier-free version, where a function is said to have *elementary* growth if it can be bounded by a tower of exponentials of fixed height.

3.3 Induction

Having infinitely many function symbols in a theory of arithmetic is not something usual and is often cumbersome to work with. Instead, when defining a theory such as PA, one typically works in a more restricted language, with only $0, succ, +$ and $\times$ as function symbols (and sometimes $<$ as a relation symbol). Other functions are then handled by defining relations that represent their graphs in terms of these basic symbols.

Adopting this restricted language leads to a revised presentation of the arithmetical hierarchy, in which the base class $\Sigma_0 = \Pi_0$ now consists of bounded formulas using only these function symbols. The relations definable by such formulas are known as the *rudimentary* relations and are strictly included in the primitive recursive relations (see e.g., Salehi [121] for an example of a primitive recursive relation that is not rudimentary). This modification, however, does not affect the higher levels of the hierarchy. Indeed, every primitive recursive relation (and even every computable relation) can be defined by both a Σ_1 - and a Π_1 -formula in that restricted language (as we will see in Chapter 4). Consequently, for all $n \geq 1$, the classes of relations definable by Σ_n - or Π_n -formulas remain unchanged. For the remainder of this thesis, we shall work in this language and with the new definition for the base class $\Sigma_0 = \Pi_0$.

Once the graph of a computable function has been defined by both a Σ_1 - and a Π_1 -formula, its basic computational properties can be established using only those of $succ, +$ and $\times$, as provided by the theory of Robinson arithmetic ($\mathcal{Q}$) introduced below. For instance, the graph of the function $(x, y) \mapsto x^y$ can be defined by a Σ_0 -formula $pow(x, y, z)$ (see [57, Theorem V.3.15]), and hence also both by a Σ_1 - and a Π_1 -formula. Then, for every $n, m, k \in \mathbb{N}$, one can derive either $\mathcal{Q} \vdash pow(n, m, k)$ or $\mathcal{Q} \vdash \neg pow(n, m, k)$, depending on whether n^m is equal to k (where n, m , and k are shorthand for iterated applications of $succ$ to 0).

Definition 3.3.1 (Robinson arithmetic). The theory of Robinson arithmetic ($\mathcal{Q}$) is given by the universal closure of

- $succ(x) \neq 0$
- $x + succ(y) = succ(x + y)$
- $succ(x) = succ(y) \rightarrow x = y$
- $x \times 0 = 0$
- $x = 0 \vee (\exists y) succ(y) = x$
- $x \times succ(y) = (x \times y) + x$
- $x + 0 = x$

The ability of $\mathcal{Q}$ (and of stronger theories) to prove the outcome of any computation is captured by the notion of Σ_1 -completeness:

Definition 3.3.2 (Σ_1 -complete). A theory $\mathcal{T}$ is said to be Σ_1 -complete if for any Σ_1 -statement φ , $\mathcal{T} \vdash \varphi$ if and only if φ is true in $\mathbb{N}$.

The theory $\mathbf{Q}$ contains no induction axiom and is too weak to establish even basic Π_1 properties of $+$ and $\times$ (such as their commutativity and associativity). Hence, although $\mathbf{Q}$ suffices to prove the outcome of any given computation, we will always extend it with additional axioms to make it a workable theory.

When considering theories without induction, a more natural standalone theory is the theory $\mathbf{PA}^-$, which includes the axioms of $\mathbf{Q}$, as well as the basic algebraic properties of $+$, $\times$ and $<$ (commutativity and associativity of $+$ and $\times$, distributivity of $\times$ over $+$, compatibility of $+$ and $\times$ with $<$, and the fact that no element lies strictly between a number and its successor). Another description of $\mathbf{PA}^-$, emphasizing its purely algebraic content, is its characterization as the axiomatization of the nonnegative part of discretely ordered rings, where an ordered ring is a commutative ring together with a linear order compatible with its two operations ($a < b$ and $c < d$ implies $a + c < b + d$ and $0 \leq a, b$ implies $0 \leq a \times b$), discrete means that there is no element between 0 and 1 and the nonnegative part are the elements of the ring greater than 0.

We can now consider the following hierarchy of induction schemes. The inclusion of $\mathbf{PA}^-$ (or other weak base theories) in the definition of $\mathbf{I}\Gamma$ and of the other schemes considered in this chapter is purely a matter of notational convenience, as it makes it possible to work with them as self-contained theories, and not have to write $\mathbf{PA}^- + \mathbf{I}\Gamma$ every time.

Definition 3.3.3 ($\mathbf{I}\Gamma$). For $\varphi(x)$ a formula (with possibly other free variables), the induction axiom associated to it is the universal closure of the formula:

$$(\varphi(0) \wedge \forall x(\varphi(x) \rightarrow \varphi(x+1))) \rightarrow \forall x\varphi(x)$$

For Γ a class of formulas, we let $\mathbf{I}\Gamma$ be the axiom scheme comprised of the induction axiom for every formula in Γ together with the axioms of $\mathbf{PA}^-$.

Proposition 3.3.4 (Paris and Kirby [107]). *For every $n \in \mathbb{N}$, $\mathbf{I}\Sigma_n$ is logically equivalent to $\mathbf{III}_n$ (and we will mostly talk about $\mathbf{I}\Sigma_n$ from now on).*

The theory $\mathbf{I}\Sigma_0$ is strong enough to prove many properties of $\text{pow}(x, y, z)$, such as the fact that $\forall x, y_0, y_1, z_0, z_1 \text{ pow}(x, y_0, z_0) \wedge \text{pow}(x, y_1, z_1) \rightarrow \text{pow}(x, y_0 + y_1, z_0 \times z_1)$. However, defining a function via its graph is different from including a corresponding function symbol. When a function symbol is part of the language, the totality of the corresponding function is guaranteed by definition. In contrast, if a function is defined through its graph, its totality must be proved, and this is not always achievable within weak theories. For example, $\mathbf{I}\Sigma_0$ does not prove the Π_2 -statement $\text{exp} := \forall x \forall y \exists z \text{ pow}(x, y, z)$ expressing the totality of the function $(x, y) \mapsto x^y$.

Fortunately, in stronger theories like $\mathbf{PA}$ (which we can now define as $\bigcup_{n \in \mathbb{N}} \mathbf{I}\Sigma_n$), the induction at our disposal is sufficient to prove the totality of many functions, which makes it possible to work with them as if they are part of the language from the start, and justify the use of a restricted language. In fact, for every primitive recursive function, $\mathbf{I}\Sigma_1$ is sufficient to prove the totality of the standard algorithm computing it. As a result, $\mathbf{I}\Sigma_1$ can be viewed as including first-order PRA.

In fact, by the following statement, $\mathbf{I}\Sigma_1$ is finitistically reducible (and, similarly to Proposition 3.2.1, this statement can be proven in PRA itself). Thus, from now on, we will freely use Σ_1 -induction when proving Π_2 -statements in PRA, and will assume, when convenient, that we work in a theory using only 0, succ , $+$, $\times$ and $<$ as function and relation symbols.

Proposition 3.3.5 (see, e.g., Avigad [5, Theorem 10.6.1]). *$\mathbf{I}\Sigma_1$ is Π_2 -conservative over PRA.*

IS_1 admits a non-elementary speedup over PRA, even for Π_1 -formulas. This was shown by Ignjatović in his dissertation [66, Theorem 2.9], with a later published account given by Caldon and Ignjatović [12]. This further justifies the use of IS_1 .

As we will see in Proposition 3.6.5, IS_2 is already a theory that is not finitistically reducible (as IS_2 proves the consistency of PRA, which is a Π_1 -statement that PRA does not prove).

In Chapter 4, we will introduce the notion of a Δ_n set as one that is definable by both a Σ_n - and a Π_n -formula. This is purely a semantic notion, as such equivalences are undecidable in general, hence we will not speak of Δ_n -formulas. Nevertheless, we will sometimes say that a given property or a given set is provably Δ_n in a theory $\mathcal{T}$ when $\mathcal{T}$ is able to prove the equivalence: for instance, in IS_1 the provably Δ_1 sets coincide exactly with the primitive recursive sets by Proposition 3.3.5. This is also not to be confused with the notion of a Δ_n set inside a specific structure, which we will introduce in Chapter 9.

Definition 3.3.6 ($\text{I}\Delta_n$). For $n \in \mathbb{N}$, the Δ_n -induction scheme ($\text{I}\Delta_n$) is the theory comprised of PA^- together with, for every Σ_n - and Π_n -formulas $\varphi(x)$ and $\psi(x)$ (with possibly other free variables), the universal closure of the following axiom:

$$(\forall x \varphi(x) \leftrightarrow \psi(x)) \rightarrow \varphi(0) \wedge (\forall x \varphi(x) \rightarrow \varphi(x+1)) \rightarrow \forall x \varphi(x)$$

3.4 Collection

Other natural classes of first-order statements to consider are the collection schemes. They correspond to the axiomatic power needed to uniformly bound a finite number of existential witnesses.

Definition 3.4.1. For $\varphi(x, y)$ a formula (with possibly other free variables), the collection axiom associated to it is the universal closure of the formula:

$$\forall n[(\forall x < n \exists y \varphi(x, y)) \rightarrow \exists b \forall x < n \exists y < b \varphi(x, y)]$$

For Γ a class of formulas, we let $\text{B}\Gamma$ be the axiom scheme comprised of the collection axiom for every formula in Γ together with the axioms of IS_0 .

A conjunction or disjunction of Σ_n - (resp. Π_n -) formulas is logically equivalent to a Σ_n - (resp. Π_n -) formula. The following result shows that, assuming $\text{B}\Sigma_n$, we can add bounded quantifications in front of a Σ_n - (resp. Π_n -) formula and obtain a formula that is still equivalent to a Σ_n - (resp. Π_n -) formula.

Proposition 3.4.2 (Parsons [109]). *Let φ be Σ_n and ψ be Π_n , then $\text{B}\Sigma_n$ proves that $\forall x < y \varphi$ and $\exists x < y \psi$ are equivalent to a Σ_n - and a Π_n -formula, respectively.*

Proposition 3.4.3 (Kirby and Paris [107]). *For every $n \in \mathbb{N}$,*

- $\text{B}\Sigma_{n+1}$ and $\text{B}\Pi_n$ are logically equivalent. In particular $\text{B}\Sigma_1$ is equivalent to $\text{B}\Pi_0 = \text{B}\Sigma_0$.
- IS_{n+1} implies $\text{B}\Sigma_{n+1}$, which itself implies IS_n and the converse implications do not hold.

The following result relates the collection and induction schemes. The assumption that the exponential function is total (which was denoted by *exp*) is only relevant in the case $n = 1$ as both $\text{B}\Sigma_n$ and $\text{I}\Delta_n$ prove it when $n \geq 2$. It remains open, however, whether this assumption can be removed entirely.

Proposition 3.4.4 (Slaman [132]). *For every $n > 0$, $\text{B}\Sigma_n$ is equivalent to $\text{I}\Delta_n$ under the assumption that the exponential function is total.*

The case $n = 1$ of the following statement yields another conservation result: $\text{B}\Sigma_2$ is Π_3 -conservative over $\text{I}\Sigma_1$, hence, by Proposition 3.3.5, it is Π_2 -conservative over PRA.

Proposition 3.4.5 (Paris [106] see, e.g., [57, Theorem IV.1.59]). *For every $n \in \mathbb{N}$, $\text{B}\Sigma_{n+1}$ is Π_{n+2} -conservative over $\text{I}\Sigma_n$.*

By Kołodziejczyk, Wong, and Yokoyama [78, Theorem 4.9], for every $n \geq 1$, every proof of a Π_{n+2} -statement in $\text{B}\Sigma_{n+1}$ can be translated to a proof in $\text{I}\Sigma_n$ with at most a polynomial increase in size. A similar result also holds for $n = 0$ but requires the inclusion of *exp* in the theories.

These results were independently discovered by Fedor Pakhomov (personal communication from Kołodziejczyk), with an argument that appears to remove the need for the inclusion of *exp* in the case $n = 0$.

3.5 Bounded comprehension

The theory PRA is strong enough to handle encodings of finite sets via integers. For instance, the relation $x \in_{\text{ack}} y$, which holds precisely when the x -th bit in the binary expansion of y is 1, is a primitive recursive relation and can even be defined by a Σ_0 -formula, using the Σ_0 -formula $\text{pow}(x, y, z)$. One can then define the usual operations on such encodings, such as taking unions, computing cardinalities, or removing elements. Moreover, PRA proves the basic properties of these finite sets, such as the finite pigeonhole principle.

However, PRA does not in general prove the existence of a finite set defined by an arbitrary arithmetical formula. This motivates the consideration of the following hierarchies of statements:

Definition 3.5.1 (BC Γ). For $\varphi(x)$ a formula (with possibly other free variables), the bounded comprehension axiom associated to it is the universal closure of the formula:

$$\forall t \exists n \forall y (y \in_{\text{ack}} n \leftrightarrow (y < t \wedge \varphi(y)))$$

For Γ a class of formulas, we let $\text{BC}\Gamma$ be the axiom scheme comprised of the bounded comprehension axiom for every formula in Γ together with the axioms of PA^- .

For $n \in \mathbb{N}$, the Δ_n bounded comprehension scheme ($\text{BC}\Delta_n$) is the theory comprised of PA^- together with, for every Σ_n - and Π_n -formulas $\varphi(x)$ and $\psi(x)$ (with possibly other free variables), the universal closure of the following axiom:

$$(\forall x \varphi(x) \leftrightarrow \psi(x)) \rightarrow \forall t \exists n \forall y (y \in_{\text{ack}} n \leftrightarrow (y < t \wedge \varphi(y)))$$

This leads to the following equivalences between bounded comprehension schemes and the various levels of the induction hierarchy.

Proposition 3.5.2 (See e.g., Hájek and Pudlák [57]). *For every $n \in \mathbb{N}$,*

- $\text{I}\Sigma_0 + \text{exp} \vdash \text{BC}\Sigma_n \leftrightarrow \text{BC}\Pi_n \leftrightarrow \text{I}\Sigma_n$
- $\text{I}\Sigma_0 + \text{exp} \vdash \text{I}\Delta_n \leftrightarrow \text{BC}\Delta_n$

3.6 Well-foundedness

Another important hierarchy of statements that will feature prominently in this thesis concerns the well-foundedness of certain ordinals. Developing a general theory of ordinals requires strong set-theoretic axioms that are not finitistically available. Even when restricted to countable ordinals, substantial axiomatic strength is needed: as Simpson [128] observed, “ ATR_0 is the weakest set of axioms which permits the

development of a decent theory of countable ordinals”, where ATR_0 is a strong theory of second-order arithmetic (see Chapter 5). Fortunately, for our purposes, developing a general theory of countable ordinals is not required.

A linear order is said to be well-founded (or well-ordered) if it contains no infinite decreasing sequences (or, equivalently, if every non-empty subset has a least element). The general theory of ordinals investigates such orders through order-preserving functions, aiming to classify and compare them. Ordinals are usually defined either as equivalence classes under order-preserving bijections, or by picking some representative of these classes, where the set membership relation $\in$ is the order relation. The ordinals themselves can then be linearly ordered, depending on whether they can be embedded into each other by an order-preserving function.

Here, as we do not have at our disposal the tools to prove the basic properties of this equivalence relation, or to define those sets, we will go in a different direction with the use of ordinal notations.

An ordinal notation is a way to encode a set of ordinals with natural numbers, so that the set of codes and the induced order relation between them are computable (and even primitive recursive in our case). Here, we will only consider a notation for ordinals below the ordinal ε_0 (where ε_0 is defined as the first solution to the equation $\omega^\alpha = \alpha$ or, equivalently, as the supremum of the sequence $(\omega_n)_{n \in \mathbb{N}}$, where $\omega_n := \omega_n^1$ and the ω_n^α are defined inductively by letting $\omega_0^\alpha := \alpha$ and $\omega_{n+1}^\alpha := \omega^{\omega_n^\alpha}$ for $\alpha < \varepsilon_0$). Our notation will be based on the Cantor normal form theorem, which can be used to represent ordinals below ε_0 as trees. Another desirable feature of this notation is the fact that the basic ordinal operations on the codes, such as addition, multiplication, exponentiation will be computable (and even primitive recursive).

We will then study the logical strength of the statement asserting the well-foundedness of the induced ordering.

Theorem 3.6.1 (Cantor normal form). *For every ordinal α , there exists a finite sequence $\beta_0 < \dots < \beta_{k-1}$ of ordinals and a finite sequence $n_0, \dots, n_{k-1} \in \mathbb{N} \setminus \{0\}$ of integers such that:*

$$\alpha = \omega^{\beta_{k-1}} \times n_{k-1} + \dots + \omega^{\beta_0} \times n_0$$

This expression is called the Cantor normal form of α , and it is unique.

For an ordinal α smaller than ε_0 , the coefficients β_i in its Cantor normal form are strictly smaller than α , and we can inductively replace them by their own Cantor normal form to express α using only integers, ω , ordinal addition, multiplication, and exponentiation (e.g., $\omega^{\omega^\omega \times 5 + \omega^2 \times 3 + \omega} \times 4 + \omega^{\omega^2 \times 8 + 1} \times 6 + 17$). Thus, every ordinal below ε_0 can be written this way using a finite number of symbols, hence can be encoded by some integer.

The statement asserting the well-foundedness of a countable order is a second-order statement, as it quantifies over all the possible sequences of elements of that order. Here, we consider a restricted version of that statement, which quantifies only over sequences that are Σ_1 . More precisely, a Σ_1 sequence refers to a total function $f : \mathbb{N} \rightarrow \mathbb{N}$ whose graph $\{(x, f(x)) : x \in \mathbb{N}\}$ is encoded by a Σ_1 -formula $\exists s \varphi(x, y, s)$ which holds exactly when $y = f(x)$. Since f is total, its graph can equivalently be given by a Π_1 -formula, namely $\forall s \forall x', x' \neq x \rightarrow \neg \varphi(x', y, s)$. As we will see in Chapter 5, the second-order version will be of similar strength.

Definition 3.6.2 ($\text{WF}(\alpha)$). For every $\alpha \leq \varepsilon_0$, let $\text{WF}(\alpha)$ be the statement that there is no infinite decreasing Σ_1 sequence of ordinals smaller than α

The statement $\text{WF}(\alpha)$ is a Π_3 -statement for every ordinal $\alpha < \varepsilon_0$. It is clear that if $\alpha > \beta$, then $\text{PRA} \vdash \text{WF}(\alpha) \rightarrow \text{WF}(\beta)$.

The following result, due to Gentzen, may be viewed as a partial realization of Hilbert’s program and is the foundational result of ordinal analysis, a field of logic that assigns to every theory an ordinal measuring its proof-theoretic strength. Although

Gödel's incompleteness theorem shows that PA cannot prove its own consistency (and in particular that $\text{PRA} \not\vdash \text{Con}(\text{PA})$), the system $\text{PRA} + \text{WF}(\varepsilon_0)$ is able to prove $\text{Con}(\text{PA})$. Consequently, PA itself cannot prove $\text{WF}(\varepsilon_0)$.

Theorem 3.6.3 (Gentzen [49]). $\text{PRA} + \text{WF}(\varepsilon_0) \vdash \text{Con}(\text{PA})$.

More generally, an analysis of Gentzen's proof yields $\text{PRA} + \text{WF}(\omega_{n+1}) \vdash \text{Con}(\text{IS}_n)$.

In particular, $\text{PRA} + \text{WF}(\omega^\omega) \vdash \text{Con}(\text{IS}_1)$, which is a Π_1 -statement not provable in IS_1 by Gödel's incompleteness theorem, hence $\text{PRA} + \text{WF}(\omega^\omega)$ is not Π_1 -conservative over PRA and is not finitistically reducible.

The following result, due to Kreuzer and Yokoyama, shows that $\text{WF}(\omega^\omega)$ is equivalent to several other well-studied fragments of first-order arithmetic.

Proposition 3.6.4 (Kreuzer and Yokoyama [84, Theorem 5]). *Over IS_1 , the following are equivalent,*

- $\text{WF}(\omega^\omega)$
- *The totality of all the functions defined by the same scheme as the Ackermann function (from a total function $f : \mathbb{N} \rightarrow \mathbb{N}$, we define $A_f : \mathbb{N}^2 \rightarrow \mathbb{N}$ by $A_f(0, n) = f(n)$, $A_f(m+1, 0) = A_f(m, 1)$ and $A_f(m+1, n+1) = A_f(m, A_f(m+1, n))$. The Ackermann function is simply the case where f is the successor operation).*
- $\text{BME}_\star$ (which will be introduced in a second-order form in Chapter 10)
- $\text{P}\Sigma_1$, which is the scheme comprised of PA^- , together with, for every Σ_1 -formula $\varphi(x, y)$ (with possibly other free variables), the universal closure of the axioms stating that: if $\varphi(x, y)$ encodes a partial function, then for every $k \in \mathbb{N}$, there exists a sequence $(s_i)_{i \leq k}$ of integers, such that, for every $i < k$, if $\varphi(x, y)$ holds for some $x \leq s_i$, then necessarily $y \leq s_{i+1}$.

However, these equivalences do not generalize well to other ordinals of the form ω_n for $n > 2$. The scheme $\text{P}\Sigma_n$, which generalizes $\text{P}\Sigma_1$ to all Σ_n -formulas $\varphi(x, y)$, lies strictly between IS_{n+1} and IS_n , while being incomparable with $\text{B}\Sigma_{n+1}$, and, by the following proposition, even $\text{PRA} + \text{WF}(\varepsilon_0)$ is not strong enough to prove IS_2 (and hence $\text{P}\Sigma_2$). Thus, the theories $\text{PRA} + \text{WF}(\alpha)$ are of peculiar strength: they are able to prove the consistency of certain strong theories of arithmetic, yet are not able to establish the truth of their axioms.

Proposition 3.6.5 (See e.g., [57, Theorem II.3.18 and Theorem IV.1.33]).

- *For every $n, k \in \mathbb{N}$, $\text{IS}_n \vdash \text{WF}(\omega_n^k)$, in particular, $\text{IS}_2 \vdash \text{WF}(\omega^\omega)$ and the IS_n are not finitistically reducible for $n > 1$.*
- *If φ is a Π_3 -formula consistent with $\text{B}\Sigma_2$, then $\text{B}\Sigma_2 + \varphi \not\vdash \text{IS}_2$. In particular, $\text{PRA} + \text{WF}(\varepsilon_0) \not\vdash \text{IS}_2$.*

Now that we have defined several hierarchies of fragments of first-order arithmetic, we will see in the next chapter various connections between the arithmetical hierarchy and computability-theoretic notions. These connections make it easier to work with arithmetical formulas by relating them to more intuitive concepts, such as the termination of an algorithm. They also provide a general form for Σ_n - and Π_n -formulas, which allows us to establish the existence of universal Σ_n - and Π_n -formulas, respectively, and deduce that the theories IS_n , $\text{B}\Sigma_{n+1}$, and $\text{IS}_1 + \text{WF}(\omega_n)$ are in fact finitely axiomatizable for $n > 0$.

Chapter 4

Computability

Among the infinite objects studied by mathematicians, not all are equally problematic. Even setting aside the very large sizes of infinity studied by set theorists, there is already a distinction to be made between the different subsets of $\mathbb{N}$. Indeed, some of these subsets can be completely described as consisting of the integers satisfying a simple property expressible with a finite number of symbols. In such cases, it is not entirely clear what is gained by treating them as completed entities, rather than simply replacing their use by their finite definition (i.e., saying “let n be such that $\varphi(n)$ ” instead of saying “let $n \in X$ ”). In contrast, other subsets cannot be defined by a finite number of symbols, and the question of what it means for such sets to “exist” raises deep philosophical concerns.

In the next chapter, we will see that some set existence axioms are indeed finitistically reducible. This chapter introduces the tools and definitions from computability theory which will be used to establish those results, by classifying sets based on their computability-theoretic content.

We will not go into detail on how to formalize what is an algorithm. In line with the Church-Turing thesis, we assume that our intuitive notion of what can be computed can be formalized, and that all algorithmic procedures described using natural language in this thesis could also be expressed using Turing machines, lambda calculus, or any other Turing-complete formalism.

The following facts and notations will be useful in our subsequent descriptions of algorithms:

- We regard all computable functions as being partial functions from $\mathbb{N}$ to $\mathbb{N}$, or equivalently as partial functions from $\{0, 1\}^{<\mathbb{N}}$ to $\{0, 1\}^{<\mathbb{N}}$, where $A^{<\mathbb{N}}$ denotes the set of finite sequences of elements of A (we also introduce the notation $A^{<n}$ and $A^{\leq n}$ for sequences of elements of A of length $< n$ and $\leq n$, respectively). For functions taking more than one parameter as input, or a different type of data, we rely on the fact that such inputs can be computably encoded as a single natural number. On a given input, an algorithm either halts and returns an integer, or never halts and the corresponding function is undefined for that input.
- There exists an enumeration $(\Phi_e)_{e \in \mathbb{N}}$ containing all the computable functions and such that this enumeration is effective, in the sense that there exists some computable function Φ (hence belonging to the list) such that $\Phi(\langle e, x \rangle)$ is equal to $\Phi_e(x)$ (and is undefined if $\Phi_e(x)$ is). The index e is said to be the *Turing index* of Φ_e .
- We write $\Phi_e(x) \downarrow$ to specify that Φ_e halts on input x , and we write $\Phi_e(x) \downarrow = v$ to specify that it halts and returns v . Conversely, we write $\Phi_e(x) \uparrow$ if the computation never halts.

- We assume that our formal model allows a computation to be broken down into a sequence of discrete steps, and we write $\Phi_e(x)[s]$ for the result of the computation after s steps (we write $\Phi_e(x)[s]\downarrow$ or $\Phi_e(x)[s]\uparrow$ to specify whether the computation has already halted after those s steps).
- The function defined by $(e, x, s) \mapsto 0$ if $\Phi_e(x)[s]\uparrow$ and $(e, x, s) \mapsto v + 1$ if $\Phi_e(x)[s]\downarrow = v$ can be assumed to be primitive recursive.

Definition 4.0.1 (computable set). A set X is said to be *computable* if its characteristic function is computable.

A set X is said to be *computably enumerable* (or c.e.) if $X = W_e$, for some Turing index e , where W_e is the domain of Φ_e .

Thus, a set X is computable if there exists an algorithm that takes an integer n as input, and answers whether n is in X or not. A set X is computably enumerable if there is an algorithm that takes an integer n as input, and only halts if n is in X .

Proposition 4.0.2. *A set $X \subseteq \mathbb{N}$ is computable if and only if X and its complement $\mathbb{N} \setminus X$ are computably enumerable.*

Contrary to what its name suggests, computability theory is mainly concerned with what is not computable. The field examines mathematical objects and theorems through the lens of their computational content, and from this point of view, the computable cases are seen as trivial. In the next section, we will see how to compare the computational content of subsets of $\mathbb{N}$.

4.1 Relativized computability and arithmetical sets

The notion of a computable function can be relativized to define computation relative to a given set A , called the oracle. In this context, it is assumed that the computation has access to all the bits of A (for example, by adding another tape to a Turing machine with A written on it, or in lambda calculus by including the characteristic function of A as a primitive in the language). We can then talk about A -computable functions, consider Φ_e^A , the e -th such function and define W_e^A as the A -c.e. sets.... More generally, this relativization can be done in a uniform manner, and there exists a computable list $(\Phi_e^X)_{e \in \mathbb{N}}$ of the so-called Turing functionals, where the oracle X is seen as a parameter. For example, we can consider the code e of the functional Φ_e^X that searches the first element of a given set X and outputs it if there is one.

Since every computation that halts does so in a finite number of steps, the result of it only depends on a finite number of bits in the set A . We can also assume that in s steps of computation, we only have access to the first s elements of A . This gives Proposition 4.1.1, which can be seen as stating the continuity of the total functionals on the Cantor space $\{0, 1\}^{\mathbb{N}}$ (for the product topology). We first introduce notation for working with finite sets and finite binary sequences.

By abuse of notation, we identify subsets of $\mathbb{N}$, infinite binary sequences, and functions from $\mathbb{N}$ to $\{0, 1\}$. We will also make use of set-theoretic notations, writing k for the set $\{0, \dots, k-1\}$. Thus, $2^{\mathbb{N}}$ denotes the Cantor space and 2^n the set of binary sequences of length n . For finite subsets $E \subseteq \mathbb{N}$ (which we denote $E \subseteq_{\text{fin}} \mathbb{N}$), the link with finite binary sequences is less direct, since two sequences that differ only by trailing zeros represent the same set.

We will often view finite sets and finite binary sequences as approximations of infinite sets. For this reason, both notations have their own purpose: finite subsets contain only positive information about which elements belong to the infinite set, whereas finite binary sequences contain both positive and negative information about which elements do or do not belong to the set.

We write $\sigma \leq \tau$ for the prefix relation between two binary sequences (finite or infinite). $|\sigma|$ will denote the length of a sequence σ and $\text{card } E$ will denote the

cardinality of the set E . For $k \in \mathbb{N}$, we write $\sigma \upharpoonright k$ for the sequence σ restricted to the domain $\{0, \dots, k-1\}$. We will abuse notation and often write $\sigma \subseteq E$ (meaning that $\sigma(i) = 1$ implies $i \in E$ for every $i < |\sigma|$), $E \upharpoonright k$ (for $E \cap \{0, \dots, k-1\}$) and $E(i)$ (equal to 1 if $i \in E$ and 0 otherwise).

Proposition 4.1.1 (use property). *Let $A \subseteq \mathbb{N}$ and $x, e, v \in \mathbb{N}$. If $\Phi_e^A(x)[s] \downarrow = v$, then for every $B \subseteq \mathbb{N}$ with $A \upharpoonright s = B \upharpoonright s$, $\Phi_e^B(x) \downarrow = v$.*

Another consequence is that it is not necessary to specify all the bits of the oracle. For a finite binary sequence σ , we define Φ_e^σ to be the computation that behaves like Φ_e^A for any set $A \subseteq \mathbb{N}$ extending σ (i.e., with $\sigma < A$) up to the $|\sigma|$ -th step, and if the computation has not halted at the $|\sigma|$ -th step, Φ_e^σ enters an infinite loop.

This relativization of the notion of computability allows us to compare sets by the amount of information they contain.

Definition 4.1.2 (Turing reducibility). For $A, B \subseteq \mathbb{N}$, we say that A is *Turing-reducible* to B and we write $A \leq_T B$ if A is B -computable.

The structure $(\mathcal{P}(\mathbb{N}), \leq_T)$ is a preorder. One may then define the equivalence relation $\equiv_T$ (Turing equivalence) by letting $A \equiv_T B$ if $A \leq_T B$ and $B \leq_T A$. An equivalence class for $\equiv_T$ is called a *Turing degree* and we let $\mathcal{D}$ be the set of all Turing degrees. By construction, $\leq_T$ is compatible with the quotient by $\equiv_T$. The resulting order $(\mathcal{D}, \leq_T)$ is a join-semilattice, with minimal element $\mathbf{0}$, the degree of $\emptyset$, i.e., the class containing exactly the computable sets, and with join $\oplus$ induced by the following operation on sets (which is compatible with $\equiv_T$).

Definition 4.1.3 (join). For $A, B \subseteq \mathbb{N}$, we define $A \oplus B$, the *join* of A and B , to be the set $2A \cup (2B + 1) = \{2n : n \in A\} \cup \{2n + 1 : n \in B\}$.

The join operator as defined here is not associative, but we can easily compute $(A \oplus B) \oplus C$ from $A \oplus (B \oplus C)$ and vice versa, thus we will abuse notation and write $A_0 \oplus \dots \oplus A_{n-1}$ for the join of n sets.

Proposition 4.1.4. *For $A, B, C, D \subseteq \mathbb{N}$, if $A \equiv_T B$ and $C \equiv_T D$, then $A \oplus C \equiv_T B \oplus D$, thus $\oplus$ can be defined for degrees.*

We will also talk of the *join* of $A_0, A_1, \dots \subseteq \mathbb{N}$, an infinite countable family of sets and define it to be the set $\bigoplus_{n \in \mathbb{N}} A_n := \{\langle k, n \rangle : k \in A_n\}$. However, it should be noted that this definition cannot be applied to degrees: if $B_0, B_1, \dots \subseteq \mathbb{N}$ is another infinite family with $A_n \equiv_T B_n$ for every $n \in \mathbb{N}$, then $\bigoplus_{n \in \mathbb{N}} A_n$ is not necessarily Turing equivalent to $\bigoplus_{n \in \mathbb{N}} B_n$. In fact, $(\mathcal{D}, \leq_T)$ is not complete in a strong sense: by Spector [135], every countable increasing sequence of degrees has no least upper bound.

Definition 4.1.5 (Turing jump). For $A \subseteq \mathbb{N}$, let $A' = \{e : \Phi_e^A(e) \downarrow\}$ be its *Turing jump* (also called the halting set relativized to A).

A' is an A -c.e. set, and we define, for each $s \in \mathbb{N}$, $A'[s] = \{e : \Phi_e^A(e)[s] \downarrow\}$, its A -c.e. approximation. We have $A'[0] \subseteq A'[1] \subseteq \dots$, and $A' = \bigcup_{s \in \mathbb{N}} A'[s]$. The *modulus* of A' is the function $\mu : \mathbb{N} \rightarrow \mathbb{N}$ which sends $x \in \mathbb{N}$ to the first integer $s \in \mathbb{N}$ such that $A'[s] \upharpoonright x = A' \upharpoonright x$.

Proposition 4.1.6. *For every set $A, B \subseteq \mathbb{N}$:*

- $A \leq_T A'$ in a uniform manner: there exists some index e such that for every set C , $\Phi_e^{C'} is the characteristic function of C .$
- $A' \not\leq_T A$ (undecidability of the halting problem, relativized to any set A)
- Let $\mu : \mathbb{N} \rightarrow \mathbb{N}$ be the modulus of A' , and let $f : \mathbb{N} \rightarrow \mathbb{N}$ be such that $\mu(n) \leq f(n)$

for every $n \in \mathbb{N}$ (that is, f dominates μ). Then $A \oplus f \geq_T A'$.

- If $A \leq_T B$, then $A' \leq_T B'$, hence we can define the jump $\mathbf{d}'$ of a degree $\mathbf{d}$ to be the degree of the jump of any of its elements.
- $A' = W_e^A$, for some Turing index e (which does not depend on A).

We write $A^{(n)}$ for the iterated Turing jumps of a set A (i.e., $A^{(0)} = A$ and $A^{(n+1)} = A^{(n)'}$), and similarly for degrees. We can continue iterating the jump operator into the transfinite by letting $A^{(\omega)} := \bigoplus_{n \in \mathbb{N}} A^{(n)}$. It should be noted that the degree of $A^{(\omega)}$ is not the least upper bound of the degrees of the $A^{(n)}$ for $n \in \mathbb{N}$ (A set can compute $A^{(n)}$ for every $n \in \mathbb{N}$, but may not compute $A^{(\omega)}$ if the $A^{(n)}$ are not computed in a uniform manner).

Let $A, B \subseteq \mathbb{N}$ be such that B is A' -computable. Then there exists an index $e \in \mathbb{N}$ such that $\Phi_e^{A'}(x) \downarrow = B(x)$ for every $x \in \mathbb{N}$. For every $s \in \mathbb{N}$, define $B_s = \{x < s : \Phi_e^{A'}(x)[s] \downarrow = 1\}$. The sequence $(B_s)_{s \in \mathbb{N}}$ is A -computable and satisfies that $\lim_{s \rightarrow \infty} B_s(x)$ exists and is equal to $B(x)$ for every $x \in \mathbb{N}$.

Indeed, let μ be the modulus of A' and fix $x \in \mathbb{N}$. Since $\Phi_e^{A'}(x) \downarrow$, there exists $y \in \mathbb{N}$ such that $\Phi_e^{A'} \upharpoonright y(x) \downarrow$ and, since $A'[s] \upharpoonright y = A' \upharpoonright y$ for every $s \geq \mu(y)$, we have $\Phi_e^{A'}(x)[s] \downarrow = B(x)$ for every sufficiently large s . From this, we obtain the following result, which connects computations relative to the jump of a set with computations in the limit.

Proposition 4.1.7 (Shoenfield's limit Lemma). *Let $A \subseteq \mathbb{N}$. A set B is A' -computable if and only if there exists an A -computable function $g : \mathbb{N}^2 \rightarrow \mathbb{N}$ such that $\lim_{n \rightarrow \infty} g(n, x)$ exists for every $x \in \mathbb{N}$ and is equal to $B(x)$.*

We now have sufficient background to classify the sets defined by arithmetical formulas in terms of computable sets and repeated applications of the jump operator.

To the syntactical notion of a Σ_n - or Π_n -formula, we associate the semantical notion of a Σ_n or Π_n set.

Definition 4.1.8 (arithmetical, Σ_n , Π_n and Δ_n sets). Let $\varphi(x)$ be an arithmetical formula with x its only free variable, and let $X = \{n : \varphi(n)\}$. The set X is said to be *arithmetical*.

If φ is a Σ_n - (resp. Π_n -) formula, then X is also said to be Σ_n (resp. Π_n).

A set $X \subseteq \mathbb{N}$ is said to be Δ_n if it is both Σ_n and Π_n .

The notion of a Δ_n set has no syntactical equivalent, as the equivalence between a Σ_n - and a Π_n -formula is not a syntactic property (and in fact not even decidable). The following central theorem of computability theory relates the levels of the arithmetical hierarchy to iterations of the Turing jump. The level $n = 1$ is due to Kleene [76] building on the work of Gödel [51].

Theorem 4.1.9 (Post's theorem [116]). *Let $A \subseteq \mathbb{N}$ and $n > 0$. Then, the following hold:*

- A is Δ_n if and only if $A \leq_T \emptyset^{(n-1)}$
- A is Σ_n if and only if A is $\emptyset^{(n-1)}$ -c.e.

One consequence of this theorem and the existence of a computable list of all the computable functions, is the existence, for every $n \geq 1$, of a universal Σ_n -formula (here the word universal does not refer to universal quantifiers, but rather to the fact that the formula is universal among the other Σ_n -formulas), i.e., a Σ_n -formula $\varphi(x)$ such that PRA proves that for every other Σ_n -formula $\psi(y)$ there exists an index $e \in \mathbb{N}$ such that $\psi(y)$ is equivalent to $\varphi(\langle e, y \rangle)$. Similarly, by taking its negation, we also obtain a universal Π_n -formula. See Hájek and Pudlák [57, I.1] for an explicit construction of such formulas.

Using this universal Σ_n -formula, we obtain that the theories $I\Sigma_n, I\Pi_n, I\Delta_{n+1}, B\Sigma_{n+1}, B\Pi_n$ and $I\Sigma_1 + WF(\alpha)$ are all finitely axiomatizable for $n > 0$ (as they all contain PRA and thus prove the existence of this universal formula). More precisely, for every $n > 0$, $I\Sigma_n$ and $I\Delta_n$ can be axiomatized by a single Π_{n+2} -formula and $B\Sigma_{n+1}$ by a single Π_{n+3} -formula.

A general phenomenon in computability theory is the ability to relativize a given result. For a fixed set A , almost all results remain valid when every occurrence of “computable” is replaced by “computable with oracle A ”. In Chapter 5, we will likewise relativize the arithmetical hierarchy by allowing set parameters to appear in formulas. The relativization of Post’s theorem will give us a characterization of the $X^{(n-1)}$ -computable sets and the $X^{(n-1)}$ -c.e. sets in terms of formulas with X as a parameter.

4.2 PA degrees

By Gödel’s first incompleteness theorem, no complete extension (or completion) of PA^- is computable. The canonical example of such a theory is the theory of true arithmetic, that is, the set of all first-order sentences satisfied by the structure $(\mathbb{N}, 0, succ, +, \times, <)$. This theory is far from being computable, as it is Turing equivalent to $\mathcal{O}^{(\omega)}$. Nevertheless, there exist completions that are computationally less complex. In this section, we investigate the exact computable strength needed to produce a complete extension of PA^- and of other stronger systems such as PA or ZFC.

Definition 4.2.1 (PA degree). A degree $\mathbf{d}$ is a PA *degree* if it computes a complete extension of PA.

We can represent all the possible completions of a computable theory $\mathcal{T}$ as the infinite paths of some computable infinite binary tree T . The idea is to make every level n of our tree represent a choice: do we take the n -th formula φ_n or its negation in our theory, and if our choices lead to a contradiction, then we simply cut the corresponding branch. Thus defined, this tree is not computable, as we cannot decide whether a given theory is consistent or not. Nevertheless, if a theory is inconsistent, then a proof of a contradiction will eventually be found, and we can rather define T as follows: a given binary sequence $\sigma \in 2^s$ belongs to T if and only if there is no proof of size $\leq s$ of a contradiction using the formulas encoded by the bits of σ together with the axioms of $\mathcal{T}$ of size $\leq s$. This property is indeed computable, and the infinite paths of T are exactly those encoding completions of $\mathcal{T}$ (as we have made a choice between φ or $\neg\varphi$ for every formula φ).

If the theory $\mathcal{T}$ is consistent and contains PA^- , then this construction defines an infinite binary computable tree with no infinite computable path (by Gödel’s first incompleteness theorem). The statement that every infinite binary tree has an infinite path is known as weak König’s lemma (or WKL), and our construction shows that it is not constructively valid (to believe in the truth of that statement requires believing in the existence of some non-computable sets). Nevertheless, as we will see in the next chapter, the statement WKL is finitistically reducible.

WKL will play an important role in this thesis, in part due to the following equivalence, which gives several characterizations of the sets of paths through a computable tree. WKL also plays an important role in reverse mathematics, as we will see in Chapter 5.

Proposition 4.2.2. Let $\mathcal{P} \subseteq 2^{\mathbb{N}}$, the following are equivalent:

- There exists a computable tree T such that $\mathcal{P}$ is the set of all infinite paths in T (denoted $[T]$).
- There exists a computable set I of binary sequences, such that $2^{\mathbb{N}} \setminus \mathcal{P} = \bigcup_{\sigma \in I} [\sigma]$,

where $[\sigma]$ denotes the basic open set $\{X \in 2^{\mathbb{N}} : \sigma < X\}$ (i.e., $\mathcal{P}$ is effectively closed in the product space $2^{\mathbb{N}}$)

- There exists a Turing functional Φ^X that takes no input, and such that $\mathcal{P} = \{A : \Phi^A \uparrow\}$

A class $\mathcal{P}$ satisfying those properties is said to be a Π_1^0 class.

The standard proof of WKL proceeds as follows. Let T be an infinite binary tree. We define an infinite path X through T as $X = \bigcup_{n \in \mathbb{N}} \sigma_n$ where $(\sigma_n)_{n \in \mathbb{N}}$ is a sequence of elements of T satisfying $\sigma_0 < \sigma_1 < \dots$, such that for every $n \in \mathbb{N}$, $|\sigma_n| = n$ and there are infinitely many elements of T extending σ_n . Such a sequence is defined inductively as follows: let σ_0 be the root of T . Since T is infinite, there are infinitely many elements of T extending σ_0 . Now assume σ_n has been defined for some $n \in \mathbb{N}$. For $i < 2$, let $A_i = \{\sigma \in T : \sigma \geq \sigma_n \cdot i\}$, then, as $\{\sigma \in T : \sigma \geq \sigma_n\} = \{\sigma_n\} \cup A_0 \cup A_1$ is infinite, at least one of the A_i is infinite and let $\sigma_{n+1} = \sigma_n \cdot i$ for such an i . Continuing this process inductively yields an infinite path through T .

The set X defined above is not computable, since we cannot effectively find an $i < 2$ for which A_i is infinite. In general, deciding whether a computable set $X \subseteq \mathbb{N}$ is infinite requires $\emptyset''$, as the formula $\forall x \exists y y > x \wedge y \in X$ is equivalent to a Π_2 -formula. However, due to the specific form of the sets A_i , namely, their closure under prefixes for elements extending $\sigma_n \cdot i$, this could be reduced to a Π_1 -formula: A_i is infinite if and only if for every level $\ell > |\sigma_n|$ of the tree, we have $A_i \cap 2^\ell \neq \emptyset$. This allows the infiniteness of A_i to be decided using only $\emptyset'$. We can go beyond that by noticing that the construction does not require deciding, for each $i < 2$, whether A_i is finite or not, but only to find one index $i < 2$ such that A_i is infinite. As we will see, this can be done without having to compute $\emptyset'$.

Definition 4.2.3 (DNC₂ degree). A degree $\mathbf{d}$ is a DNC₂ degree (for diagonally non-computable) if it computes a total function $f : \mathbb{N} \rightarrow \{0, 1\}$ such that $f(n) \neq \Phi_n(n)$ for every $n \in \mathbb{N}$. Any such function is referred to as a DNC₂ function.

Consider the program Φ_e that performs a breadth-first search on the trees A_0 and A_1 in parallel, and outputs $i < 2$ if it discovers that A_i is finite, while never halting if both A_0 and A_1 are infinite. Then, for any DNC₂ function f , the value of $f(e)$ indicates which branch to follow in order to continue the construction of our path. As this could be done in a uniform manner, any DNC₂ degree computes a path through any infinite computable subtree of $2^{\mathbb{N}}$.

Conversely, given a complete theory $\mathcal{T}$ extending PA (in fact, PRA suffices), we can compute a DNC₂ function f as follows. For each $e \in \mathbb{N}$, search for a proof in $\mathcal{T}$ of a statement of the form $\neg(\Phi_e(e) \downarrow = i)$ for some $i < 2$, and define $f(e) = i$. This function is well-defined. Indeed, PRA proves that $\Phi_e(e)$ cannot take two distinct values, so $\mathcal{T}$ cannot prove $\Phi_e(e) \downarrow = i$ for both $i < 2$. By completeness of $\mathcal{T}$, it must therefore prove $\neg(\Phi_e(e) \downarrow = i)$ for at least one $i < 2$, ensuring that the search eventually succeeds. Finally, by the Σ_1 -completeness of PRA, the function f is indeed DNC₂: if $\Phi_e(e) \downarrow = i$, then this fact is provable in PRA, and hence cannot be refuted in $\mathcal{T}$, so $f(e) \neq i$.

Thus, we have the following characterization of PA degrees:

Proposition 4.2.4. A degree $\mathbf{d}$ is PA if and only if it is DNC₂ if and only if it computes an element of any non-empty Π_1^0 class.

As stated before, we don't need to go as high as $\emptyset^{(\omega)}$ or $\emptyset'$ to find sets of PA degree.

Theorem 4.2.5 (Jockusch and Soare [71] (cone avoidance basis theorem)). For every non-computable degree $\mathbf{d}$, there exists a PA degree $\mathbf{a}$ such that $\mathbf{d} \not\leq_T \mathbf{a}$. In particular, there exists some PA degree $\mathbf{a}$ such that $\mathbf{a} \not\geq_T \mathbf{0}'$.

The jump operator is not injective on Turing degrees: there exist multiple degrees whose jump is equal to $\mathbf{0}'$. Such degrees are called *low*, and they will play an important role in this thesis. Indeed, by the relativization of Post's theorem, if X is a low set, then every Σ_2 - and Π_2 -formula in which we allow X to appear as a parameter (which will be referred to as $\Sigma_2(X)$ - and $\Pi_2(X)$ -formulas in the next chapter) is equivalent to a Σ_2 - and a Π_2 -formula respectively, and we will use this fact to show that, from a proof-theoretic viewpoint, assuming the existence of some low sets is often no more problematic than working with computable sets.

Definition 4.2.6 (low degrees). A degree $\mathbf{d}$ is *low* if $\mathbf{d}' \equiv_T \mathbf{0}'$. For $n > 0$, we say that $\mathbf{d}$ is *low_n* if $\mathbf{d}^{(n)} \equiv_T \mathbf{0}^{(n)}$.

Theorem 4.2.7 (low basis theorem (see e.g., [57, I.3.8])). *There exist PA degrees that are low.*

Again, all the results in this section relativize. For example, Proposition 4.2.2 remains true when considering A -computable trees T , A -computable sets I of binary sequences and Turing functionals $\Phi^{X \oplus A}$, resulting in the definition of a $\Pi_1^0(A)$ class. Similarly, we can say that a set B is of *PA degree relative to A* (which we write $A \ll B$) or *PA over A* , and say that a degree $\mathbf{d}$ is *low relative to $\mathbf{a}$* if $(\mathbf{a} \oplus \mathbf{d})' \equiv_T \mathbf{a}'$.

Chapter 5

Second-order arithmetic

We have seen in Chapter 3 several first-order theories extending PRA, and we have compared their strength as well as whether they are finitistically reducible. In this chapter, we will extend PRA in another direction, by adding the possibility to talk and quantify not only over integers, but also on sets of integers.

Contrary to Zermelo-Fraenkel set theory (ZFC) where everything is a set, the language of second-order arithmetic $\mathcal{L}_{Z_2}$ is two-sorted, meaning that the universe of discourse comprises two distinct kinds of objects (i.e., the integers and sets of integers).

The language of second-order arithmetic extends the one of first-order arithmetic with the addition of set variables (written in uppercase $A, B, X, Y, \dots$), set quantifiers (written $\forall X$ and $\exists X$) and of one binary relation: the membership relation $\in$ between one integer and a set of integers.

We can now extend the arithmetical hierarchy to include statements involving set variables and quantifiers.

5.1 Relativized arithmetical and analytical hierarchies

Before classifying second-order formulas by their number of alternations of second-order quantifiers, we can consider second-order formulas with no second-order quantifiers (i.e., where the set variables are not quantified over and appear free). Those formulas can be seen as almost first-order, but relative to a collection of set parameters.

For a (possibly infinite) collection of set variables $X_0, X_1, \dots$, the $\Sigma_n(X_0, \dots)$ - and $\Pi_n(X_0, \dots)$ -formulas are defined in the same way as the Σ_n - and Π_n -formulas, except that atomic formulas of the form $x \in X_i$ are allowed in the base case (e.g., $\varphi(X) := \forall x \exists y \ y > x \wedge y \in X$ is a $\Pi_2(X)$ -formula asserting that the set X is infinite).

In the absence of second-order quantification, it is often helpful to view set variables not as second-order objects, but rather as unary relations on $\mathbb{N}$ without any additional assumptions.

More generally, we will often not specify the set parameters and simply say that a formula is Σ_n^0 (respectively Π_n^0) if there exist some set variables $\vec{X}$ such that the formula is $\Sigma_n(\vec{X})$ (respectively $\Pi_n(\vec{X})$).

For $n > 0$ and $C_0, C_1, \dots \subseteq \mathbb{N}$ a (possibly infinite) sequence of sets, we define similarly the $\Sigma_n(C_0, C_1, \dots)$, $\Pi_n(C_0, C_1, \dots)$ and $\Delta_n(C_0, C_1, \dots)$ sets. The link between computability and the arithmetical hierarchy can also be relativized: a set A is $\Delta_n(C_0, C_1, \dots)$ if and only if it is $(C_{i_0} \oplus \dots \oplus C_{i_{k-1}})^{(n-1)}$ -computable for some indices $i_0 < \dots < i_{k-1}$ and A is $\Sigma_n(C_0, C_1, \dots)$ if and only if it is $(C_{i_0} \oplus \dots \oplus C_{i_{k-1}})^{(n-1)}$ -c.e. for some indices $i_0 < \dots < i_{k-1}$.

Definition 5.1.1 (Σ_n^1 - and Π_n^1 -formulas). We define inductively the classes Σ_n^1 and Π_n^1 of formulas for an integer n :

- The classes Σ_0^1 and Π_0^1 are equal and correspond to formulas with no set quantifiers, or equivalently to the arithmetic formulas with second-order parameters.
- If φ is Σ_n^1 , then $\forall X_0 \forall X_1 \dots \forall X_{k-1} \varphi(\vec{X})$ is Π_{n+1}^1 for every $k \in \mathbb{N}$.
If ψ is Π_n^1 , then $\exists X_0 \exists X_1 \dots \exists X_{k-1} \psi(\vec{X})$ is Σ_{n+1}^1 for every $k \in \mathbb{N}$.

As with first-order quantifiers, we will work in theories strong enough to combine several second-order universal (resp. existential) quantifiers into one. To do so, it is sufficient to be able to prove the existence of the join $A_0 \oplus \dots \oplus A_{k-1}$, from the existence of $A_0, \dots, A_{k-1}$. We can then write $\forall X_0 \oplus \dots \oplus X_{k-1} \varphi(\vec{X})$ in place of $\forall X_0 \dots \forall X_{k-1} \varphi(\vec{X})$.

The existence of universal formulas also relativizes and there exists a universal $\Sigma_n(X)$ - (resp. $\Pi_n(X)$ -) formula for every $n \in \mathbb{N}$. That is, a formula $\varphi(x, y, Y)$ such that, for every formula $\psi(y, Y)$ in the corresponding class, there exists some $e \in \mathbb{N}$ such that $\mathbb{I}\Sigma_1^0 \vdash \forall Y \forall y (\varphi(e, y, Y) \leftrightarrow \psi(y, Y))$.

Another consequence of the link between computability-theoretic notions and the arithmetical hierarchy is the following result, due to Kleene. It can be seen as relying on the use property (see Proposition 4.1.1) to show that the truth of a Σ_1^0 -formula depends only on a finite portion of its set parameters. This observation then allows us to rewrite such formulas in a simpler form in which this finite dependence is made explicit.

Proposition 5.1.2 (Kleene normal form for Σ_1^0 -formulas (Kleene [77], see also Simpson [128, Theorem II.2.7])). $\mathbb{I}\Sigma_1^0$ proves that for every Σ_1^0 -formula $\varphi(Y, \vec{y})$, there exists a Σ_0^0 -formula $\psi(x, Y, \vec{y})$ such that

$$\forall Y \forall \vec{y} (\varphi(Y, \vec{y}) \leftrightarrow \exists x \psi(x, Y \upharpoonright x, \vec{y}))$$

and such that, if $\psi(x, Y \upharpoonright x, \vec{y})$ holds for some $x, Y, \vec{y}$, then $\psi(x, Y \upharpoonright x', \vec{y})$ also holds for every $x' \geq x$.

By taking negations of the formulas under consideration, one can also speak of Kleene normal form for Π_1^0 -formulas. More generally, let $\varphi(Y) := \forall x_0 \exists x_1 \dots Q x_{n-1} \theta(x_0, \dots, x_{n-1}, Y)$ a Σ_n^0 -formula (resp. a Π_n^0 -formula), where $Q \in \{\exists, \forall\}$ is determined by the parity of n . Then, the subformula $Q x_{n-1} \theta(x_0, \dots, x_{n-1}, Y)$ can be replaced by its Kleene normal form $Q x_{n-1} \theta'(x_0, \dots, x_{n-1}, Y \upharpoonright x_{n-1})$ yielding a new formula $\psi(Y) := \forall x_0 \exists x_1 \dots Q x_{n-1} \theta'(x_0, \dots, x_{n-1}, Y \upharpoonright x_{n-1})$ equivalent to $\varphi(Y)$ which will be referred to as its Kleene normal form.

In this thesis, we will primarily be concerned with statements of the following form: a Π_2^1 problem P is a formula

$$\forall X \exists Y [\varphi(X) \rightarrow \psi(X, Y)]$$

where both φ and ψ are arithmetic formulas. Such a statement can be interpreted as claiming that for any set X such that $\varphi(X)$ holds (called a P -instance), there is a set Y such that $\psi(X, Y)$ holds (called a P -solution to X).

For example, the Bolzano–Weierstrass theorem is a Π_2^1 problem, where the instances are bounded sequences of reals and the solutions are convergent subsequences.

A major problem in the computable analysis of theorems is to relate the complexity of an instance of a problem to the complexity of its solutions (given a computable instance X , does there exist a computable solution, a $0'$ -computable one, or at least an arithmetically definable one?). These kinds of questions are related to

the study of the proof-theoretic strength of Π_2^1 -statements.

Following Hilbert's program, the infinite sets of integers added to our theories can be seen as *ideal elements*: useful in proving statements about the integers, but whose existence requires justification through consistency proofs. To this end, we study the *first-order part* of second-order theories, namely the set of first-order statements provable within them.

Given a computable second-order theory $\mathcal{T}$, one can consider the computably enumerable set consisting of all first-order statements proved by $\mathcal{T}$. Using Craig's trick [31], it is then possible to obtain a computable (and even primitive recursive) axiomatization of this set. However, this approach is unsatisfactory, as it is a mere tautology that provides no meaningful insight into the first-order consequences of $\mathcal{T}$. A more interesting answer is to search for a "natural" first-order theory $\mathcal{T}^1$ axiomatizing those first-order consequences (or, equivalently said, proving that $\mathcal{T}$ is conservative over $\mathcal{T}^1$ for all first-order sentences).

For example, consider the second-order counterparts of the theories introduced in Chapter 3: $\text{IS}_n^0, \text{I}\Pi_n^0, \text{I}\Delta_n^0, \text{BS}_n^0, \text{B}\Pi_n^0$ For all those theories, the relativization of formulas to include set parameters does not give new proofs of first-order statements, as shown by the result below.

Theorem 5.1.3 (See Clote et al. [27]). *For every $n \geq 1$, the first-order parts of $\text{IS}_n^0, \text{I}\Pi_n^0, \text{I}\Delta_n^0, \text{BS}_n^0, \text{B}\Pi_n^0$... are $\text{IS}_n, \text{I}\Pi_n, \text{I}\Delta_n, \text{BS}_n, \text{B}\Pi_n$... respectively.*

The search for the first-order part of a theory $\mathcal{T}$ becomes more interesting when the theory $\mathcal{T}$ contains some set existence axioms, as those sets can sometimes be used (in an essential way) in the proof of some first-order statement. An important part of this thesis will focus on determining the first-order part of many statements, primarily coming from Ramsey theory.

For the statement $\text{WF}(\alpha)$, we can define its second-order counterpart $\text{WF}^0(\alpha)$ asserting the non-existence of a set encoding an infinite decreasing sequence of ordinals below α . For $\text{WF}^0(\alpha)$ to imply $\text{WF}(\alpha)$, we need to be able to prove the existence of a set encoding every Σ_1 decreasing sequence of ordinals. This will be the case in the finitistically reducible theory RCA_0 , which will be introduced in Section 5.3.

5.2 Henkin models

Just as the independence of the fifth postulate of Euclidean geometry was proven by studying structures different from the Euclidean plane, the same thing can be done with arithmetic, by studying models different from the intended one $\mathbb{N}$.

Definition 5.2.1 (first-order structure). A *first-order structure* is a 6-tuple $(M, 0_M, \text{succ}_M, +_M, \times_M, <_M)$ where M is a set, $0_M \in M$, $\text{succ}_M : M \rightarrow M$ is a unary function, $+_M$ and $\times_M$ are binary functions on M , and $<_M$ is a binary relation on M . When the interpretations of these symbols are clear from context, we simply denote the structure by M .

Given a first-order sentence φ , we say that M is a *model* of φ (denoted $M \models \varphi$) if φ holds true when interpreted in M . For example, $M \models \forall x \forall y (x + y = y + x)$ means that for all $x, y \in M$ we have $x +_M y = y +_M x$.

More generally, if $\mathcal{T}$ is a theory of first-order arithmetic, we say that M is a *model* of $\mathcal{T}$ (denoted $M \models \mathcal{T}$) if $M \models \varphi$ for every axiom $\varphi \in \mathcal{T}$.

All the first-order structures M we will consider will at least be models of PA^- . In that case, the axioms of PA^- guarantee that 0_M is the smallest element for $<_M$, that $\text{succ}_M(0_M)$ is the next smallest, followed by $\text{succ}_M(\text{succ}_M(0_M))$ and so on. Thus, every model of PA^- has an initial segment that is isomorphic to the standard integers, and we can show that $+_M$ and $\times_M$ do coincide with the standard addition and multiplication on this set.

But, as shown by the following example, there can be elements in M that do not belong to that copy of $\mathbb{N}$, and are thus bigger than all the standard integers. Such elements are said to be *non-standard*, and, accordingly, a model with non-standard elements is also said to be *non-standard*.

Example 5.2.2. Consider the order $<$ defined on $\mathbb{Z}[X]$ by $P < Q$ if the leading coefficient of $P - Q$ is positive. Then, $(\mathbb{Z}[X], 0, 1, +, \times, <)$ forms a discretely ordered ring. Let M be its nonnegative part (i.e., the polynomials with positive leading coefficient or 0), then, $(M, 0, succ, +, \times, <)$ forms a model of PA^- where *succ* is the function sending a polynomial P to $P + 1$.

In this model, we have $0 < 1 < 2 < \dots < X - 1 < X < X + 1 < \dots < 2X < \dots < X^2 - 1 < X^2 < \dots < X^3 < \dots$. Thus, the non-constant polynomials form non-standard integers.

The existence of non-standard models can be used to prove non-implication results. For example, in the above model, there is no polynomial $Y \in M$ such that $2Y = X$ or such that $2Y + 1 = X$, thus, the first-order sentence $\varphi := \forall x \exists y, x = 2y \vee x = 2y + 1$ claiming that every integer is odd or even is not provable in PA^- . By contrapositive, they can be used to prove the consistency of a given theory (here, we have a proof that $\text{PA}^- + \neg\varphi$ is a consistent theory).

There are several possible ways to define second-order structures. The most natural approach is given by *full semantics*, where a second-order structure M is just a first-order structure, and set quantifiers are interpreted as ranging over all subsets of M . However, this choice leads to significant drawbacks: most of the major theorems of model theory now fail, such as completeness, compactness, and the Löwenheim-Skolem theorems. Moreover, the truth of a formula in a given model may depend on set-theoretic assumptions about the powerset of M like the continuum hypothesis (which led Quine to refer to second-order logic as “set theory in sheep’s clothing”).

For these reasons, we instead adopt *Henkin semantics* where not every subset of a first-order structure M needs to be in the model. This allows us to obtain counterparts of the main model-theoretic tools from first-order logic.

Definition 5.2.3 (Henkin Structure). A *Henkin structure* $\mathcal{M}$ is a 7-tuple $(M, S, 0_{\mathcal{M}}, succ_{\mathcal{M}}, +_{\mathcal{M}}, \times_{\mathcal{M}}, <_{\mathcal{M}})$ (often shortened to (M, S)) where:

- $(M, 0_{\mathcal{M}}, succ_{\mathcal{M}}, +_{\mathcal{M}}, \times_{\mathcal{M}}, <_{\mathcal{M}})$ is a first-order structure (called the *first-order part* of $\mathcal{M}$).
- $S \subseteq \mathcal{P}(M)$ (called the *second-order part* of $\mathcal{M}$)

For φ a second-order sentence and $\mathcal{T}$ a second-order theory, we define similarly $\mathcal{M} = (M, S) \models \varphi$ and $\mathcal{M} = (M, S) \models \mathcal{T}$ by interpreting second-order quantifiers as ranging over the elements of S , and by interpreting the symbol $\in$ as the membership relation between elements of M and elements of S .

A special kind of model that will be of particular interest in this thesis is the ω -model, which allows for the direct application of computability-theoretic methods in their construction and study.

Definition 5.2.4 (ω -models). An ω -model is a second-order structure of the form $(\mathbb{N}, S, 0, succ, +, \times, <)$ for $S \subseteq \mathcal{P}(\mathbb{N})$ where $(\mathbb{N}, 0, succ, +, \times, <)$ are the standard integers.

Thus, an ω -model is fully specified by its second-order part.

The following two theorems justify the use of model-theoretic arguments in proof theory, by showing a correspondence between syntax (proofs) and semantics (models). If a theory $\mathcal{T}$ proves something, then it holds for every model of $\mathcal{T}$ (soundness), and

reciprocally, if something holds for every model of a theory, then the theory proves it (completeness).

Theorem 5.2.5 (Soundness). *Let $\mathcal{T}$ be a theory of first-order (resp. second-order) arithmetic and φ be a first-order (resp. second-order) statement. If $\mathcal{T} \vdash \varphi$, then for every model M (resp. $\mathcal{M}$) of $\mathcal{T}$, we have $M \models \varphi$ (resp. $\mathcal{M} \models \varphi$). In particular, if $\mathcal{T}$ is not consistent, then $\mathcal{T}$ has no model.*

Theorem 5.2.6 (Completeness (Gödel [50] and Henkin [60])). *Let $\mathcal{T}$ be a theory of first-order (resp. second-order) arithmetic and φ be a first-order (resp. second-order) statement. If $\mathcal{T} \not\vdash \varphi$, then there exists a model $M \models \neg\varphi$ (resp. $\mathcal{M} \models \neg\varphi$). In particular, if $\mathcal{T}$ is consistent, then $\mathcal{T}$ has a model.*

Even in theories $\mathcal{T}$ containing infinitely many axioms, a proof $\mathcal{T} \vdash \varphi$ of a statement uses only finitely many of them. Combined with the completeness theorem, this observation gives us the following compactness result:

Theorem 5.2.7 (Compactness). *Let $\mathcal{T}$ be a theory of first-order (resp. second-order) arithmetic. $\mathcal{T}$ has a model if and only if every finite subset of $\mathcal{T}$ has a model.*

Although the intended model of second-order arithmetic $(\mathbb{N}, \mathcal{P}(\mathbb{N}), 0, \text{succ}, +, \times, <)$ is uncountable, the following theorem shows that it is possible to work with only countable structures. This makes it possible to encode entire models of second-order arithmetic using a single set, hence making it possible to talk about some models of second-order arithmetic in the language of second-order arithmetic itself.

Theorem 5.2.8 ((downward) Löwenheim-Skolem). *Let $\mathcal{T}$ be a theory of first-order (resp. second-order) arithmetic. If $\mathcal{T}$ has a model, then $\mathcal{T}$ has a countable model.*

Those model-theoretic statements, as stated here, are in the language of set theory and not second-order arithmetic. Since we will only work with countable models and countable theories, we only need a version of the completeness theorem stating the existence of a countable model for every countable consistent theory (this version of the completeness theorem also encapsulates one use of the Löwenheim-Skolem theorem). This version is formalizable in the language of second-order arithmetic.

Non-standard models of arithmetic are not finitist objects. However, in the next section we introduce the theory WKL_0 , which captures precisely the strength required to establish the model-theoretic results discussed above. WKL_0 is a finitistically reducible theory, thereby allowing us to use model-theoretic tools for the remainder of this thesis.

5.3 Reverse mathematics

The field of reverse mathematics attempts to find the optimal axioms required to prove various theorems of second-order arithmetic. To ensure that a given axiom was necessary to prove a theorem, it is sufficient to go in the opposite direction and prove the axiom from the theorem. Of course, a theorem is rarely sufficient to prove anything on its own, and we need to work modulo a weak base theory. For example, in the context of set theory, we can prove modulo ZF that the axiom of choice is equivalent to Zorn's lemma, and is thus necessary for its proof (in a theory weaker than ZF the equivalence doesn't necessarily hold, and working modulo ZFC makes this equivalence uninteresting, as all true propositions are equivalent). In reverse mathematics, the choice fell on a theory called RCA_0 encapsulating computable mathematics. Theorem 5.3.5 shows us that this choice is also justified from the point of view of Hilbert's program, as RCA_0 is finitistically reducible.

Definition 5.3.1 (Comprehension scheme). For $\varphi(x)$ a formula (possibly with additional free first- or second-order variables), the comprehension axiom associated to it is the universal closure of the formula:

$$\exists X \forall x (x \in X \leftrightarrow \varphi(x))$$

This defines the corresponding axiom scheme for every class of formulas.

We use the same trick as with ID_n and define the Δ_1^0 -comprehension scheme to include, for every Σ_1^0 - and Π_1^0 -formulas $\varphi(x)$ and $\psi(x)$ (with possibly other free variables), the universal closure of the following axiom:

$$(\forall x \varphi(x) \leftrightarrow \psi(x)) \rightarrow \exists X \forall y (y \in X \leftrightarrow \varphi(y))$$

Definition 5.3.2 (RCA_0). RCA_0 (stands for recursive comprehension axiom) is the theory of second-order arithmetic composed of IS_1^0 together with the Δ_1^0 -comprehension scheme.

From the existence of a universal Σ_1^0 -formula, it follows that the theory RCA_0 is finitely axiomatizable.

The following definition gives a useful characterization of the ω -models of RCA_0 in terms of computably-theoretic notions.

Definition 5.3.3 (Turing ideal). A *Turing ideal* I is a subset of $\mathcal{P}(\mathbb{N})$ which is:

- non-empty.
- stable under Turing reducibility (if $A \in I$ and $B \leq_T A$, then $B \in I$)
- stable under (finite) join (if $A, B \in I$, then $A \oplus B \in I$)

Proposition 5.3.4. *The ω -models of RCA_0 are exactly those whose second-order part is a Turing ideal. Hence, there exists a minimal ω -model of RCA_0 : the model COMP whose second-order part consists of the computable sets.*

The following theorem confirms our intuition about the harmlessness of computable sets. The original proof given by Friedman [48] is not syntactic but model-theoretic, see Sieg [126] for one formalizable in PRA .

Theorem 5.3.5 (Friedman [48]). RCA_0 is Π_1^1 -conservative over IS_1^0 (and hence Π_2^1 -conservative over PRA).

In particular, the first-order part of RCA_0 is IS_1 and, by Ignjatovic [66, Theorem 3.5], RCA_0 has at most a polynomial speedup over IS_1 .

In second-order arithmetic, there are two ways to formalize the notion of a finite set. One approach is to use sets coded by integers, as introduced in Section 3.5. Alternatively, one may consider second-order objects X that are bounded by some first-order object b , in the sense that every $x \in X$ satisfies $x < b$.

In RCA_0 , these two notions coincide. In particular, if $\mathcal{M} = (M, S) \models \text{RCA}_0$, then the sets $X \in S$ that are bounded by some element of M do not depend on the second-order part S of the model. We will refer to such sets as *M -finite sets*.

Note, however, that not every bounded subset of M is M -finite. For instance, if $\mathcal{M} = (M, S) \models \text{RCA}_0$ and M is non-standard, then the set ω of standard integers is bounded in M but not M -finite. We will refer to arbitrary subsets of M (not necessarily elements of S) as being either *M -bounded* or *M -unbounded*. An M -unbounded set that belongs to S will be called *$\mathcal{M}$ -infinite*.

5.3.1 The Big Five

We can then begin comparing the strength of various theorems from across different areas of mathematics by examining the implications between them over RCA_0 , and studying the resulting lattice structure. In doing so, an interesting empirical phenomenon emerges: most theorems fall into one of five large equivalence classes (see Simpson [128] for examples of such equivalences). These classes are known as the “Big Five” (see Montálban [104]).

The first of those classes consists of statements already provable in the base theory RCA_0 . The subsequent four classes can be axiomatized by the following theories, listed in order of increasing strength. Although some of these systems, such as ACA_0 or $\Pi_1^1\text{-CA}_0$ are canonically defined with axiom schemes, each of them can be finitely axiomatized by a theory of the form $\text{RCA}_0 + \varphi$ where φ is any theorem belonging to the class.

WKL₀: The theory WKL_0 is RCA_0 together with weak König’s lemma (WKL). It corresponds to the power needed for compactness arguments. Examples of statements equivalent to WKL_0 modulo RCA_0 are: the Brouwer fixed point theorem, the compactness of closed intervals of $\mathbb{R}$ and many others. More interestingly for our purposes, the countable version of the completeness theorem is also equivalent to WKL_0 modulo RCA_0 (the soundness theorem is already provable in RCA_0).

ACA₀: The theory ACA_0 is RCA_0 together with the comprehension scheme for arithmetical formulas. Examples of statements equivalent to ACA_0 modulo RCA_0 are the Bolzano–Weierstrass theorem and the Arzelà–Ascoli theorem.

ATR₀: The theory ATR_0 is RCA_0 together with the axiom of arithmetic transfinite recursion.

$\Pi_1^1\text{-CA}_0$: The theory $\Pi_1^1\text{-CA}_0$ is RCA_0 together with the comprehension scheme for Π_1^1 -formulas.

We will not use the theories ATR_0 and $\Pi_1^1\text{-CA}_0$ in this thesis and mostly consider statements provable in ACA_0 .

We also have a characterization of the ω -models of WKL_0 and ACA_0 in terms of computability-theoretic notions:

Definition 5.3.6 (Scott ideal). A *Scott ideal* I is a Turing ideal that additionally satisfies the condition that for every $X \in I$, there is a set $Y \in I$ that is of PA degree relative to X .

Proposition 5.3.7. *The ω -models of WKL_0 are exactly those whose second-order part is a Scott ideal.*

By iterating and relativizing Theorem 4.2.5, we can construct a Scott ideal that does not contain any given non-computable set. In particular, there is no minimal Scott ideal.

Proposition 5.3.8. *The ω -models of ACA_0 are exactly those whose second-order part is a Turing ideal that is closed under Turing jump, such ideals are called jump ideals. Hence, there exists a minimal ω -model of ACA_0 : the model ARITH whose second-order part consists of the arithmetical sets.*

In the same way that the Von Neumann–Bernays–Gödel set theory (NBG) is conservative over ZF, ACA_0 is conservative over Peano arithmetic.

Theorem 5.3.9 (See e.g., [128, Corollary IX.1.6]). *The first-order part of ACA_0 is PA.*

And, as in the case of NBG and ZF, Solovay, Pudlák, and Friedman (see e.g., [117]) showed that ACA_0 has a non-elementary speedup over PA.

For WKL_0 , the following proposition, combined with Theorem 5.3.5 yields the Π_2 -conservation of WKL_0 over PRA . The original proof from Harrington was not syntactic as it uses model-theoretic arguments. A direct syntactic proof (formalizable in PRA) of the Π_2 -conservation of WKL_0 over PRA was later given by Sieg [126], and more recently, a syntactical proof of Harrington's result by Ferreira and Ferreira [45].

Theorem 5.3.10 (Harrington (see Simpson [128, Corollary IX.2.6])). *WKL_0 is Π_1^1 -conservative over RCA_0 .*

By Avigad [4, Theorem 10.1], every proof of a Π_1^1 -statement in WKL_0 can be transformed into a proof of the same statement in RCA_0 with at most a polynomial increase in size.

Thus, WKL_0 is finitistically reducible. In view of this result, and since WKL_0 is sufficiently strong to formalize most model-theoretic constructions, the use of model-theoretic arguments to derive conservation and non-implication results is finitistically justified. And we will primarily rely on this kind of argument throughout the remainder of this thesis instead of resorting to syntactical proofs.

One important point to keep in mind when working with conservation results is that, even after the conservation of a statement φ over a theory $\mathcal{T}$ has been established for a class Γ of formulas, there is no guarantee that this will remain true when replacing $\mathcal{T}$ with a stronger theory $\mathcal{T}'$. Indeed, a procedure that transforms a proof in $\mathcal{T} + \varphi$ into a proof in $\mathcal{T}$, does not necessarily yield a corresponding transformation from proofs in $\mathcal{T}' + \varphi$ to proofs in $\mathcal{T}'$. Fortunately for us, this seems to be the case for most “natural” examples of theories and statements, and the first-order parts of theories such as $\text{RCA}_0 + \text{IS}_n^0$, $\text{RCA}_0 + \text{BS}_n^0$, $\text{WKL}_0 + \text{IS}_n^0$ and $\text{WKL}_0 + \text{BS}_n^0$ are respectively IS_n , BS_n , IS_n and BS_n , as shown by Clote et al. [27].

For $\text{WF}^0(\alpha)$, we can see that it implies $\text{WF}(\alpha)$ modulo RCA_0 . Indeed, as discussed in Chapter 3, any Σ_1 -definable decreasing sequence of ordinals below α can also be defined by a Π_1 -formula and thus forms a computable set. By similar arguments, we get that the first-order part of $\text{RCA}_0 + \text{WF}^0(\alpha)$ is $\text{IS}_1 + \text{WF}(\alpha)$.

5.4 Semantic proofs

There are several ways to compare the strength of two second-order statements P and Q :

First, one may compare their first-order strength by identifying first-order sentences that are provable in $\text{RCA}_0 + P$ but not in $\text{RCA}_0 + Q$. Similarly, their consistency strength can be compared by determining whether $\text{RCA}_0 \vdash \text{Con}(\text{RCA}_0 + P) \rightarrow \text{Con}(\text{RCA}_0 + Q)$.

One can also compare their second-order strength by checking whether every ω -model of $\text{RCA}_0 + P$ is again a model of Q .

Finally, one may ask whether $\text{RCA}_0 \vdash P \rightarrow Q$, which reflects both the first-order and second-order strength of P relative to Q .

Two theories may share the same first-order consequences, while differing in second-order strength. For instance, WKL_0 and RCA_0 prove the same arithmetical statements, yet COMP , the ω -model whose second-order part consists of the computable sets, is not a model of WKL_0 .

Conversely, two theories may have the same ω -models but differ in their first-order consequences. For example, $\text{RCA}_0 + \text{IS}_n^0$ and $\text{RCA}_0 + \text{IS}_m^0$ for $n \neq m > 0$ have distinct first-order parts (IS_n and IS_m), but both hold in every ω -model of RCA_0 .

If $\text{RCA}_0 \vdash P \rightarrow Q$, then, in particular, the first-order part of $\text{RCA}_0 + Q$ is included in that of $\text{RCA}_0 + P$, and every ω -model of $\text{RCA}_0 + P$ is an ω -model of Q (since this holds for every model).

Semantic proofs showing that a theory $\mathcal{T}_1$ is Γ -conservative over a theory $\mathcal{T}_0$ for Γ a class of formulas typically proceed as follows. Arguing by contrapositive, suppose that $\mathcal{T}_0 \not\models \varphi$ for some $\varphi \in \Gamma$. Then, by Gödel's completeness theorem, there exists a model $\mathcal{M}$ of $\mathcal{T}_0 + \neg\varphi$. From this model, we construct a new model $\mathcal{M}'$ that is sufficiently similar to $\mathcal{M}$ to still satisfy $\neg\varphi$, while also ensuring that $\mathcal{M}' \models \mathcal{T}_1$, thereby showing that $\mathcal{T}_1 \not\models \varphi$.

As mentioned earlier, our main focus will be on comparing the strength of Π_2^1 -statements, particularly those arising from Ramsey theory. The remaining three parts of this thesis are organized according to the different kinds of model-theoretic arguments used to build models.

In Part II we will only consider ω -models. In Chapter 7, we compare Ramsey's theorem for pairs and two colors (RT_2^2) to several of its restrictions by proving that they hold in the same ω -models. In Chapter 8, we prove that a statement called the rainbow Ramsey theorem for exactly ω -large sets (RRT^ω) does not imply ACA_0 over ω -models, by constructing an ω -model of RRT^ω that does not contain $\emptyset'$ and is therefore not a model of ACA_0 . This is achieved by proving that RRT^ω admits a property called cone avoidance, as is the case for WKL_0 by Theorem 4.2.5.

In Part III, we consider models $\mathcal{M} = (M, S)$ of RCA_0 and see ways to expand them into new models $\mathcal{M}' = (M, S')$ by adding second-order elements. These models may have a non-standard first-order part M and we will provide a general overview of them in Chapter 9. Since $\mathcal{M}$ and $\mathcal{M}'$ share the same first-order part M , any Σ_1^1 -formula that holds in $\mathcal{M}$ also holds in $\mathcal{M}'$. This fact will be used in Chapter 10 to prove the Π_1^1 -conservativity of RT_2^2 over the theory $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\varepsilon_0)$. In Chapter 11 we construct a model of the ascending/descending sequence principle (ADS) that is not a model of the bounded Ramsey theorem (BRT_2^2), thereby showing that $\text{RCA}_0 \not\models \text{ADS} \rightarrow \text{BRT}_2^2$. Such a construction could not have been carried out in an ω -model as every such model satisfies BRT_2^2 .

Finally, in Part IV, we consider models $\mathcal{M} = (M, S)$ and see ways to define initial segments $I \subseteq M$ called cuts, in order to build new models whose first-order part is I . In Chapter 12 we introduce the general framework for constructing such I so that any Π_4 -formula true in M remains true in I . In Chapter 13 we prove some additional combinatorial properties related to this framework. Then, we apply it in Chapter 14 and Chapter 15 to prove the $\forall\Pi_4^0$ -conservativity of RT_2^2 and of the ordered variable word theorem (OVW) over $\text{RCA}_0 + \text{B}\Sigma_2^0$. Here, a $\forall\Pi_4^0$ -formula is any formula of the form $\forall X \varphi(X)$ where φ is Π_4^0 .

The next chapter will introduce several statements from Ramsey theory that will be used all along this thesis and compare their first-order, computability-theoretic and reverse-mathematical strength.

Chapter 6

Ramsey's theorem

Introduced by Frank Ramsey in 1930 [119], Ramsey's theorem was originally conceived as a tool in formal logic for deciding the truth of certain formulas. It has since evolved well beyond its original purpose, giving rise to Ramsey theory, which studies related results. Broadly speaking, Ramsey theory investigates the extent to which particular substructures must necessarily appear within sufficiently large structures.

Definition 6.0.1. For $A \subseteq \mathbb{N}$ a set, and $n \in \mathbb{N}$ an integer, we write $[A]^n$ for the set of subsets of A of cardinality n . For $b \in \mathbb{N}$, we will also write $[b]^n$ for $[\{0, \dots, b-1\}]^n$.

For a function f with domain $[A]^n$, we will write $f(x_0, \dots, x_{n-1})$ as shorthand for $f(\{x_0, \dots, x_{n-1}\})$ whenever $x_0 < \dots < x_{n-1}$.

Statement 6.0.2 ((infinite) Ramsey's theorem). For $n, k \in \mathbb{N}$, RT_k^n is the statement that: "For every coloring $f : [\mathbb{N}]^n \rightarrow k$, there exists some infinite subset $X \subseteq \mathbb{N}$ such that f is constant on $[X]^n$." Equivalently, we say that the set X is f -homogeneous.

RT^n is the statement $\forall k \text{RT}_k^n$, and RT is the statement $\forall n \forall k \text{RT}_k^n$.

RT^1 corresponds to the infinite pigeonhole principle, which states that, for any partition of $\mathbb{N}$ into finitely many parts, one of the parts must be infinite.

Ramsey's original paper also features a finitary version of this theorem, focusing on colorings of $[A]^n$ for a finite set A .

Theorem 6.0.3 ((finite) Ramsey's theorem). *For every $n, k, d \in \mathbb{N}$, there exists a bound $b := b(n, k, d) \in \mathbb{N}$ such that, for every coloring $f : [b]^n \rightarrow k$, there exists some f -homogeneous subset $B \subseteq \{0, \dots, b-1\}$ with $\text{card } B \geq d$.*

The finite version is provable in RCA_0 by a proof giving explicit bounds $b(n, k, d)$ from the parameters n, k and d . For instance, in the case of colorings of pairs, let $R_k(d)$ be the optimal bound $b(2, k, d)$. The standard upper bound for the finite Ramsey theorem for pairs and two colors, obtained by Erdős and Szekeres in [42], states that $R_2(d) \leq \binom{2d-1}{d-1} \leq 4^d$. This upper bound generalizes to $R_k(d) \leq k^{kd}$ (see Graham, Rothschild, and Spencer [53, Section 1.1]).

The following proof shows that the finite Ramsey theorem follows from the infinite by a compactness argument (i.e., is provable in $\text{WKL}_0 + \text{RT}$). It is a far less optimal proof than the one in RCA_0 , but is a good example of the kind of compactness argument that will be used all along this thesis to obtain finite bounds from results holding for every infinite set. It is also a good example of a proof of a finitary result using infinitary means, and the last part of this thesis, Part IV, will study in more detail the relation between finitary and infinitary versions of Ramsey's theorem. In

fact, we will see in Proposition 12.0.3 how a slight modification of that proof yields, in $\text{WKL}_0 + \text{RT}$, a proof of a variant of the finite Ramsey theorem for pairs that is not provable in ACA_0 (and hence not in PA either).

Proof. Fix $n, k \in \mathbb{N}$ and consider the set T_k^n of all colorings $f : [a]^n \rightarrow k$ for some $a \in \mathbb{N}$. This set naturally forms a tree under inclusion: a coloring $f : [a]^n \rightarrow k$ is a predecessor of $g : [c]^n \rightarrow k$ if $a \leq c$ and $g \upharpoonright [a]^n = f$. The tree is finitely branching, since for each a there are only finitely many colorings $g : [a+1]^n \rightarrow k$, and hence only finitely many immediate extensions of a given coloring $f : [a]^n \rightarrow k$. The set of infinite paths $[T_k^n]$ corresponds exactly to the set of all colorings $f : [\mathbb{N}]^n \rightarrow k$.

Now fix $d \in \mathbb{N}$ and consider the subtree $S_k^n(d) \subseteq T_k^n$ consisting of those colorings $f : [a]^n \rightarrow k$ for which no f -homogeneous subset of cardinality d can be found. This tree has no infinite path: indeed, any infinite path through $S_k^n(d)$ would yield a coloring $f : [\mathbb{N}]^n \rightarrow k$ that admits no f -homogeneous subset of cardinality d , contradicting RT_k^n which proves the existence of an infinite f -homogeneous set.

Since the statement WKL is stated in terms of binary trees, we cannot apply its contrapositive to $S_k^n(d)$ directly, as its branching may exceed 2. However, since we can computably find the code of the finite set of immediate extensions of any given node in $S_k^n(d)$, one can computably replace each branching of $S_k^n(d)$ by a bounded number of successive binary branchings, yielding a binary tree with no infinite path either, and which is finite iff $S_k^n(d)$ is. Therefore, by the contrapositive of WKL , the tree $S_k^n(d)$ is finite. In particular, there exists $b \in \mathbb{N}$ such that $S_k^n(d)$ has no node of height $b \in \mathbb{N}$, that is, for every coloring $f : [b]^n \rightarrow k$, there exists an f -homogeneous subset of cardinality d . $\square$

A lot is already known concerning the reverse mathematical strength of the RT_k^n (see Hirschfeldt [61] for a gentle introduction to the reverse mathematics of Ramsey-like theorems).

For every $n, k \in \mathbb{N}$, the statements RT_0^n and RT_k^0 are vacuously true and RT_1^n is trivial, so we only consider statements RT_k^n for $k \geq 2$ and $n \geq 1$.

The infinite pigeonhole principle is *computably true*, in the sense that for every instance $f : [\mathbb{N}]^1 \rightarrow k$ of RT_k^1 (i.e., every partition of $\mathbb{N}$ into k sets) there exists a $\Delta_1(f)$ infinite f -homogeneous set. In particular, RT_k^1 is provable in RCA_0 for every fixed $k \in \mathbb{N}$.

Proposition 6.0.4. *For every $k \in \mathbb{N}$, $\text{RCA}_0 \vdash \text{RT}_k^1$.*

Nevertheless, proving that $\forall k \text{RT}_k^1$ holds requires more inductive power than is available in RCA_0 .

Theorem 6.0.5 (Hirst [64]). $\text{RCA}_0 \vdash \text{RT}^1 \leftrightarrow \text{B}\Sigma_2^0$, hence $\text{RCA}_0 \not\vdash \text{RT}^1$.

The RT_k^n form a hierarchy of statements with the following implications between them:

Proposition 6.0.6.

- $\text{RCA}_0 \vdash \forall n \forall k \geq 2 \forall m < n \text{RT}_k^n \rightarrow \text{RT}_k^m$
- $\text{RCA}_0 \vdash \forall n \forall k \forall \ell \leq k \text{RT}_k^n \rightarrow \text{RT}_\ell^n$
- $\text{RCA}_0 \vdash \forall n \forall k \geq 2 \text{RT}_k^n \rightarrow \text{RT}_{k+1}^n$

But the hierarchy collapses after $n = 3$ and everything is then equivalent to ACA_0 . This is due to the existence of a computable instance of RT_2^3 , every solution of which computes the halting set $\emptyset'$.

Theorem 6.0.7 (Jockusch [70] and Simpson [128, Theorem III.7.6]). *For $n \geq 3$ and $k \geq 2$, $\text{RCA}_0 \vdash \text{RT}_k^n \leftrightarrow \text{ACA}_0 \leftrightarrow \text{RT}^n$.*

Thus, there remains the question of the reverse mathematical strength of Ramsey's theorem for pairs. As $\text{RCA}_0 \vdash \text{RT}_2^2 \rightarrow \text{RT}_k^2$ for every $k \in \mathbb{N}$, we will only focus on Ramsey's theorem for pairs and two colors. RT_2^2 has been shown to escape the structural phenomenon of the Big Five by being strictly between RCA_0 and ACA_0 but incomparable with WKL_0 .

Theorem 6.0.8 (Jockusch [70] and Liu [95]).

- $\text{WKL}_0 \not\vdash \text{RT}_2^2$
- $\text{RCA}_0 + \text{RT}_2^2 \not\vdash \text{WKL}$

The unprovability of RT_2^2 in WKL_0 follows from the following proposition.

Proposition 6.0.9 (Jockusch [70]). *There exists a computable instance of RT_2^2 for which no Σ_2 infinite set is homogeneous.*

Indeed, from the low basis theorem (Theorem 4.2.7), every computable instance of WKL admits a solution that is low (and in particular Δ_2). This result can then be used (in a relativized form) to build an ω -model of WKL_0 in which every set is Δ_2 (see Proposition 8.1.5 for a similar construction but for the cone avoidance basis theorem). By Proposition 6.0.9, this model is not a model of RT_2^2 (as it contains an instance of RT_2^2 but no solution for it).

Nevertheless, Jockusch [70] showed that every computable instance of RT_2^2 admits a Π_2 solution. Moreover, even though no analogue of the low basis theorem (where $\emptyset'$ can compute the jump of a solution) can be proven for RT_2^2 , the following result shows that solutions of weak strength still exist (where a PA degree relative to $\emptyset'$ can compute the jump of a solution).

Theorem 6.0.10 (Cholak, Jockusch, and Slaman [19]). *Let $\mathbf{d}$ be of PA degree over $\emptyset'$ (i.e., $\mathbf{d}$ computes a path in any $\emptyset'$ -computable binary tree). Then every computable instance of RT_2^2 admits an infinite homogeneous set H such that $\mathbf{d} \geq_T H'$.*

In particular, by the low basis theorem (Theorem 4.2.7) relativized to $\emptyset'$, there exists some $\mathbf{d}$ of PA degree over $\emptyset'$ that is low over $\emptyset'$ (i.e., such that $(\mathbf{d} \oplus \emptyset')' \leq_T \emptyset''$). Thus, we obtain the following result:

Theorem 6.0.11 (Cholak, Jockusch, and Slaman [19]). *Every computable instance of RT_2^2 admits an infinite homogeneous set that is low₂, i.e., a set H such that $H'' \leq_T \emptyset''$.*

Ramsey's theorem for pairs and two colors is often formulated in graph-theoretic terms: for every 2-coloring of the edges of the infinite complete graph $(\mathbb{N}, [\mathbb{N}]^2)$, there exists an infinite monochromatic clique (where a *clique* is a set of vertices with edges between each pair).

6.1 First-order part

Concerning the first-order part of RT_2^2 , since $\text{RCA}_0 \vdash \text{RT}_2^2 \rightarrow \text{RT}_1^1$, we have the following lower bound:

Proposition 6.1.1 (Hirst [64]). *RT_2^2 implies $\text{B}\Sigma_2^0$ over RCA_0 .*

Whether the first-order part of RT_2^2 is $\text{B}\Sigma_2$ or not is a major open problem in reverse mathematics and this thesis makes several contributions to the subject.

Question 6.1.2. Is RT_2^2 Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$?

By the following result, the first-order consequences of RT_2^2 are included in those of $\text{I}\Sigma_2$.

Theorem 6.1.3 (Cholak, Jockusch and Slaman [19]). RT_2^2 is Π_1^1 -conservative over $\text{RCA}_0 + \text{IS}_2^0$.

But, by Chong, Slaman, and Yang [23], this upper bound is strict in that RT_2^2 does not imply IS_2^0 over RCA_0 .

Patey and Yokoyama [115] extended this result by proving RT_2^2 to be finitistically reducible:

Theorem 6.1.4 (Patey and Yokoyama [115]). $\text{WKL}_0 + \text{RT}_2^2$ is $\forall\Pi_3^0$ -conservative over RCA_0 (where a $\forall\Pi_3^0$ -formula is any formula of the form $\forall X\varphi(X)$ where φ is Π_3^0).

Later, Kołodziejczyk et al. [78] were able to prove that $\text{WKL}_0 + \text{RT}_2^2$ offers at most a polynomial speedup over RCA_0 for $\forall\Pi_3^0$ -formulas by adapting the general scheme of forcing interpretation as introduced by Avigad [4].

Even if, for each $k \in \mathbb{N}$, we have $\text{RCA}_0 \vdash \text{RT}_2^2 \leftrightarrow \text{RT}_k^2$, there is not enough induction available to derive $\forall k \text{RT}_k^2$ from $\text{RCA}_0 + \text{RT}_2^2$. In fact, the exact first-order strength of RT^2 is known to be stronger than that of RT_2^2 .

Theorem 6.1.5 (Slaman and Yokoyama [133]). The first-order part of $\text{RCA}_0 + \text{RT}^2$ is BS_3 .

6.2 Ramsey-like theorems

A common tool in the study of Ramsey's theorem for pairs and two colors was to decompose it into several weaker statements, each focusing on one part of the combinatorics. One decomposition that will be used in all parts of this thesis is the decomposition of RT_2^2 into the Erdős-Moser theorem about infinite tournaments and the ascending/descending sequence principle about linear orders. We will formulate both these theorems in terms of coloring so that they form special cases of RT_2^2 for so-called transitive colorings.

Definition 6.2.1. Let $A \subseteq \mathbb{N}$ and let $f : [A]^2 \rightarrow 2$ be a coloring. Then, f is said to be *transitive* if for every $x < y < z \in A$ and every $i < 2$, if $f(x, y) = f(y, z) = i$, then $f(x, z) = i$.

A subset $H \subseteq A$ is said to be *f-transitive* if $f \upharpoonright [H]^2$ is transitive.

Definition 6.2.2. A *tournament* on a domain $D \subseteq \mathbb{N}$ is an irreflexive binary relation $R \subseteq D^2$ such that for every $a, b \in D$ with $a \neq b$, exactly one of $R(a, b)$ and $R(b, a)$ holds. A tournament R is *transitive* if for every $a, b, c \in D$, if $R(a, b)$ and $R(b, c)$ then $R(a, c)$.

A subtournament of $R \subseteq \mathbb{N}^2$ is fully characterized by its domain and the parent tournament R . We therefore identify subtournaments with infinite sets $H \subseteq \mathbb{N}$. The following statement is known as the Erdős-Moser theorem, and is an infinitary version of a theorem of Erdős and Moser [41].

Statement 6.2.3 (Erdős-Moser theorem). **EM** is the statement “Every infinite tournament admits an infinite transitive subtournament”.

There exists a one-to-one correspondence between tournaments R and colorings $f : [\mathbb{N}]^2 \rightarrow 2$, by letting $f(\{x, y\}) = 1$ if $(R(x, y) \leftrightarrow x < y)$. **EM** can be restated as “for every coloring $f : [\mathbb{N}]^2 \rightarrow 2$, there exists an infinite *f-transitive* subset $H \subseteq \mathbb{N}$ ”.

Statement 6.2.4 (Ascending/descending sequence principle). **ADS** is the statement “Every infinite linear order admits an infinite ascending or descending subsequence”.

Here again, there is a one-to-one correspondence between linear orders $\mathcal{L} = (\mathbb{N}, <_{\mathcal{L}})$ and transitive colorings $f : [\mathbb{N}]^2 \rightarrow 2$, by letting $f(\{x, y\}) = 1$ if $(x < y \leftrightarrow x <_{\mathcal{L}} y)$.

Then ADS can be restated as “for every transitive coloring $f : [\mathbb{N}]^2 \rightarrow 2$, there exists an infinite f -homogeneous set”. Using the reformulations of EM and ADS, the following proposition is immediate:

Proposition 6.2.5 (Bovykin and Weiermann [11]). $\text{RCA}_0 \vdash \text{RT}_2^2 \leftrightarrow \text{EM} \wedge \text{ADS}$.

Proof. EM and ADS are both particular cases of RT_2^2 (as every f -homogeneous set is f -transitive), hence $\text{RCA}_0 \vdash \text{RT}_2^2 \rightarrow \text{EM} \wedge \text{ADS}$.

On the other hand, given a coloring $f : [\mathbb{N}]^2 \rightarrow 2$, first apply EM to obtain an infinite f -transitive subset $X = \{x_0 < x_1 < \dots\}$. Let $g : [\mathbb{N}]^2 \rightarrow 2$ be the transitive coloring defined by $g(a, b) = f(x_a, x_b)$. By ADS, there exists an infinite g -homogeneous subset $H \subseteq \mathbb{N}$. The set $\{x_a : a \in H\}$ is infinite and f -homogeneous. $\square$

From a combinatorial viewpoint, the decomposition of RT_2^2 into EM and ADS divides the difficulty of Ramsey’s theorem for pairs as follows: the combinatorics of EM are very close to those of RT_2^2 , but without the disjunctive nature of RT_2^2 (there is no need to pick a color at one point). On the other hand, ADS has very simple combinatorics, but is a disjunctive statement.

From a reverse mathematical viewpoint, the decomposition is non-trivial, in that EM and ADS are both strictly weaker than RT_2^2 (see Lerman, Solomon, and Towsner [94]).

From a proof-theoretic viewpoint, both statements imply $\text{B}\Sigma_2^0$ over RCA_0 (see Hirschfeldt and Shore [63, Proposition 4.5] and Kreuzer [83, Proposition 16]), and $\text{RCA}_0 + \text{ADS}$ is known to be Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$ (see Chong, Slaman, and Yang [22]). By an amalgamation theorem of Yokoyama [142], it follows that RT_2^2 is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$ iff so is EM.

The second (and historically the first) decomposition of RT_2^2 we will consider is the decomposition by Cholak, Jockusch, and Slaman [19] in their seminal article on Ramsey’s theorem for pairs. It decomposes RT_2^2 in terms of the cohesive set principle (COH) and the stable Ramsey theorem (SRT_2^2).

Definition 6.2.6 (stable coloring). A coloring $f : [\mathbb{N}]^2 \rightarrow k$ is *stable* if for every $x \in \mathbb{N}$, there exists some color $i < k$ and some rank $b \in \mathbb{N}$ such that, for every $y \geq b$, $f(x, y) = i$ (i.e., $\lim_y f(x, y)$ exists for every $x \in \mathbb{N}$).

Statement 6.2.7 (Stable Ramsey’s theorem for pairs). SRT_k^2 denotes the restriction of RT_k^2 to stable colorings.

By Shoenfield’s limit lemma (Proposition 4.1.7), there is a correspondence between stable colorings and Δ_2^0 sets. For $f : [\mathbb{N}]^2 \rightarrow 2$ a stable coloring, we can consider the partition $\mathbb{N} = A_0 \sqcup A_1$ where $A_i = \{x : \lim_y f(x, y) = i\}$ for $i < 2$. Both A_i are $\Delta_2(f)$ sets, and, if for some $i < 2$ we have some infinite subset $C \subseteq A_i$, then we can compute from C and f an infinite f -homogeneous set. Indeed, proceeding inductively, suppose that for some $n \in \mathbb{N}$, some $x_0 < \dots < x_{n-1} \in C$ have been chosen so that $f(x_j, x_k) = i$ for all $j < k < n$. Then, as $C \subseteq A_i$, we have $\lim_y f(x_j, y) = i$ for every $j < n$ and we will eventually find some $x_n \in C$ with $x_n > x_{n-1}$ and such that $f(x_j, x_n) = i$ for every $j < n$.

Continuing this way yields an infinite f -homogeneous set. Notice that we don’t need to compute the stage at which the limit stabilizes (which is not computable), but need only to know that such a stage exists.

This proof can then be formalized in RCA_0 to establish the equivalence between SRT_2^2 and the infinite pigeonhole principle for Δ_2^0 sets (D_2^2): for every set X and every partition $A_0 \sqcup A_1 = \mathbb{N}$ with each A_i a $\Delta_2(X)$ set, there exist $i < 2$ and an infinite set $H \subseteq A_i$.

The original proof of this equivalence, due to Cholak et al. [19, Lemma 7.10], contained a gap: proving the existence of a uniform bound after which all the $x_0 <$

$\dots < x_{n-1}$ have reached their limit behavior requires $\text{B}\Sigma_2^0$. This gap was later closed by Chong et al. [20] who showed that D_2^2 implies $\text{B}\Sigma_2^0$.

D_2^2 is often combinatorially simpler to use than SRT_2^2 .

Definition 6.2.8. Given an infinite sequence of sets $\vec{R} = R_0, R_1, \dots$, an infinite set C is $\vec{R}$ -cohesive if for every $s \in \mathbb{N}$, $C \subseteq^* R_s$ or $C \subseteq^* \mathbb{N} \setminus R_s$, where $\subseteq^*$ denotes inclusion up to finite changes.

Statement 6.2.9 (Cohesive set principle). COH is the statement “Every infinite sequence of subsets of $\mathbb{N}$ has a cohesive set”.

The following equivalence is essentially due to Cholak, Jockusch, and Slaman [19], with the proof $\text{RCA}_0 \vdash \text{RT}_2^2 \rightarrow \text{COH}$ fixed by Mileti [99, Appendix A].

Proposition 6.2.10. $\text{RCA}_0 \vdash \text{RT}_2^2 \leftrightarrow \text{SRT}_2^2 \wedge \text{COH}$.

Proof. We omit the proof of the forward direction. Let us prove that $\text{RCA}_0 \vdash \text{SRT}_2^2 \wedge \text{COH} \rightarrow \text{RT}_2^2$. Given a coloring $f : [\mathbb{N}]^2 \rightarrow 2$, define the sequence of sets $\vec{R} = R_0, R_1, \dots$ by $R_x = \{y : f(x, y) = 1\}$. Given an infinite $\vec{R}$ -cohesive set $C = \{x_0 < x_1 < \dots\}$, let $g : [\mathbb{N}]^2 \rightarrow 2$ be defined by $g(\{a, b\}) = f(\{x_a, x_b\})$. By cohesiveness of C , the coloring g is stable, so by SRT_2^2 , there is an infinite g -homogeneous set H . The set $\{x_a : a \in H\}$ is infinite and f -homogeneous. $\square$

From a reverse mathematical viewpoint, the decomposition of RT_2^2 into SRT_2^2 and COH is also non-trivial (see Hirschfeldt et al. [62] and Monin and Patey [103]).

The computability-theoretic content of the cohesive set principle is well-understood: it admits a maximally complex instance, which is the sequence of all primitive recursive sets. A cohesive set for this sequence is called *p-cohesive*. For every uniformly computable sequence of sets $\vec{R}$, every *p*-cohesive set computes an $\vec{R}$ -cohesive set. Moreover, by Jockusch and Stephan [69], a set G computes a *p*-cohesive set iff G' is PA over $0'$.

From a proof-theoretic viewpoint, by Hirschfeldt and Shore [63], COH is $r\text{-}\Pi_2^1$ -conservative over RCA_0 , where a $r\text{-}\Pi_2^1$ -formula (for *restricted* Π_2^1) is a formula of the form $\forall X[\eta(X) \rightarrow \exists Y\rho(X, Y)]$ where η is arithmetic and ρ is Σ_3^0 , in particular COH is Π_1^1 -conservative over RCA_0 . SRT_2^2 implies $\text{B}\Sigma_2^0$ over RCA_0 and, again, by the amalgamation theorem, RT_2^2 is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$ iff so is SRT_2^2 .

The idea of using cohesiveness to transform a coloring into a stable one can be applied to EM and ADS. We call SEM and SADS the restrictions of EM and ADS to stable tournaments and linear orders, respectively. In a stable linear order, every element is either smaller than everyone except a finite number of elements, or bigger than everyone except a finite number of elements. Therefore, such an order can only be of three types: $\omega + \omega^*$, $\omega + k^*$ and $k + \omega^*$, for some $k \in \mathbb{N}$ (where α^* is the opposite order to α). The two latter cases are computably true (provable in RCA_0), so the only interesting case is when it is of order type $\omega + \omega^*$. This is why SADS is often defined as the statement “Every linear order of type $\omega + \omega^*$ admits an infinite ascending or descending sequence”.

Combining the two kinds of decomposition, we shall use the following equivalence:

Proposition 6.2.11. $\text{RCA}_0 \vdash \text{RT}_2^2 \leftrightarrow \text{SEM} \wedge \text{SADS} \wedge \text{COH}$.

We will often prove conservation, implication, and non-implication results about COH , SEM and SADS independently, and use the previous decomposition to obtain our main theorems.

The fact that RT_2^2 escapes the Big Five phenomenon has led to the study of many other Ramsey-like theorems, with the goal of finding other statements of unusual strength. In this context, a phenomenon opposite to that of the Big Five emerges: instead of a small number of equivalence classes over RCA_0 corresponding to well-studied subsystems of second-order arithmetic, one finds an entire zoo of equivalence classes¹. In this thesis, we shall study some of these Ramsey-like theorems of peculiar strength.

¹<https://www.computability.org/zoo-viewer/>

Part II

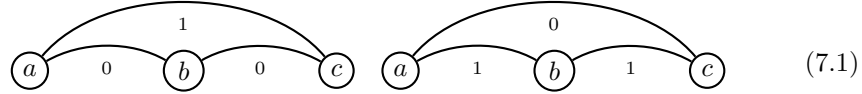
Computational content of theorems

Chapter 7

Ramsey-like theorems for separable permutations

7.1 Introduction

In this chapter, we pursue the computability-theoretic study of weakenings of Ramsey's theorem for pairs by considering infinite cliques avoiding some finite coloring patterns instead of monochromatic cliques. For instance, the weakening of RT_2^2 to cliques avoiding the following two patterns yields the Erdős-Moser theorem (EM) as they correspond to the two ways a set can be non-transitive.



We also consider dual weakenings of Ramsey's theorem for pairs by restricting colorings to those avoiding some pattern. For instance, RT_2^2 for colorings avoiding (7.1) yields the ascending/descending sequence principle (ADS).

We give a particular focus on transitive forbidden patterns, that is, patterns which do not include (7.1). These patterns are in one-to-one correspondence with permutations. Among these, the class of *separable permutations* introduced by Bose, Buss, and Lubiw [10] will play an important role. Our main result is a characterization of the patterns whose avoidance implies RT_2^2 over ω -models, as the separable permutations (Theorem 7.1.6). For this, we use an elaborate computability-theoretic analysis of a probabilistic construction.

7.1.1 RT_2^2 -patterns

An RT_2^2 -*pattern* or simply *pattern* is a function $p : [\ell]^2 \rightarrow 2$. We then write $|p|$ for its *size* ℓ . Fix a function $f : [A]^2 \rightarrow 2$ for some $A \subseteq \mathbb{N}$. We say that a finite set $F = \{x_0 < \dots < x_{\ell-1}\} \subseteq A$ *f-realizes* p if for every $i < j$, $f(x_i, x_j) = p(i, j)$. We say that $H \subseteq A$ *f-avoids* the pattern p if no subset of H *f-realizes* p . When $H = A$, we simply say that the coloring f *avoids* p .

A pattern $q : [\ell']^2 \rightarrow 2$ is a *sub-pattern* of p if some subset of $\{0, \dots, \ell-1\}$ p -realizes q . In that case, every coloring avoiding q will also avoid p .

We are primarily interested in the following two families of statements:

Statement 7.1.1 ($\text{RT}_2^2(p)$). Given a pattern p , $\text{RT}_2^2(p)$ is the statement “For every coloring $f : [\mathbb{N}]^2 \rightarrow 2$, there exists an infinite set f -avoiding p .”

Statement 7.1.2 ($\text{coRT}_2^2(p)$). Given a pattern p , $\text{coRT}_2^2(p)$ is the statement “For every coloring $f : [\mathbb{N}]^2 \rightarrow 2$ avoiding p , there exists an infinite homogeneous set.”

The statements $\text{RT}_2^2(p)$ and $\text{coRT}_2^2(p)$ were studied both classically and from a reverse mathematical viewpoint for small patterns p . The Ramsey-like statements for permutations of size 3 were studied by multiple authors, and summarized in Cervelle, Gaudelier, and Patey [16, Section 1.2].

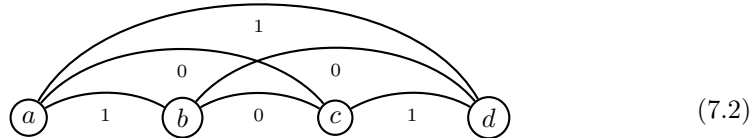
Note that if the pattern p is a constant function $i < 2$, then the statement $\text{RT}_2^2(p)$ is combinatorially false, as witnessed by the function $f : [\mathbb{N}]^2 \rightarrow 2$ also equal to i on its whole domain. By Ramsey's theorem for pairs, $\text{RT}_2^2(p)$ is true for any other pattern p . The statements $\text{RT}_2^2(p)$ and $\text{coRT}_2^2(p)$ play a dual role with respect to RT_2^2 , as $\text{RCA}_0 \vdash \text{RT}_2^2 \leftrightarrow \text{RT}_2^2(p) + \text{coRT}_2^2(p)$ for any non-constant pattern p . We will also consider stable instances of $\text{coRT}_2^2(p)$ and define $\text{coSRT}_2^2(p)$ to be coRT_2^2 restricted to stable colorings, i.e., colorings $f : [\mathbb{N}]^2 \rightarrow 2$ such that $\lim_y f(x, y)$ exists for every $x \in \mathbb{N}$. As with the decomposition of RT_2^2 into $\text{SEM} + \text{SADS} + \text{COH}$, we have the decomposition $\text{RCA}_0 \vdash \text{RT}_2^2 \leftrightarrow \text{RT}_2^2(p) + \text{coSRT}_2^2(p) + \text{COH}$ (we will not consider a stable variant of $\text{RT}_2^2(p)$ for our purposes).

Every pattern $p : [\ell]^2 \rightarrow 2$ admits a *dual pattern* $\bar{p} : [\ell]^2 \rightarrow 2$ defined by $\bar{p}(x, y) = 1 - p(x, y)$. In particular, the non-transitivity patterns (7.1) are dual. Over RCA_0 , $\text{RT}_2^2(p)$ and $\text{RT}_2^2(\bar{p})$ are equivalent, as given a coloring $f : [\mathbb{N}]^2 \rightarrow 2$, one can always consider the dual coloring $\bar{f} : [\mathbb{N}]^2 \rightarrow 2$ and apply $\text{RT}_2^2(\bar{p})$. Thus, over RCA_0 , EM is equivalent to the statement $\text{RT}_2^2(p)$, where p is any of (7.1).

7.1.2 Permutations

A *permutation* on $\ell \in \mathbb{N}$ is a bijection $\pi : \ell \rightarrow \ell$. A permutation π can be seen as a linear order $\mathcal{L}_\pi = (\ell, <_\pi)$ defined for every $\{x, y\} \in [\ell]^2$ by $x <_\pi y$ iff $\pi(x) <_\mathbb{N} \pi(y)$. There is therefore a one-to-one correspondence between permutations on ℓ and patterns of size ℓ avoiding both non-transitivity patterns (see (7.1)).

Indeed, given a permutation π on ℓ , one can define a unique pattern $p_\pi : [\ell]^2 \rightarrow 2$ coding the graph of $<_\pi$, that is, for $x <_\mathbb{N} y <_\mathbb{N} \ell$, $p_\pi(x, y) = 0$ if $x <_\pi y$ and $p_\pi(x, y) = 1$ otherwise. Thanks to this correspondence, we shall consider permutations as particular RT_2^2 -patterns, and will use either representation when most convenient. Permutations enjoy a simple notation as their range sequence. For instance, the sequence 2031 denotes the permutation π on 4 defined by $\pi(0) = 2$, $\pi(1) = 0$, $\pi(2) = 3$ and $\pi(3) = 1$. Under the pattern representation, this corresponds to



Permutations enjoy two natural associative operators, namely, the skew sum $\ominus$ and the direct sum $\oplus^1$, combining shorter permutations into longer ones. They are defined as follows: for π and σ two permutations of size m and n , respectively,

$$(\pi \ominus \sigma)(x) = \begin{cases} \pi(x) + n & \text{for } x < m, \\ \sigma(x - m) & \text{for } m \leq x < m + n, \end{cases}$$

$$(\pi \oplus \sigma)(x) = \begin{cases} \pi(x) & \text{for } x < m, \\ \sigma(x - m) + m & \text{for } m \leq x < m + n. \end{cases}$$

These sums enable us to define the class of separable permutations, which will play a central role in this chapter.

¹We will also use the symbol $\oplus$ to denote the join of two sets, as defined by Definition 4.1.3

Definition 7.1.3. A permutation is *separable* if it can be obtained from the trivial permutation 0 by direct sums and skew sums.

Bose, Buss, and Lubiw [10] introduced the notion of separable permutation as a natural family for which multiple problems (sub-pattern decision, longest common sub-pattern) admit a polynomial-time procedure. Separable permutations admit the following historical characterization in terms of pattern avoidance:

Theorem 7.1.4 (Bose, Buss and Lubiw [10]). *A permutation is separable if and only if it does not contain the permutations 1302 and 2031 as sub-patterns.*

The main interest of this characterization in this article is the existence of a finite number of patterns (the two non-transitivity patterns and the permutations 1302 and 2031) such that every coloring avoiding all of them also avoids all non-separable patterns.

By Lerman, Solomon, and Towsner [94, Theorem 1.15], EM does not imply RT_2^2 over ω -models. Hence, $\text{RT}_2^2(p)$ does not imply RT_2^2 over ω -models when p is not a permutation.

This result was generalized by a theorem of Mimouni and Patey [102, Main Theorem 1.8], who showed that $\text{RT}_2^2(p)$ does not imply RT_2^2 over ω -models when p is not a separable permutation. Their result was not directly expressed in the formalism of separable permutations, but rather using another equivalent characterization of them, more suitable for their computability-theoretic analysis (see [90, Section 2] for details on this equivalence).

Their theorem implied that the separable permutations are exactly the patterns p for which the statement $\text{RT}_2^2(p)$ does not admit so-called *preservation of ω hyperimmunities*, where a function $f : \mathbb{N} \rightarrow \mathbb{N}$ is *(X-)hyperimmune* if it is not dominated by any *(X-)computable* function.

Definition 7.1.5. A problem P admits *preservation of ω hyperimmunities* if for every set Z , every countable sequence $f_0, f_1, \dots$ of Z -hyperimmune functions and every Z -computable P -instance X , there is a P -solution Y to X such that $f_0, f_1, \dots$ are all $Y \oplus Z$ -hyperimmune.

Being able to preserve ω hyperimmunities is a weakness property of problems that RT_2^2 does not enjoy (in fact RT_2^2 can only preserve at most 1 hyperimmunity). This distinction can be used to build an ω -model of $\text{RT}_2^2(p)$ that is not an ω -model of RT_2^2 for every pattern p that is not a separable permutation (with a construction similar to that of the proof of Proposition 8.1.5).

In this chapter, we prove the converse, showing that every ω -model of $\text{RT}_2^2(p)$ where p is a separable permutation is also a model of RT_2^2 , thereby establishing the following characterization.

Main Theorem 7.1.6. *For every pattern p , $\text{RT}_2^2(p)$ implies RT_2^2 over ω -models if and only if p is a separable permutation.*

To do so, we show that in the decomposition of RT_2^2 into $\text{RT}_2^2(p) + \text{COH} + \text{coSRT}_2^2(p)$ for non-constant separable permutations, the statements COH and $\text{coSRT}_2^2(p)$ are already implied by $\text{RT}_2^2(p)$ over ω -models.

More precisely, for every separable permutation p , we prove that $\text{RCA}_0 \vdash \text{RT}_2^2(p) \rightarrow \text{ADS}$ in Section 7.2, and thus that $\text{RCA}_0 \vdash \text{RT}_2^2(p) \rightarrow \text{COH}$ as $\text{RCA}_0 \vdash \text{ADS} \rightarrow \text{COH}$ (Hirschfeldt and Shore [63]).

And we prove that for every separable permutation p , every computable instance of $\text{coSRT}_2^2(p)$ admits a probabilistic solution. To formalize this, we consider the product topology on the space $2^{\mathbb{N}}$, whose basic open sets are of the form $[\sigma] = \{X \in 2^{\mathbb{N}} : \sigma < X\}$ for each $\sigma \in 2^{<\mathbb{N}}$. We then equip this space with the standard measure λ , defined by $\lambda([\sigma]) = \frac{1}{2^{|\sigma|}}$ and then extended to all Borel sets. We then prove that, for every such instance, the set of oracles computing a solution has measure 1.

Most statements we consider do not admit probabilistic solutions. For example, there are instances of EM, ADS and WKL for which the class of oracles computing a solution has measure 0 (Astor, Bienvenu, Dzhafarov, Patey, Shafer, Solomon, and Westrick [3]).

A central theme in the field of algorithmic randomness (see Nies [105] or Downey and Hirschfeldt [37] for a good introduction to the topic) is to quantify how much randomness is needed for a probabilistic statement to hold. In our case, we consider the sets with the property (true almost everywhere) of computing a so-called $\emptyset'$ -DNC function, and prove that any such set computes a solution to any computable instance of $\text{coSRT}_2^2(p)$. This is also the case for computable instances of $\text{coRT}_2^2(p)$, although the proof is more involved (as done in the original article [90, Section 6]).

7.1.3 Diagonal non-computation

In Chapter 4, we introduced the $X\text{-DNC}_2$ functions as those functions $f : \mathbb{N} \rightarrow \{0, 1\}$ satisfying $f(e) \neq \Phi_e^X(e)$ for every $e \in \mathbb{N}$. If we relax the requirement that the codomain be $\{0, 1\}$ and instead allow values in $\mathbb{N}$, we obtain the historical definition of the DNC functions:

Definition 7.1.7 ($X\text{-DNC}$). A function $f : \mathbb{N} \rightarrow \mathbb{N}$ is *diagonally non- X -computable* or $X\text{-DNC}$ if for every $e \in \mathbb{N}$, $f(e) \neq \Phi_e^X(e)$. A Turing degree is $X\text{-DNC}$ if it computes an $X\text{-DNC}$ function.

Intuitively, it is easy to construct an $X\text{-DNC}$ function. Let $f : \mathbb{N} \rightarrow \mathbb{N}$ be defined by choosing each $f(e)$ uniformly at random from $\{0, \dots, 10^{e+1} - 1\}$. Then, f is $X\text{-DNC}$ with probability at least $\prod_{e \in \mathbb{N}} (1 - \frac{1}{10^{e+1}}) > 0$. By replacing the sequence $e + 1$ with $e + C$ for $C \in \mathbb{N}$ large enough, this probability can be made arbitrarily close to 1. Thus, although no $X\text{-DNC}$ function is X -computable by definition, the set of oracles that compute one can be shown to be of measure 1, regardless of the set X .

An $X\text{-DNC}$ function does not need to compute the set X . In fact, when X is not computable, most $X\text{-DNC}$ functions do not compute X , since the set of oracles computing a non-computable set has measure 0. This contrasts with $X\text{-DNC}_2$ functions, which always compute X . Indeed, equivalently, every set Y of PA degree over X satisfies $Y \geq_T X$, as witnessed by the X -computable infinite tree whose only path is X .

The diagonally non-computable degrees have many characterizations, using various paradigms, such as fixpoint-free functions ([68]), infinite subsets of random sequences ([74, 54], see [37, Theorem 8.10.2]), sequences of high Kolmogorov complexity ([75]). The following characterization, in terms of uniform avoidance of finite $\Sigma_1(Z)$ sets with bounded size, is arguably the most useful:

Definition 7.1.8. A function $g : \mathbb{N}^2 \rightarrow \mathbb{N}$ is *Z -escaping* if for every $e, k \in \mathbb{N}$, if $\text{card } W_e^Z \leq k$, then $g(e, k) \notin W_e^Z$ where W_e^Z is the e -th Z -computably enumerable set, as defined in Chapter 4.

Lemma 7.1.9 (Folklore (see [7, Lemma 3.3])). *Fix a set Z and a Turing degree $\mathbf{d}$. The following are equivalent:*

- (1) $\mathbf{d}$ computes a $Z\text{-DNC}$ function;
- (2) $\mathbf{d}$ computes a Z -escaping function.

Proof. (1) $\rightarrow$ (2). Given a $Z\text{-DNC}$ function f and $e, k \in \mathbb{N}$, define $g(e, k)$ as follows: For every $i \leq k$, let $\Phi_{a_i}^Z$ be the function which, on every input, searches for the i -element x_i of W_e^Z in order of appearance, if it exists, and interprets x_i as a $(k+1)$ -tuple $\langle x_i^0, \dots, x_i^k \rangle$ and outputs x_i^i . Let $g(e, k) = \langle f(a_0), \dots, f(a_k) \rangle$. Suppose for contradiction that $\text{card } W_e^Z \leq k$ but $g(e, k) = x_i \in W_e^Z$ for some $i \leq k$. Then $\Phi_{a_i}^Z(a_i) = x_i^i = f(a_i)$, contradicting the fact that f is $Z\text{-DNC}$.

(2) $\rightarrow$ (1). Given a Z -escaping function g , and $e \in \mathbb{N}$, define $f(e)$ as follows: Let W_a^Z be the Z -c.e. set whose unique element is $\Phi_e^Z(e)$, if it exists, and $W_a^Z = \emptyset$ otherwise. Let $f(e) = g(a, 1)$. By construction, $\text{card } W_a^Z \leq 1$, so $g(a, 1) \notin W_a^Z$, hence $f(e) = g(a, 1) \neq \Phi_e^Z(e)$. $\square$

One can refine the previous lemma to require the Z -escaping function to output its value within an infinite computable set $X \subseteq \mathbb{N}$, by transforming any Z -c.e. set $W_e^Z \subseteq X$ into the set of positions of its elements in X .

By a result of Mimouni and Patey [102, Main Theorem 1.2], there exists a computable coloring $f : [\mathbb{N}]^2 \rightarrow 2$ such that for every infinite set H f -avoiding any pattern, H computes a $\emptyset'$ -DNC function. Thus, by relativization, for every set X and every pattern p , there is an X -computable instance of $\text{RT}_2^2(p)$ such that for every solution Y , $Y \oplus X$ -computes an X' -DNC function. And, as we will see in Theorem 7.5.6, if p is a separable permutation, every X -computable instance of $\text{coSRT}_2^2(p)$ can be solved from $X \oplus Y$. Hence, together with Proposition 7.2.1 implying that $\text{RCA}_0 \vdash \text{RT}_2^2(p) \rightarrow \text{COH}$, we get that every ω -model of $\text{RCA}_0 + \text{RT}_2^2(p)$ is a model of $\text{RCA}_0 + \text{RT}_2^2(p) + \text{COH} + \text{coSRT}_2^2(p)$ or equivalently of $\text{RCA}_0 + \text{RT}_2^2$, which is what we wanted.

7.2 Separable permutations and ADS

The goal of this section is to prove that $\text{RCA}_0 \vdash \text{RT}_2^2(p) \rightarrow \text{COH}$ for any separable permutation p . We shall actually prove a stronger implication, namely, $\text{RCA}_0 \vdash \text{RT}_2^2(p) \rightarrow \text{ADS}$, by showing that ADS restricted to linear orders avoiding any separable permutation is computably true.

Proposition 7.2.1. *Let $\mathcal{L} = (\mathbb{N}, <_{\mathcal{L}})$ be a computable linear order avoiding a separable permutation p . Then there is an infinite computable $\mathcal{L}$ -ascending or $\mathcal{L}$ -descending sequence.*

Proof. We show by induction on p that for every computable linear order $\mathcal{L} = (\mathbb{N}, <_{\mathcal{L}})$ without any infinite computable $\mathcal{L}$ -increasing or $\mathcal{L}$ -decreasing sequence, the pattern p appears. More precisely, we proceed by structural induction on the construction of p by a series of direct and skew summations, starting from the trivial permutation with one element. This statement holds for the trivial permutation. The direct sum and the skew sum playing a dual role, we prove the inductive step only for the direct sum.

Let p be the direct sum of two permutations p_0 and p_1 satisfying the induction hypothesis and let $\mathcal{L} = (\mathbb{N}, <_{\mathcal{L}})$ be a computable linear order without any infinite computable $\mathcal{L}$ -increasing or $\mathcal{L}$ -decreasing sequence, hence both of these patterns appear in $\mathcal{L}$.

If the pattern p_0 appears in $\mathcal{L}$ at position F_0 with the element $x_0 = \max_{\mathcal{L}} F_0$ such that the set $\{x \in \mathbb{N} : x >_{\mathcal{L}} x_0\}$ is infinite, then we are done, as we can consider the induced order $(\{x \in \mathbb{N} : x >_{\mathcal{L}} x_0 \wedge x > \max F_0\}, <_{\mathcal{L}})$ and apply the induction hypothesis on it (by first taking a computable bijection between this order and one having $\mathbb{N}$ as base set), thus, the pattern p_1 appears in that order at some position F_1 , and by construction, $F_0 < F_1$ and $F_0 <_{\mathcal{L}} F_1$ holds (where $A < B$ and $A <_{\mathcal{L}} B$ are shorthand for $\forall x \in A \forall y \in B \ x < y$ and $\forall x \in A \forall y \in B \ x <_{\mathcal{L}} y$ respectively), thus $F_0 \cup F_1$ $\mathcal{L}$ -realizes $p_0 \oplus p_1$.

If for every set F_0 $\mathcal{L}$ -realizing p_0 , we have that $x_0 = \max_{\mathcal{L}} F_0$ satisfies that $\{x \in \mathbb{N} : x >_{\mathcal{L}} x_0\}$ is finite, then we can contradict our assumption on the order by computing an infinite $\mathcal{L}$ -decreasing sequence. Indeed, under our assumptions, there exists an infinite computable sequence $F^0 < F^1 < \dots$ of sets realizing p_0 and such that the sequence $(x_i)_{i \in \mathbb{N}}$, where $x_i = \max_{\mathcal{L}} F^i$, is infinite and $\mathcal{L}$ -decreasing.

Such a sequence can be defined as follows: for every $i \in \mathbb{N}$, assuming that F_j has been defined for every $j < i$, the set $\bigcup_{j < i} \{x \in \mathbb{N} : x >_{\mathcal{L}} x_j\}$ is finite, as it is

equal to $\{x \in \mathbb{N} : x >_{\mathcal{L}} x_{i-1}\}$ if $i > 0$ and to $\emptyset$ if $i = 0$, hence its complement is infinite and computable and we can apply the induction hypothesis on the induced order $(\{x \in \mathbb{N} : \forall j < i, x <_{\mathcal{L}} x_j \wedge x > \max F^j\}, <_{\mathcal{L}})$ and computably find an instance of p_0 at a position F^i in that ordering. By construction, $x_i := \max_{\mathcal{L}} F^i <_{\mathcal{L}} x_j$ and $F^i > F^j$ for every $j < i$.

□

Proposition 7.2.2. *For every separable permutation p , $\text{RCA}_0 \vdash \text{RT}_2^2(p) \rightarrow \text{ADS}$.*

Proof. The proof of Proposition 7.2.1 is formalizable over RCA_0 , except the induction on the structure of p which is too complex. Thus, for every separable permutation p , we have a proof of $\text{RT}_2^2(p) \rightarrow \text{ADS}$ in RCA_0 , but this does not give us a proof of $\text{RCA}_0 \vdash \forall p \text{ separable}(\text{RT}_2^2(p) \rightarrow \text{ADS})$. □

In particular, for every separable permutation p , $\text{RT}_2^2(p)$ implies the cohesiveness principle (COH) and the Σ_2^0 -bounding scheme ($\text{B}\Sigma_2^0$) over RCA_0 , as ADS implies both statements (see Hirschfeldt and Shore [63, Proposition 2.10, Proposition 4.5]).

We shall now see a direct proof of the second statement in the more general case of all permutations.

Proposition 7.2.3. *Let p be a permutation. Then $\text{RCA}_0 \vdash \text{RT}_2^2(p) \rightarrow \text{B}\Sigma_2^0$.*

Proof. As $\text{B}\Sigma_2^0$ is equivalent to RT^1 modulo RCA_0 , we consider by contrapositive some coloring $f : \mathbb{N} \rightarrow a$ for some $a \in \mathbb{N}$ with no infinite f -homogeneous subset.

Let $g : [\mathbb{N}]^2 \rightarrow 2$ be defined by $g(x, y) = 1$ if $f(x) < f(y)$ and $g(x, y) = 0$ otherwise. We claim that for every permutation p , there is no infinite set g -avoiding p .

Firstly, by external induction on n (see Chapter 9 for a more precise distinction between internal induction which is the induction available within a (potentially non-standard) model, used to prove a property for all elements of that model, and external induction, the usual form of induction, which proves a property only for the standard integers, but without any restriction on the complexity of the formulas involved), we prove that for every infinite set $H = \{x_0, x_1, \dots\}$ there exists some $x_{i_0} < \dots < x_{i_{n-1}}$ such that $f(x_{i_0}) < \dots < f(x_{i_{n-1}})$. Note that if there exists one such n -tuple, there exists an infinite amount of them, by truncating the beginning of H .

If $n = 1$ (or $n = 0$), the property holds. So assume the property to be true for some standard integer n and fix such an infinite set $H = \{x_0, x_1, \dots\}$. Let $H' \subseteq H$ be the set of all $x_{i_{n-1}} \in H$ such that there exists some $x_{i_0} < \dots < x_{i_{n-1}} \in H$ with $f(x_{i_0}) < \dots < f(x_{i_{n-1}})$. By the inductive hypothesis, this set is infinite. If for some $x_{i_{n-1}} < x_{j_{n-1}} \in H'$ we have $f(x_{i_{n-1}}) < f(x_{j_{n-1}})$, then $x_{i_0} < x_{i_1} < \dots < x_{i_{n-1}} < x_{j_{n-1}}$ is the desired sequence. Otherwise, the values taken by f on the elements of H' form an infinite non-increasing sequence of integers smaller than a , which must eventually be constant by $|\Sigma_1^0|$, contradicting the assumption that no infinite f -homogeneous set exists. This proves that for every standard integer n , no infinite set can g -avoid the homogeneous set of size n and of color 1.

Let $p : n \rightarrow n$ be a permutation on n elements and let $H \subseteq \mathbb{N}$ be an infinite set. There exists an infinite sequence of n -tuples of elements of H

$$(x_0^0 < \dots < x_{n-1}^0) < (x_0^1 < \dots < x_{n-1}^1) < \dots$$

each g -homogeneous for the color 1. Notice that for a given n -tuple $(x_0, \dots, x_{n-1})$, there are a^n possible values for $(f(x_0), \dots, f(x_{n-1}))$, so, by the finite pigeonhole principle for a^n colors (provable in RCA_0), there exists a sequence $k_0 < \dots < k_{n-1}$ such that $f(x_i^{k_\ell}) = f(x_i^{k_{\ell'}})$ for every $\ell < \ell' < n$ and every $i < n$. Then, the set $\{x_0^{k_{p(0)}} < x_1^{k_{p(1)}} < \dots < x_{n-1}^{k_{p(n-1)}}\}$ g -realizes p . □

7.3 Basis for separable permutations

The separable permutations form a relatively large family. In this section, we define canonical sub-families for separable permutations, the k -fractals for $k \in \mathbb{N}$, which enjoy some universality properties: for every $k \geq 2$, every separable permutation is a sub-pattern of a k -fractal. These families will play an essential role in Section 7.5 to prove our main theorem.

Definition 7.3.1. A family $\mathcal{P}$ of patterns is a *basis* for another family $\mathcal{Q}$ of patterns if every pattern $q \in \mathcal{Q}$ is a sub-pattern of some pattern $p \in \mathcal{P}$.

Suppose $\mathcal{P}$ is a basis for $\mathcal{Q}$. Then the statement “For every $p \in \mathcal{P}$, every coloring avoiding p satisfies a given property” implies “For every $q \in \mathcal{Q}$, every coloring avoiding q satisfies the same property”. Indeed, if a coloring does not satisfy the property, then every pattern in $\mathcal{P}$ appears in it, and as $\mathcal{P}$ is a basis for $\mathcal{Q}$, every pattern in $\mathcal{Q}$ also appears.

Definition 7.3.2. The family $\mathcal{F}_k$ of k -fractals is defined inductively as follows:

- The k -fractal of dimension 0 is the trivial permutation $p : [1]^2 \rightarrow 2$.
- Let p_n be the k -fractal of dimension n . The k -fractal of dimension $n + 1$ is the permutation p_{n+1} defined as the direct sum $p_n \oplus \cdots \oplus p_n$ of k copies of p_n if n is even, and as the skew sum $p_n \ominus \cdots \ominus p_n$ of k copies of p_n if n is odd.

Note that since the skew sum and the direct sum are associative, the previous definition is well-defined.

Remark 7.3.3. For a fixed coloring $f : [\mathbb{N}]^2 \rightarrow 2$, by a small abuse of notation, we shall simply call k -fractal of dimension n any finite set f -realizing the k -fractal of dimension n .

Example 7.3.4. Let $f : [\mathbb{N}]^2 \rightarrow 2$ be a coloring. The sets f -realizing the k -fractal of dimension 0 are the singletons. The sets f -realizing the k -fractals of dimension 1 are the f -homogeneous sets for color 0 of cardinality k . The sets f -realizing the k -fractals of dimension 2 are the finite sets of cardinality k^2 which can be partitioned into k blocks $F_0 < F_1 < \cdots < F_{k-1}$, such that each F_i is f -homogeneous for color 0, and the pairs between any two blocks have color 1.

Proposition 7.3.5. For every $k \geq 2$, the family $\mathcal{F}_k$ is a basis for the family of separable permutations.

Proof. Inductively on the structure of a separable permutation p , obtained as a series of direct and skew sums, we can show that there exists some n such that p is a sub-pattern of the k -fractal of dimension n .

This is the case for the trivial permutation of a single point. It is a sub-pattern of the k -fractal of dimension 0.

If $p = p_0 \oplus p_1$, with n_0, n_1 such that p_i is a sub-pattern of the k -fractal of dimension n_i for every $i < 2$. Let n be any odd number greater than $\max(n_0, n_1)$. Then p_0 and p_1 are sub-patterns of the k -fractal of dimension $n - 1$, hence $p = p_0 \oplus p_1$ is a sub-pattern of the k -fractal of dimension n (since $k \geq 2$).

The case where p is a skew sum of two separable permutations p_0, p_1 is similar, with n taken to be even. $\square$

7.4 The example of $\text{coSRT}_2^2(120)$

We will prove in Proposition 7.5.5 that for all $n, k \in \mathbb{N}$, any stable coloring f that avoids the k -fractal of dimension n has the property that every $\emptyset'$ -DNC degree $\mathbf{d}$

computes an infinite f -homogeneous set. Hence, as the k -fractals of dimension n form a basis for the separable permutations, the statement will also hold for stable colorings avoiding some separable permutation. Our proof will proceed by induction on the dimension n of the k -fractal. The case $n = 0$ is immediate, since no coloring avoids a k -fractal of dimension 0. The case $n = 1$ is more interesting, and a more detailed analysis of its general form, not restricted to stable colorings, will be carried out in Chapter 11, where it will be referred to as the *bounded Ramsey theorem for pairs*, denoted by BRT_2^2 .

Statement 7.4.1 (Bounded Ramsey theorem for pairs). BRT_2^2 is the statement “For every coloring $f : [\mathbb{N}]^2 \rightarrow 2$ and every $\ell \in \mathbb{N}$ such that there is no f -homogeneous set for color 1 of size ℓ , there is an infinite f -homogeneous set for color 0.”

In contrast with the general situation, this case does not only admit a probabilistic solution: it is in fact computably true, in the sense that every computable instance has a computable solution.

Lemma 7.4.2. *Let $f : [\mathbb{N}]^2 \rightarrow 2$ be a computable coloring and $k \in \mathbb{N}$ such that no clique of size k is f -homogeneous for the color 0. Then there exists some infinite computable f -homogeneous set for the color 1.*

Belanger’s proof over IS_2^0 . Let $T \subseteq \mathbb{N}^{<\mathbb{N}}$ be a computable tree such that

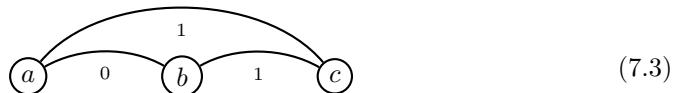
- (1) every $x \in \mathbb{N}$ appears in some node $\sigma \in T$;
- (2) every $\sigma \in T$ is f -homogeneous for color 0;
- (3) for every $\sigma \in T$ and every $x, y \in \mathbb{N}$ such that $\sigma \cdot x, \sigma \cdot y \in T$, $f(x, y) = 1$.

T can be constructed via a greedy algorithm: at each stage, we insert the next element $x \in \mathbb{N}$ into a suitable position in the tree, ensuring that properties (1) to (3) are preserved.

By (2) and by assumption, T has finite depth, so by (1) and BS_2^0 , there is some level $\ell_0 \in \mathbb{N}$ such that $\{\sigma \in T : |\sigma| = \ell_0\}$ is infinite. By IS_2^0 , there is a least such level $\ell \leq \ell_0$. Since T has only one node at level 0, namely $\langle \rangle$, $\ell > 0$. By minimality of ℓ , $\{\sigma \in T : |\sigma| = \ell - 1\}$ is finite, so by BS_2^0 , there is some $\sigma \in T$ of length $\ell - 1$ with infinitely many immediate children. The set $\{x \in \mathbb{N} : \sigma \cdot x \in T\}$ is therefore infinite, and by (3), it is f -homogeneous for color 1. $\square$

In Chapter 11, we will show the statement BRT_2^2 to be of peculiar strength. On the one hand, the mere existence of a solution can be established using RT_2^2 , since BRT_2^2 is a special case of it, hence $\text{RCA}_0 + \text{BRT}_2^2$ does not imply IS_2^0 either. On the other hand, we will show that proving the existence of a computable solution does require IS_2 , demonstrating that the above proof is optimal in this respect. Since we restrict attention to ω -models throughout this chapter, we will not track the amount of induction required in our proofs, and will instead focus on the second-order content of the statements under consideration. From this perspective, the statement BRT_2^2 is trivial, as it holds in every ω -model of RCA_0 .

Before proceeding with our induction and proving the general case for arbitrary k -fractals of dimension n , we consider the case of the permutation 120 (corresponding to the pattern below), and show that the statement $\text{coSRT}_2^2(120)$ admits probabilistic solutions.



The statement $\text{coRT}_2^2(120)$ was already studied by various authors under the names **SHER** (Dorais et al. [36]), **TAC** (Conidis [29]), **CAC for trees**, and **SAC** (Cervelle, Gaudelier, and Patey [16]). In particular, Cervelle, Gaudelier, and Patey [16, Section 3] proved that every computable instance of **SAC** admits a probabilistic solution. We give here the argument in the language of $\text{coSRT}_2^2(120)$ as it already contains the core ideas of the generalization to all separable permutations.

Lemma 7.4.3 (Cervelle, Gaudelier, and Patey [16, Section 3]). *$\text{coSRT}_2^2(120)$ admits probabilistic solutions.*

Proof. Let $f : [\mathbb{N}]^2 \rightarrow 2$ be a stable computable coloring avoiding the permutation 120.

Given an infinite set $X \subseteq \mathbb{N}$, we say that an element $x \in \mathbb{N}$ is *f-bad* if $\lim_{y \rightarrow \infty} f(x, y) = 1$. Otherwise (i.e., if $\lim_{y \rightarrow \infty} f(x, y) = 0$), x is *f-good*.

Since f avoids 120, every finite f -homogeneous set $F \subseteq \mathbb{N}$ for color 0 contains at most one f -bad element. Thus, if one picks an element $x \in F$ at random, the probability that it is f -bad is at most $\frac{1}{\text{card } F}$. By Lemma 7.4.2, if there exists some bound $b \in \mathbb{N}$ and some $\ell \in \mathbb{N}$ such that no subset $F > b$ of cardinality ℓ is f -homogeneous for color 0, then there exists an infinite computable subset of $\{b, b+1, \dots\}$ that is f -homogeneous for color 1. In that case, no probabilistic argument is needed, so we may assume that this situation does not occur.

Consider the following probabilistic algorithm to compute an infinite f -homogeneous set Y for color 0:

Y will be defined as $\bigcup_{s \in \mathbb{N}} E_s$ for $E_0 \subseteq E_1 \subseteq \dots$ an increasing sequence of finite sets, satisfying the following additional constraints for every $s \in \mathbb{N}$:

- E_s is f -homogeneous for the color 0, and hence Y is as well.
- $\text{card } E_s = s$, and hence Y is infinite.

At stage 0, set $E_0 = \emptyset$.

At stage s , assume that E_s has been defined. Search for some finite f -homogeneous set $F_s > E_s$ for color 0 of size 2^{s+2} , such that $f(x, y) = 0$ for every $x \in E_s$ and every $y \in F_s$. Pick an element $x \in F_s$ uniformly at random, and let $E_{s+1} = E_s \cup \{x\}$ and go to the next step.

If every element of E_s is f -good, then there exists a bound $b \in \mathbb{N}$ such that for all $y > b$ and all $x \in E_s$, we have $f(x, y) = 0$. It follows, by our assumption, that a suitable set F_s will eventually be found. Hence, as long as the property that every element of E_s is f -good is preserved, the construction can proceed.

Since every f -homogeneous set for color 0 contains at most one f -bad element, the probability that an f -bad element is chosen at stage s is at most $\frac{1}{\text{card } F_s} = 2^{-s-2}$. Therefore, the probability that the construction fails at some stage is bounded by $\sum_s 2^{-s-2} = \frac{1}{2}$. More generally, one can parameterize the construction by a constant $k \in \mathbb{N}$, and require $\text{card } F_s = 2^{s+1+k}$, in which case the probability of failure is at most 2^{-k} .

By encoding the choices within the blocks $\{F_s\}_{s \in \mathbb{N}}$ into chunks of an infinite binary sequence, that is, decomposing any $R \in 2^{\mathbb{N}}$ into an infinite concatenation $\sigma_0^R \cdot \sigma_1^R \cdot \dots$ where σ_s^R is a string of length 2^{s+k+1} coding for the choice of an element in F_s , the measure of oracles $R \in 2^{\mathbb{N}}$ such that this decomposition yields a successful construction is at least $1 - 2^{-k}$. This tends to 1 when k tends to infinity, so the measure of oracles computing a solution is 1. $\square$

An analysis of the complexity of the above argument using the tools of algorithmic randomness shows that a so-called 2-random degree is sufficient to carry out the construction. On the other hand, at first sight, the existence of an $\emptyset'$ -DNC function is not sufficient. Indeed, although the property of being f -bad is Σ_2 , or equivalently $\Sigma_1(\emptyset')$, an $\emptyset'$ -DNC function is only sufficient to avoid selecting a finite number of f -bad elements from an infinite set and doesn't provide a way of picking an element in F_s .

Every 2-random degree is $\mathcal{O}'$ -DNC, but the converse doesn't hold. Unfortunately, contrary to $\mathcal{O}'$ -DNC degrees, there is no hope to prove the existence of a 2-random degree from $\text{RT}_2^2(p)$, as Liu [96] proved that every computable coloring $f : [\mathbb{N}]^2 \rightarrow 2$ admits an infinite f -homogeneous set which does not compute any 1-random, and a fortiori any 2-random.

Thankfully, a more involved argument, based on the following lemma, allows us to reduce the required amount of randomness to that provided by $\mathcal{O}'$ -DNC degrees. For that, we will exploit the fact that, instead of picking one f -good element from a finite f -homogeneous set F_s , we may instead consider infinitely many such f -homogeneous sets of arbitrarily large sizes.

Lemma 7.4.4. *Let $B \subseteq \mathbb{N}$ be a Σ_2 set of “bad elements”, $(F_i)_{i \in \mathbb{N}}$ be a computable family of subsets of $\mathbb{N}$ and $k \in \mathbb{N}$ such that: $\text{card } F_i \cap B \leq k$ and $\text{card } F_i \geq i$ for every i . Then, every $\mathcal{O}'$ -DNC degree $\mathbf{d}$ computes an infinite subset of $\bigcup_i F_i \setminus B$.*

Proof. Fix such $B \subseteq \mathbb{N}$, $(F_i)_{i \in \mathbb{N}}$, $k \in \mathbb{N}$ and $\mathbf{d}$. We have two cases.

Case 1: $\mathbf{d} \geq_T \mathcal{O}'$. Let $\ell \leq k$ be the largest integer such that $\text{card } F_i \cap B = \ell$ for infinitely many $i \in \mathbb{N}$. Let $t \in \mathbb{N}$ be such that for every $i \geq t$, $\text{card } F_i \cap B \leq \ell$. As B is Σ_2 , one can $\mathcal{O}'$ -computably find infinitely many $i \geq t$ such that $\text{card } F_i \cap B \geq \ell$, hence such that $\text{card } F_i \cap B = \ell$, and pick an element in $F_i \setminus B$.

Case 2: $\mathbf{d} \not\geq_T \mathcal{O}'$. Let $g \leq_T \mathbf{d}$ be the function of Lemma 7.1.9(2). Let μ_s be the standard left-c.e. approximation of the modulus μ of $\mathcal{O}'$, that is, $\mu_s(x)$ is the least $t \leq s$ such that $\mathcal{O}'_t \upharpoonright x = \mathcal{O}'_s \upharpoonright x$. For every $x \in \mathbb{N}$, the sequence $(\mu_s(x))_{s \in \mathbb{N}}$ can only change values at most x times (every time a program with Turing index less than x stops), and let $m_0^x < \dots < m_i^x (= \mu(x))$ be those values for some $i < x$. Then, consider the $\mathcal{O}'$ -enumerable set $W_{e(x)}^{\mathcal{O}'}$ containing $\{0, \dots, x-1\}$ and the position in $F_{m_j^x}$ of the elements of $F_{m_j^x} \cap B$ for every $j < i$.

Note that $\text{card } W_{e(x)}^{\mathcal{O}'} \leq x \times (k+1)$, hence, by hypothesis on g , letting $p_x = g(e(x), x \times (k+1))$, the sequence $(p_x)_{x \in \mathbb{N}}$ is such that $p_x \notin W_{e(x)}^{\mathcal{O}'}$ for every x . In particular, $p_x \geq x$ for every x as $\{0, \dots, x-1\} \subseteq W_{e(x)}^{\mathcal{O}'}$. Since $\mathbf{d} \not\geq_T \mathcal{O}'$, the sequence $(p_x)_{x \in \mathbb{N}}$ does not eventually dominate μ , hence, for infinitely many x , we can $\mathbf{d}$ -computably find some stage s such that $\mu_s(x) \geq p_x$ and if s is chosen to be minimal, this ensures that the element in position p_x of $F_{\mu_s(x)}$ does not belong to B . Let S be this set of elements. Then $S \subseteq \bigcup_i F_i \setminus B$, and since $p_x \geq x$, S is infinite. $\square$

We can now obtain an effective version of Lemma 7.4.3.

Lemma 7.4.5. *For every computable instance of $\text{coSRT}_2^2(120)$, every $\mathcal{O}'$ -DNC degree computes a solution.*

Proof. Let f be an instance of $\text{coSRT}_2^2(120)$. As explained in the proof of Lemma 7.4.3, for every finite f -homogeneous set $F \subseteq \mathbb{N}$ for color 0, at most one element has limit 1. If we let $B = \{x : \lim_y f(x, y) = 1\}$ and $(F_n)_{n \in \mathbb{N}}$ be a sequence of f -homogeneous sets for color 0 such that $\text{card } F_n = n$ (if no such sequence exists, then by Lemma 7.4.2 there is a computable f -homogeneous set and we are done), then by Lemma 7.4.4, any $\mathcal{O}'$ -DNC degree computes an infinite subset of $A = \{x : \lim_y f(x, y) = 0\}$, hence computes an infinite f -homogeneous set (see e.g., the proof of the equivalence between SRT_2^2 and D_2^2 in Chapter 6). $\square$

This yields a new proof for the following theorem.

Theorem 7.4.6 (Cervelle, Gaudelier, and Patey [16]). *$\text{RCA}_0 + \text{RT}_2^2(120)$ implies RT_2^2 over ω -models.*

Proof. Let $\mathcal{M} = (\mathbb{N}, S)$ be an ω -model of $\text{RCA}_0 + \text{RT}_2^2(120)$. By Proposition 7.2.2, we have that $\mathcal{M} \models \text{ADS}$ and thus $\mathcal{M} \models \text{COH}$.

Let $f : [\mathbb{N}]^2 \rightarrow 2$ be an instance of $\text{coSRT}_2^2(120)$ coded by a set in S . By a relativization of Mimouni and Patey [102, Main theorem 1.2], there exists an f -computable coloring $g : [\mathbb{N}]^2 \rightarrow 2$ such that for every infinite set $H \subseteq \mathbb{N}$ that g -avoids any pattern, $f \oplus H$ computes an f' -DNC function.

As $\mathcal{M} \models \text{RCA}_0$, we have that $g \in S$, and, as $\mathcal{M} \models \text{RT}_2^2(120)$, there exists an infinite set $H \in S$ that g -avoids the pattern 120. Hence, as $f \oplus H \in S$, there exists a set G of f' -DNC degree in S .

By a relativization of Lemma 7.4.5, $G \oplus f$ computes an infinite f -homogeneous set. Thus, every instance of $\text{coSRT}_2^2(120)$ in S admits a solution in S , showing that $\mathcal{M} \models \text{coSRT}_2^2(120)$.

Thus, $\mathcal{M} \models \text{RT}_2^2(120) + \text{COH} + \text{coSRT}_2^2(120)$, and as $\text{RCA}_0 \vdash \text{RT}_2^2 \leftrightarrow \text{RT}_2^2(120) + \text{COH} + \text{coSRT}_2^2(120)$, we have $\mathcal{M} \models \text{RT}_2^2$ as required. $\square$

7.5 General case

We now turn to the study of $\text{coSRT}_2^2(p)$ for arbitrary separable permutations p . We shall first prove the theorem for fractals, and deduce the general case since they form a basis for separable permutations.

Definition 7.5.1. Fix $k, n \geq 1$. Let $f : [\mathbb{N}]^2 \rightarrow 2$ be a coloring and F be a set f -realizing the k -fractal of dimension n . The *fractal decomposition* of F is the sequence $F_0 < \dots < F_{k-1}$ of subsets of F such that for every $i < k$, F_i is a k -fractal of dimension $n - 1$. We shall refer to the F_i 's as *blocks*.

Given a k -fractal F of dimension n , and $\sigma \in \mathbb{N}^{\leq n}$, let $F(\sigma)$ be defined inductively as follows: $F(\langle \rangle) = F$ and $F(w \cdot \tau) = F_w(\tau)$, where $F_0 < \dots < F_{k-1}$ is the fractal decomposition of F , if $w < k$. If $w \geq k$, then $F(w \cdot \tau) = \emptyset$. Note that if $F(\sigma) \neq \emptyset$, then $F(\sigma)$ is a k -fractal of dimension $n - |\sigma|$.

For a stable coloring $f : [\mathbb{N}]^2 \rightarrow 2$, we will often omit explicit reference to f when it is clear from the context, and simply say that an element $x \in \mathbb{N}$ has limit $c < 2$ if $\lim_{y \in \mathbb{N}} f(x, y) = c$.

The following definition generalizes the notion of f -good/ f -bad from Lemma 7.4.3. Indeed, in the case $n = 0$, a finite set F f -realizing the i -fractal of dimension 0 for some $i > 0$ is a singleton, and its unique element has limit c if and only if F is (f, k) -good for color c .

Definition 7.5.2. Fix $k, i \geq 1$ and $n \in \mathbb{N}$. Let $f : [\mathbb{N}]^2 \rightarrow 2$ be a stable coloring and $c < 2$ be a color. A finite set $F \subseteq \mathbb{N}$ f -realizing the i -fractal of dimension n is (f, k) -good for color c if there is no subset $G \subseteq F$ f -realizing the k -fractal of dimension n with all its elements having limit $1 - c$. Otherwise, F is (f, k) -bad for color c .

The following lemma bounds the number of bad choices of fractals of dimension $n - 1$ given a good fractal of dimension n . This yields a probabilistic algorithm to find an element which has limit c given an i -fractal F of dimension n which is (f, k) -good for color c : pick a block F' at random in the fractal decomposition of F . With probability $1 - \frac{k-1}{i}$, F' is an i -fractal of dimension $n - 1$ which is (f, k) -good for color c . Iterate the process until we obtain an i -fractal of dimension 0 which is (f, k) -good for color c , in other words, a singleton element which has limit c . The probability of making a bad choice at some point is at most $\frac{(k-1) \times n}{i}$, which can be made arbitrarily small by considering i -fractals for very large i 's.

Lemma 7.5.3. Fix $k, i, n \geq 1$. Let $f : [\mathbb{N}]^2 \rightarrow 2$ be a stable coloring and $c < 2$ be a color. Let $F \subseteq \mathbb{N}$ be a set f -realizing the i -fractal of dimension n and let $F_0 < \dots < F_{i-1}$ be its fractal decomposition. If F is (f, k) -good for color c , then there are at most $k - 1$ many blocks which are (f, k) -bad for color c .

Proof. Assume, by contraposition, that there exist indices $\ell_0 < \dots < \ell_{k-1} < i$ such that, for every $j < k$, the block F_{ℓ_j} is (f, k) -bad for color c . Since $n \geq 1$ and by definition of a (f, k) -bad block, for every $j < k$ there exists a subset $G_{\ell_j} \subseteq F_{\ell_j}$ such that G_{ℓ_j} f -realizes the k -fractal of dimension $n - 1$ and all its elements have limit $1 - c$.

Then, $G_{\ell_0} \cup \dots \cup G_{\ell_{k-1}}$ f -realizes the k -fractal of dimension n and all its elements have limit $1 - c$, hence, F is (f, k) -bad for color c . $\square$

We now prove Lemma 7.5.4, which is the counterpart of Lemma 7.4.4. There is however an additional level of difficulty: the probabilistic algorithm given before Lemma 7.5.3 requires making n iterated queries to the $\emptyset'$ -escaping function, asking each time to give an answer smaller than i .

Lemma 7.5.4. *Let $f : [\mathbb{N}]^2 \rightarrow 2$ be a stable coloring, $c < 2$ some color and $n, k \in \mathbb{N}$. Let $(F_i)_{i \in \mathbb{N}}$ be a computable family of finite subsets of $\mathbb{N}$ such that, for every $i \in \mathbb{N}$:*

- F_i f -realizes the i -fractal of dimension n .
- F_i is (f, k) -good for color c .

Then, every $\emptyset'$ -DNC degree $\mathbf{d}$ computes an infinite f -homogeneous set for color c .

Proof. Case 1: $\mathbf{d} \geq_T \emptyset'$. Let $i \geq k$. By Lemma 7.5.3, every set F that is (f, k) -good for color c and f -realizes the i -fractal of dimension d for some $d > 0$ contains at least $i - k - 1$ blocks that are (f, k) -good for color c and f -realize the i -fractal of dimension $d - 1$. Thus, by iterating, every such set F contains at least $(i - k - 1)^d$ disjoint subsets that f -realize the i -fractal of dimension 0 and are (f, k) -good for color c , hence F contains at least $(i - k - 1)^d$ elements having limit c .

Thus, under our assumption, there exists an infinite number of elements having limit c . These elements are $\emptyset'$ -computable, and from them we can compute an infinite f -homogeneous set for color c .

Case 2: $\mathbf{d} \not\geq_T \emptyset'$. Let $g \leq_T \mathbf{d}$ be the function of Lemma 7.1.9(2) and let μ_s be the standard left-c.e. approximation of the modulus μ of $\emptyset'$, as defined in Lemma 7.4.4. For every $x \in \mathbb{N}$, the sequence $(\mu_s(x))_{s \in \mathbb{N}}$ can only change values at most x times (every time a program with Turing index less than x stops), and let $m_0^x < \dots < m_i^x (= \mu(x))$ be those values for some $i < x$.

For every $\ell < n$, and $(w_0, \dots, w_{\ell-1}) \in \mathbb{N}^\ell$, let $W_{e(x, \langle w_0, \dots, w_{\ell-1} \rangle)}^{\emptyset'}$ be the $\emptyset'$ -c.e. set containing $\{0, \dots, x - 1\}$ and the indices of the (f, k) -bad blocks in the fractal decomposition of $F_{m_j^x}(\langle w_0, \dots, w_{\ell-1} \rangle)$ for every $j < x$.

For every $\ell < n$, let w_ℓ be the g -computable function defined inductively as follows:

$$w_\ell(x) = \begin{cases} g(e(x, \langle \rangle), k \times x) & \text{if } \ell = 0 \\ g(e(x, \langle w_0(x), \dots, w_{\ell-1}(x) \rangle), k \times x) & \text{if } \ell > 0 \end{cases}$$

Claim: For every $x, j \in \mathbb{N}$ such that $\max\{w_\ell(x) : \ell < n\} < m_j^x$ and every $\ell \leq n$,

(a) $F_{m_j^x}(\langle w_0(x), \dots, w_{\ell-1}(x) \rangle)$ is (f, k) -good for color c ;

(b) $\text{card } W_{e(x, \langle w_0(x), \dots, w_{\ell-1}(x) \rangle)}^{\emptyset'} \leq k \times x$

We proceed by induction on ℓ . If $\ell = 0$, (a) is due to the fact that every F_i is (f, k) -good for color c , (b) follows from Lemma 7.5.3. Then, assuming the property to be true for some $\ell < n$, the set $F_{m_j^x}(\langle w_0(x), \dots, w_{\ell-1}(x) \rangle)$ is therefore an m_j^x -fractal of dimension $n - \ell$ which is (f, k) -good for color c . By Lemma 7.5.3, there are at most $k - 1$ blocks in the fractal decomposition of $F_{m_j^x}(\langle w_0(x), \dots, w_{\ell-1}(x) \rangle)$ that are (f, k) -bad for color c . The indices of these blocks are the ones output by $W_{e(x, \langle w_0(x), \dots, w_{\ell-1}(x) \rangle)}^{\emptyset'}$ for every $j < x$ such that m_j^x is defined. Hence $\text{card } W_{e(x, \langle w_0(x), \dots, w_{\ell-1}(x) \rangle)}^{\emptyset'} \leq (k -$

1) $\times x + x = k \times x$ and (b) holds. Thus, by definition of g , $w_\ell(x)$ does not belong to $W_{e(x, \langle w_0(x), \dots, w_{\ell-1}(x) \rangle)}^{\emptyset'}$ and $F_{m_j^x}(\langle w_0(x), \dots, w_{\ell-1}(x), w_\ell(x) \rangle)$ is an m_j^x -fractal of dimension $n - \ell - 1$ which is (f, k) -good for color c , so (a) holds. This proves our claim.

Letting $p_x = \max\{w_\ell(x) : \ell < n\}$, the sequence $(p_x)_{x \in \mathbb{N}}$ is $\mathbf{d}$ -computable. Since $\mathbf{d} \not\geq_T \emptyset'$, there is no bound $b \in \mathbb{N}$ such that $p_x > \mu(x)$ for every $x > b$, as otherwise $\mathbf{d}$ would compute a function dominating μ and therefore compute $\emptyset'$.

Hence, for infinitely many x , we can $\mathbf{d}$ -computably find some $j \in \mathbb{N}$ such that $w_\ell(x) < m_j^x$ for every $\ell < n$. Once such x, j are found, we get by our claim that $F_{m_j^x}(\langle w_0(x), \dots, w_{n-1}(x) \rangle)$ is an m_j^x -fractal of dimension 0 that is (f, k) -good for color c . Thus, we can $\mathbf{d}$ -compute an infinite set of elements having limit c , hence $\mathbf{d}$ computes an infinite f -homogeneous set for color c . $\square$

We can now prove our main proposition in the case of k -fractals.

Proposition 7.5.5. *Let $f : [\mathbb{N}]^2 \rightarrow 2$ be a stable computable coloring. For every $n, k \in \mathbb{N}$, if f avoids the k -fractal of dimension n , then every $\emptyset'$ -DNC degree $\mathbf{d}$ computes an infinite f -homogeneous set.*

Proof. Let $\mathbf{d}$ be an $\emptyset'$ -DNC degree.

By induction on n , we show that for every $k \in \mathbb{N}$ and every stable computable coloring $f : [\mathbb{N}]^2 \rightarrow 2$ that avoids the k -fractal of dimension n , $\mathbf{d}$ computes an infinite f -homogeneous set.

The case $n = 0$ is vacuously true, as no set f -avoids the singleton pattern. The case $n = 1$ is true by Lemma 7.4.2.

Now assume the property holds for some $n \in \mathbb{N}$ and fix $k \in \mathbb{N}$. Let $f : [\mathbb{N}]^2 \rightarrow 2$ be a stable coloring avoiding the k -fractal of dimension $(n + 1)$. Let c be equal to 1 if n is even and 0 if n is odd.

Suppose first that there is some $i, b \in \mathbb{N}$ such that no finite set $F > b$ f -realizes the i -fractal of dimension n . Define $g : [\mathbb{N}]^2 \rightarrow 2$ by $g(x, y) = f(x + b, y + b)$. Then g is stable, computable and avoids the i -fractal of dimension n . By the inductive hypothesis, $\mathbf{d}$ computes an infinite g -homogeneous set, and hence computes an infinite f -homogeneous set.

Thus, assume that for every $i, b \in \mathbb{N}$, there exists some finite set $F > b$ that f -realizes the i -fractal of dimension n . We distinguish two cases:

Case 1: There exists some bound $b \in \mathbb{N}$, such that every $F > b$ that f -realizes the k -fractal of dimension n contains an element whose limit is c .

In that case, for every $i \in \mathbb{N}$, every $F > b$ that f -realizes the i -fractal of dimension n is (f, k) -good for color c , as every subset $G \subseteq F$ f -realizing the k -fractal of dimension n contains an element whose limit is c . By our assumption, such sets F exist for every $i \in \mathbb{N}$. It then follows from Lemma 7.5.4 that $\mathbf{d}$ computes an infinite f -homogeneous set.

Case 2: For every bound $b \in \mathbb{N}$, there exists $F > b$ that f -realizes the k -fractal of dimension n and whose elements all have limit $1 - c$.

In that case, construct a sequence $F_0 \cup \dots \cup F_{k-1}$ as follows. Let $F_0 \subseteq \mathbb{N}$ f -realizing the k -fractal of dimension n , with all its elements having limit $1 - c$, and let $b_0 \in \mathbb{N}$ be large enough that every element of F_0 has reached its limit by stage b_0 . Next, let $F_1 > b_0$ with the same properties, and let $b_1 > b_0$ witness stabilization for all elements of F_1 . Continuing this way, we obtain a sequence $F_0 < \dots < F_{k-1}$ such that $F_0 \cup \dots \cup F_{k-1}$ f -realizes the k -fractal of dimension $n + 1$, contradicting our assumption on f . $\square$

Theorem 7.5.6. *For every separable permutation p and every computable instance of $\text{coSRT}_2^2(p)$, every $\emptyset'$ -DNC degree computes a solution.*

Proof. Let $f : [\mathbb{N}]^2 \rightarrow 2$ be a stable computable coloring avoiding p . By Proposition 7.3.5, there exists some $k, n \in \mathbb{N}$ such that it avoids the k -fractal of dimension n . Then by Proposition 7.5.5 every $\emptyset'$ -DNC degree computes an infinite f -homogeneous set. $\square$

Theorem 7.5.7. *For every separable permutation p , $\text{RCA}_0 + \text{RT}_2^2(p)$ implies RT_2^2 over ω -models.*

Proof. Let $\mathcal{M} = (\mathbb{N}, S)$ be an ω -model of $\text{RT}_2^2(p)$. By Proposition 7.2.2, $\mathcal{M} \models \text{ADS}$ and thus $\mathcal{M} \models \text{COH}$.

Let $f : [\mathbb{N}]^2 \rightarrow 2$ be an instance of $\text{coSRT}_2^2(p)$ in S . By a relativization of Mimouni and Patey [102, Main theorem 1.2], there exists an f -computable coloring $g : [\mathbb{N}]^2 \rightarrow 2$ such that for every infinite set $H \subseteq \mathbb{N}$ that g -avoids any pattern, $f \oplus H$ computes an f' -DNC function.

As $\mathcal{M} \models \text{RCA}_0 + \text{RT}_2^2(p)$, there exists an infinite set $H \in S$ that g -avoids the pattern p . Hence, there exists a set G of f' -DNC degree in S .

By a relativization of Theorem 7.5.6, $G \oplus f$ computes an infinite f -homogeneous set. Thus, every instance of $\text{coSRT}_2^2(p)$ in S admits a solution in S , showing that $\mathcal{M} \models \text{coSRT}_2^2(p)$.

Thus, $\mathcal{M} \models \text{RT}_2^2(p) + \text{COH} + \text{coSRT}_2^2(p)$, and as $\text{RCA}_0 \vdash \text{RT}_2^2 \leftrightarrow \text{RT}_2^2(p) + \text{COH} + \text{coSRT}_2^2(p)$, we have $\mathcal{M} \models \text{RT}_2^2$ as required. $\square$

We are now ready to prove our main theorem.

Main Theorem 7.1.6. *For every pattern p , $\text{RT}_2^2(p)$ implies RT_2^2 over ω -models if and only if p is a separable permutation.*

Proof. If p is a separable permutation, then $\text{RT}_2^2(p)$ implies RT_2^2 over ω -models by Theorem 7.5.7. If p is not a separable permutation, then by Mimouni and Patey [102, Main Theorem 1.8], $\text{RT}_2^2(p)$ does not imply RT_2^2 over ω -models. $\square$

Note that the bound of Theorem 7.5.6 is not optimal, as for every separable permutation p , every computable instance of $\text{coSRT}_2^2(p)$ admits an $\emptyset'$ -computable solution, while there is no such $\emptyset'$ -DNC function. Moreover, for the case $p = 120$, Cervelle, Gaudelier, and Patey [16] proved that ADS implies $\text{coRT}_2^2(120)$ over RCA_0 , and by Hirschfeldt and Shore [63], ADS does not even imply the existence of a DNC function over ω -models.

7.6 Open questions

Many questions concerning the computability-theoretic and proof-theoretic strength of the statements $\text{RT}_2^2(p)$ and $\text{coRT}_2^2(p)$ remain open. The known relations over ω -models of RCA_0 are given in Figure 7.6.

The most fundamental question is about the general structure of this partial order.

Question 7.6.1. What is the structure of the Ramsey-like statements, ordered by implication over RCA_0 ? over ω -models?

Over ω -models, if $\text{RT}_2^2(p)$ is strictly weaker than RT_2^2 , then by Theorem 7.1.6, p is not a separable permutation, so either it is not a permutation, in which case $\text{RT}_2^2(p)$ follows from EM , or it is a permutation, but not separable, and by Theorem 7.1.4 it follows from $\text{RT}_2^2(1302)$. By Mimouni and Patey [102], $\text{RT}_2^2(1302)$ does not imply EM over ω -models, but the other direction is unknown.

Question 7.6.2. Does EM imply $\text{RT}_2^2(1302)$ over ω -models?

The main theorem of this chapter states that $\text{RT}_2^2(p)$ implies RT_2^2 over ω -models whenever p is a separable permutation. The proof, however, seems to make a strong use of ω -models, in particular with Lemma 7.4.2 which uses Σ_2 -induction.

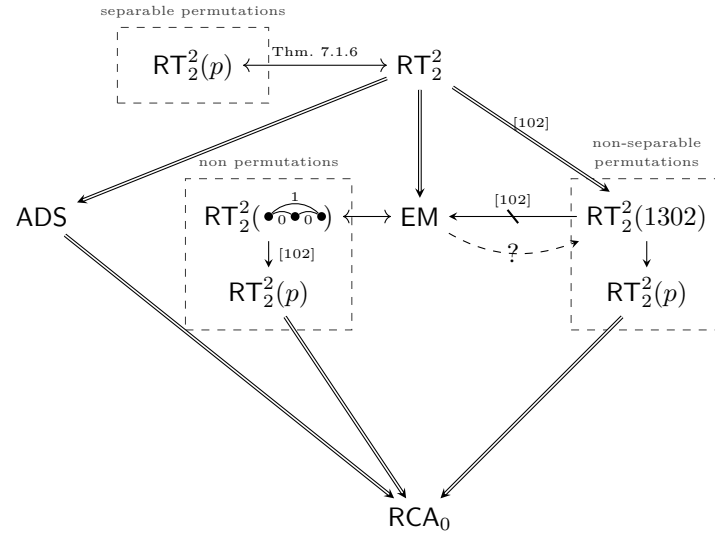


Figure 7.1: Implications over ω -models of RCA_0 . A simple (double) arrow represents a (strict) implication. A dashed arrow with a question mark denotes that the implication is still open and a crossed arrow represents a non-implication. Note that by [102], there exist non-permutations p such that $\text{RT}_2^2(p)$ is strictly weaker than EM over ω -models.

Question 7.6.3. For every separable permutation p , does $\text{RCA}_0 \vdash \text{RT}_2^2(p) \rightarrow \text{RT}_2^2$?

Chapter 8

Ramsey-like theorems for the Schreier barrier

8.1 Introduction

Among the variants of Ramsey's theorem whose computational and proof-theoretic power have been investigated, the free set, thin set and rainbow Ramsey theorems have attracted significant interest in recent decades (see, e.g., [18, 140, 113, 17, 98]), due to the peculiar behavior of these theorems when compared to Ramsey's theorem. For these three theorems, the instances are again colorings of the set $[\mathbb{N}]^n$, but this time with infinitely many colors allowed.

The following thin set and free set theorems were introduced by Harvey Friedman [47] and first studied in [18]. The thin set theorem is a weakening of Ramsey's theorem, where, instead of starting from a k -coloring of $[\mathbb{N}]^n$ and seeking to avoid all colors except one, one considers a coloring of $[\mathbb{N}]^n$ with infinitely many colors and aims to avoid just a single one.

Statement 8.1.1 (Thin set theorem). Let $n > 0$, TS^n is the statement “For every coloring $f : [\mathbb{N}]^n \rightarrow \mathbb{N}$, there exists an infinite set $H \subseteq \mathbb{N}$ such that $f([H]^n) \neq \mathbb{N}$ ”. Such a set H is called *thin* for f (or f -thin).

Statement 8.1.2 (Free set theorem). Let $n > 0$, FS^n is the statement “For every coloring $f : [\mathbb{N}]^n \rightarrow \mathbb{N}$, there exists an infinite set $H \subseteq \mathbb{N}$ such that for every $s \in [H]^n$, if $f(s) \in H$ then $f(s) \in s$ ”. Such a set H is called *free* for f (or f -free).

The notion of free set might seem ad hoc at first sight, but can be better understood in the light of the thin set theorem. Indeed, an infinite set $H \subseteq \mathbb{N}$ is f -free if and only if, for every $x \in H$, the set $H \setminus \{x\}$ is f -thin with witness color x (in the sense that f never takes value x on $[H \setminus \{x\}]^n$). Thus, the free set theorem is a natural generalization of the thin set theorem.

The rainbow Ramsey theorem is sometimes called an anti-Ramsey theorem as it requires the coloring f to be injective on the resulting set $[H]^n$ instead of monochromatic.

Statement 8.1.3 (Rainbow Ramsey theorem). Let $n, k > 0$, RRT_k^n is the statement “For every coloring $f : [\mathbb{N}]^n \rightarrow \mathbb{N}$ such that every color is used at most k times, there exists an infinite set $H \subseteq \mathbb{N}$ such that f is injective on $[H]^n$ ”. We then let RRT^n be the statement $\forall k \text{RRT}_k^n$. Such a coloring f is said to be k -bounded, and such a set H is called a *rainbow* for f (or f -rainbow).

While the hierarchy $RT^1, RT^2, RT^3, \dots$ collapses at the third level, where every RT^n for $n \geq 3$ already implies the existence of $\emptyset'$, and is equivalent to ACA_0 modulo RCA_0 , this is not the case anymore for these three principles. This surprising result is due to Wang [140], who showed that the free set, thin set and rainbow Ramsey theorems do not code the halting set and satisfy the cone avoidance property.

Definition 8.1.4. A problem P admits *cone avoidance* if for every set Z , every non- Z -computable set C and every Z -computable P -instance X , there exists a P -solution Y to X such that $C \not\leq_T Y \oplus Z$.

Proving that a problem P admits the cone avoidance property is the standard method for showing that it does not imply ACA_0 over RCA_0 . We will present another method in Chapter 15, based on the study of the first-order part of problems, to show that a statement about colorings of infinite trees, namely the ordered variable word theorem, does not imply ACA_0 over RCA_0 .

We already mentioned the cone avoidance basis theorem (Theorem 4.2.5) of Jockusch and Soare [71] that demonstrates that WKL has this property. In the case of RT_2^2 , the result was established by Seetapun and Slaman [124].

Proposition 8.1.5. *If a Π_2^1 problem P admits cone avoidance, then $RCA_0 + P \not\vdash ACA_0$. In fact, this non-implication already holds over ω -models.*

Proof. We aim to construct an ω -model $\mathcal{M} = (\mathbb{N}, S)$ of $RCA_0 + P + \neg ACA_0$. Since ACA_0 proves the existence of all arithmetical sets, it suffices that $\emptyset' \notin S$ in order for $\mathcal{M}$ not to be a model of ACA_0 .

For $\mathcal{M}$ to be a model of RCA_0 , we require S to be a Turing ideal, that is, closed under Turing join $\oplus$ and downward-closed under Turing reducibility $\leq_T$. Moreover, for $\mathcal{M}$ to be a model of P , every instance of P in S must admit a solution within S .

To achieve this, we construct an infinite sequence $Y_0, Y_1, \dots$ of subsets of $\mathbb{N}$ such that:

- For every $n \in \mathbb{N}$, $Y_0 \oplus \dots \oplus Y_{n-1} \not\leq_T \emptyset'$
- Whenever $Y_0 \oplus \dots \oplus Y_{n-1}$ computes an instance of P , there exists some $m \in \mathbb{N}$ such that Y_m is a solution to that instance.

We then define S to consist of all subsets of $\mathbb{N}$ that are computable from $Y_0 \oplus \dots \oplus Y_{n-1}$ for some $n \in \mathbb{N}$. By construction, S is a Turing ideal. Furthermore, the conditions imposed on the sequence (Y_i) ensure that $\emptyset' \notin S$ and that every instance of P in S has a solution in S .

The construction of the Y_i proceeds as follows: fix $n \in \mathbb{N}$ and assume $Y_0, \dots, Y_{n-1}$ have already been defined so that $Y_0 \oplus \dots \oplus Y_{n-1} \not\leq_T \emptyset'$. Write $n = \langle a, b \rangle$ using a coding of the pairs of integers. Since there are only countably many sets computable from $Y_0 \oplus \dots \oplus Y_b$, there are countably many $Y_0 \oplus \dots \oplus Y_b$ -instances of P . Let X_n denote the a -th such instance under some fixed enumeration.

We have $Y_0 \oplus \dots \oplus Y_{n-1} \not\leq_T \emptyset'$ and $Y_0 \oplus \dots \oplus Y_{n-1} \geq_T X_n$, thus, by the cone avoidance of P , there exists a P -solution Y_n to X_n such that $Y_0 \oplus \dots \oplus Y_n \not\leq_T \emptyset'$. This completes the inductive step and allows the construction to proceed. $\square$

The cone avoidance property reflects the computational weakness of a problem P , as it ensures that solving its instances does not enable one to compute any fixed non-computable set. In fact, Wang [140] showed that FS^n , TS^n and RRT^n satisfy even stronger forms of computational weakness.

Definition 8.1.6. A problem P admits *strong cone avoidance* if for every set Z , every non- Z -computable set C and every P -instance X , there exists a P -solution Y to X such that $C \not\leq_T Y \oplus Z$.

Theorem 8.1.7 (Wang [140]). *For every $n \in \mathbb{N}$, FS^n , TS^n and RRT^n admit strong cone avoidance.*

In the definition of the strong cone avoidance, no computability constraint is given on the P-instance X which can thus be arbitrarily complex compared to Z and in particular might compute C . Thus, the strong cone avoidance reflects the combinatorial weakness of P, in the sense that no matter how complex the instance is, it cannot code in its solutions an infinite binary sequence.

Neither RT_2^2 nor WKL admits the strong cone avoidance property. For WKL, consider the $\emptyset'$ -computable tree whose only path is $\emptyset'$. For RT_2^2 , let $g : \mathbb{N} \rightarrow \mathbb{N}$ be a fast-growing function (such as the modulus of $\emptyset'$) and define $f : [\mathbb{N}]^2 \rightarrow 2$ by $f(x, y) = 0$ if and only if $y < g(x)$. Then, every infinite f -homogeneous set $H = \{x_0 < x_1 < \dots\}$ must have color 1, and moreover satisfies $x_{i+1} > g(i)$. It follows that if g is the modulus of $\emptyset'$, then any solution to that $\emptyset'$ -instance of RT_2^2 computes $\emptyset'$.

In this chapter, we consider generalizations of the free set, thin set and rainbow Ramsey theorems to colorings of objects of unbounded dimension. We show that, among these, only the extension of RRT^n to colorings of so-called exactly ω -large sets still satisfies the strong cone avoidance property. By contrast, for the corresponding generalizations of TS^n and FS^n , this property fails dramatically, as these principles imply the existence of the ω -jump $\emptyset^{(\omega)}$.

8.2 Colorings of all finite sets

For $A \subseteq \mathbb{N}$, we let $[A]^{<\mathbb{N}}$ denote the set of all finite subsets of A . One should not confuse this with $A^{<\mathbb{N}}$, the set of all finite sequences of elements of A .

When generalizing Ramsey-like statements to colorings of finite sets of arbitrarily large sizes, one must proceed with caution, as these kinds of statements can easily be false. For example, consider the statement $\text{RT}^{<\mathbb{N}}$ stating that “For $k \in \mathbb{N}$ and every coloring $f : [\mathbb{N}]^{<\mathbb{N}} \rightarrow k$ of the finite subsets of $\mathbb{N}$, there exists an infinite set H such that f is monochromatic on $[H]^{<\mathbb{N}}$ ”. Then $\text{RT}^{<\mathbb{N}}$ is false, as witnessed by the coloring $f : [\mathbb{N}]^{<\mathbb{N}} \rightarrow 2$ which maps s to the parity of $\text{card } s$. Indeed, every infinite set contains finite subsets of even and odd cardinalities.

Similarly, the generalizations $\text{RRT}_k^{<\mathbb{N}}$, $\text{TS}^{<\mathbb{N}}$ and $\text{FS}^{<\mathbb{N}}$ of the rainbow Ramsey theorem, free set theorem and thin set theorem to colorings of $[\mathbb{N}]^{<\mathbb{N}}$ do not hold.

Proposition 8.2.1. *There exists a coloring $f : [\mathbb{N}]^{<\mathbb{N}} \rightarrow \mathbb{N}$ such that for every infinite set X , $f([X]^{<\mathbb{N}}) = \mathbb{N}$ (i.e., $\text{TS}^{<\mathbb{N}}$ is false).*

Proof. Let $f : [\mathbb{N}]^{<\mathbb{N}} \rightarrow \mathbb{N}$ be defined by setting $f(s) = \text{card } s$. Let X be an infinite subset of $\mathbb{N}$. Then $f([X]^{<\mathbb{N}}) = \mathbb{N}$. $\square$

As $\text{FS}^{<\mathbb{N}}$ can be seen as a generalization of $\text{TS}^{<\mathbb{N}}$, it is also a false statement.

For $\text{RRT}^{<\mathbb{N}}$, we will make use of the following class of finite sets. These sets will then play a central role in this chapter and later in Part IV for different reasons.

Definition 8.2.2 ((Exactly) ω -large sets). A finite $s \subseteq \mathbb{N}$ is ω -large if $\text{card } s \geq 1 + \min s$ and is exactly ω -large if $\text{card } s = 1 + \min s$. For $A \subseteq \mathbb{N}$, we write $[A]^{! \omega}$ to denote the exactly ω -large subsets of A .

Proposition 8.2.3. *There is a 2-bounded coloring $f : [\mathbb{N}]^{<\mathbb{N}} \rightarrow \mathbb{N}$ such that for every infinite set $H \subseteq \mathbb{N}$, f is not injective on $[H]^{<\mathbb{N}}$ (that is, $\text{RRT}_2^{<\mathbb{N}}$ is false).*

Proof. A finite set $s \subseteq \mathbb{N}$ is called *quasi-exactly ω -large* if there exists an exactly ω -large set t such that $s = t \setminus \{\min t\}$. Obviously if t is exactly ω -large, then $t \setminus \{\min t\}$ is quasi-exactly ω -large. Also, it is easy to see that if s is quasi-exactly ω -large then

there exists a unique exactly ω -large t such that $s = t \setminus \{\min t\}$. Let $b : [\mathbb{N}]^{<\mathbb{N}} \rightarrow \mathbb{N}$ be a bijection. We define $f : [\mathbb{N}]^{<\mathbb{N}} \rightarrow \mathbb{N}$ as follows. If s is exactly ω -large then let $f(s) = b(s \setminus \{\min s\})$, and otherwise, let $f(s) = b(s)$. Let $H \subseteq \mathbb{N}$ be infinite. Then for any $s \in [H]^{! \omega}$ we have $s \setminus \{\min s\} \in [H]^{<\mathbb{N}}$. But for any such s we have $f(s) = b(s \setminus \{\min s\}) = f(s \setminus \{\min s\})$. Thus H is not a rainbow for f . $\square$

In every case, the failure of the generalization comes from the fact that, for every large enough finite set $F \subseteq \mathbb{N}$, the colors of the subsets of F make it impossible to build a solution. For $\text{TS}^{<\mathbb{N}}$, our coloring ensures that every set F of cardinality $\geq x$ contains a subset of cardinality x making the color x unavoidable. For $\text{RRT}^{<\mathbb{N}}$, every exactly ω -large set contains a quasi-exactly ω -large subset whose color is the same.

To avoid this issue, we restrict ourselves to generalizations involving colorings $f : \mathcal{C} \rightarrow \mathbb{N}$, where $\mathcal{C} \subseteq [\mathbb{N}]^{<\mathbb{N}}$ is such that, for every $F, G \in \mathcal{C}$, if $F \subseteq G$, then $F = G$. In other words, $\mathcal{C}$ is an antichain under inclusion. This is the case for the family of the exactly ω -large sets.

The ω -large sets are also known as *relatively large* sets in the literature. The family $[\mathbb{N}]^{! \omega}$ coincides with the famous Schreier barrier used in better quasi ordering theory and Banach space theory [138].

The concept of ω -large finite subset of the natural numbers is well-known in the proof theory of arithmetic, as it is the basic ingredient for the celebrated Paris-Harrington independence result for Peano arithmetic [108]. It is worth noting that while Paris and Harrington characterized these sets using the condition $\text{card } s \geq \min s$, this variation is essentially equivalent to the one considered here.

The following is the generalization of Ramsey's theorem to colorings of exactly ω -large sets. It appears as a special case of more general theorems by Pudlák and Rödl [118] and by Farmaki and Negrepontis [43]. Besides, it is a particular case of Nash-Williams' generalization of Ramsey's theorem to barriers, which in turn is a consequence of the clopen Ramsey theorem (see [128]).

Statement 8.2.4 (Large Ramsey theorem). Let $k > 0$. For every coloring $f : [\mathbb{N}]^{! \omega} \rightarrow k$, there exists an infinite set $H \subseteq \mathbb{N}$ such that f is constant on $[H]^{! \omega}$. Following [15], we abbreviate this statement by $\text{RT}_k^{! \omega}$.

This principle is arguably the simplest example of a true version of Ramsey's theorem for colorings of objects of *unbounded dimensions*. Formulated this way, this theorem can arguably be considered as a non-robust statement, in the sense that even simple variations of it need not remain equivalent to it over RCA_0 . For example, consider the variant in which, instead of coloring $[\mathbb{N}]^{! \omega}$, one colors $[X]^{! \omega}$ for some infinite set $X \subseteq \mathbb{N}$. In the case of RT_k^n , we can use the canonical bijection between X and $\mathbb{N}$ to transform a coloring of $[X]^n$ into one of $[\mathbb{N}]^n$. However, this argument no longer applies to colorings of the exactly ω -large sets, since the image of such a set under the canonical bijection need not itself be exactly ω -large.

A first solution, adopted by Carlucci and Zdanowski [15], consisted in directly studying the theorem in its general formulation, where an instance is a pair (X, f) with X an infinite subset of $\mathbb{N}$ and f a k -coloring of $[X]^{! \omega}$. However, making the domain part of the instance raises some issues when considering strong cone avoidance as we do in Section 8.5 for the generalization of the rainbow Ramsey theorem. Indeed, one can for example consider an $\emptyset'$ -computable instance (X, f) where X is sufficiently sparse that every infinite subset of X computes $\emptyset'$ (for instance, by letting the principal function p_X of X dominate the modulus of $\emptyset'$), and where f is an arbitrary coloring. In this case, every solution to the instance computes $\emptyset'$, but for reasons unrelated to the combinatorial complexity of the problem.

A different approach was therefore considered in [14], where the strong cone avoidance result was obtained on a generalized version formulated in the language of barriers. This generalization can be understood by observing that the property

$\text{card } s = 1 + \min s$ in the definition of an exactly ω -large set can be replaced by $\text{card } s = h(\min s)$, where $h : \mathbb{N} \rightarrow \mathbb{N}$ is a computable increasing function, without significantly affecting the underlying combinatorics. Crucially, this reformulation is preserved under bijections between an infinite set X and $\mathbb{N}$, at the cost of adjusting the bound function h .

The large Ramsey theorem has been studied from the perspective of computability theory and reverse mathematics by Carlucci and Zdanowski [15], and its generalization to barriers in computability theory by Clote [26]. They proved that it is computationally and proof-theoretically stronger than the usual Ramsey's theorem for each fixed finite dimension (RT^n) and even stronger than Ramsey's theorem for all finite dimensions ($\forall n \text{RT}^n$). From the point of view of computability, the theorem corresponds to the ω -th Turing jump $\mathcal{O}^{(\omega)}$; in reverse mathematics terms, it is equivalent to the theory ACA_0^+ over RCA_0 , where ACA_0^+ extends RCA_0 by the axiom of closure under the ω -th Turing jump.

The following theorem summarizes the known bounds for $\text{RT}_k^{l\omega}$.

Theorem 8.2.5 (Carlucci and Zdanowski [15] and Clote [26]).

1. All computable finite colorings of $[\mathbb{N}]^{l\omega}$ admit an infinite monochromatic set computable in $\mathcal{O}^{(\omega)}$.
2. There exists a computable coloring of $[\mathbb{N}]^{l\omega}$ in 2 colors such that all infinite monochromatic sets compute $\mathcal{O}^{(\omega)}$.
3. $\text{RT}_2^{l\omega}$ is equivalent to ACA_0^+ over RCA_0 .

As stated before, the ω -large sets will feature prominently in Part IV. Rather than using them to measure the size of the finite subsets of $\mathbb{N}$ being colored, we will use them to formulate variants of the finite Ramsey theorem where the resulting monochromatic set is required to be ω -large.

8.3 Colorings of exactly ω -large sets

In this section, we introduce and prove the generalizations of the free set, thin set and rainbow Ramsey theorems to colorings of exactly ω -large sets. We also establish some basic implications and relations among those principles.

Statement 8.3.1 (Large thin set theorem). $\text{TS}^{l\omega}$ is the statement “For every coloring $f : [\mathbb{N}]^{l\omega} \rightarrow \mathbb{N}$, there exists an infinite set $H \subseteq \mathbb{N}$ such that $f([H]^{l\omega}) \neq \mathbb{N}$ ”. Such a set H will be called f -thin.

Statement 8.3.2 (Large free set theorem). $\text{FS}^{l\omega}$ is the statement “For every coloring $f : [\mathbb{N}]^{l\omega} \rightarrow \mathbb{N}$, there exists an infinite set $H \subseteq \mathbb{N}$ such that for every $s \in [H]^{l\omega}$, if $f(s) \in H$ then $f(s) \in s$ ”. Such a set H will be called f -free.

Statement 8.3.3 (Large rainbow Ramsey theorem). Let $k > 0$, $\text{RRT}_k^{l\omega}$ is the statement “For every coloring $f : [\mathbb{N}]^{l\omega} \rightarrow \mathbb{N}$ such that every color is used at most k times, there exists an infinite set $H \subseteq \mathbb{N}$ such that f is injective on $[H]^{l\omega}$ ”. We then let $\text{RRT}^{l\omega}$ be the statement $\forall k \text{RRT}_k^{l\omega}$. Again, such a coloring f will be called k -bounded, and such a set H an f -rainbow.

It is quite natural to ask whether these generalizations hold. We will show that they do, by proving that they are implied by $\text{RT}^{l\omega}$.

When studying the computability-theoretic strength of various statements, we have so far restricted attention to their relative strength over ω -models of RCA_0 . To show that a statement P implies a statement Q over such models, it is sufficient to

show that a solution to any instance of Q can be given by solving a finite number of instances of P one after the other.

However, this is not the only computability-theoretic notion of reducibility between problems that one may consider. For example, we introduce the following notion of reducibility whose variants became of central interest in computability theory and reverse mathematics in recent years (see [39] for background and motivation).

Definition 8.3.4. Q is *strongly Weihrauch reducible* to P (denoted $Q \leq_{\text{SW}} P$) if there exist Turing functionals Φ and Ψ such that for every instance X of Q we have that Φ^X is an instance of P , and if Y is a solution to P for Φ^X then Ψ^Y is a solution to Q for X .

This form of reducibility is far more restrictive than implication over ω -models. First, only a single instance of P can be used. Moreover, the second functional Ψ has access only to the solution Y of the P -instance, and not to the original instance X itself. Finally, the reduction is more uniform in the sense that the Turing functionals Φ and Ψ are fixed in advance and cannot depend on the particular instance X or on the solution Y . This contrasts with implication over ω -models, where some statements like RT_2^1 and BRT_2^2 can be computably true, but such that no single Turing functional is able to solve all instances.

The more restrictive the notion of reducibility, the more properties can be transferred between the two statements. For example, we will use the fact that $\text{TS}^{\omega} \leq_{\text{SW}} \text{FS}^{\omega}$ to prove, in Section 8.4, the existence of a computable instance of FS^{ω} every solution of which computes $\mathcal{O}^{(\omega)}$, based on the existence of an instance of TS^{ω} with similar properties. Such a deduction would not be possible from a mere implication over ω -models.

We establish FS^{ω} by reduction to RT^{ω} . The proof combines ideas from the proof of FS^n from RT_{2n+2}^n (Theorem 5.2 and Corollary 5.3 in [18]) and from the proof of Theorem 4.1 in [15].

Theorem 8.3.5. $\text{FS}^{\omega} \leq_{\text{SW}} \text{RT}_2^{\omega}$.

Proof. For $s = \{s_0, \dots, s_{s_0}\}$ an exactly ω -large set with $s_0 > 0$, and $s_0 < \dots < s_{s_0}$ define $s \ominus 1$ to be the following exactly ω -large set: $\{s_0 - 1, s_1 - 1, \dots, s_{s_0-1} - 1\}$. The idea is to keep one degree of freedom: s_{s_0} , as in the proof of Carlucci and Zdanowski [15, Proposition 4.1].

Let $f : [\mathbb{N}]^{\omega} \rightarrow \mathbb{N}$ be an instance of FS^{ω} . Consider the following function $g : [\mathbb{N}]^{\omega} \rightarrow 2$ defined by induction. Note that the function g is called recursively on lexicographically smaller parameters, so the induction is well-defined since the lexicographic order is a well-order. For $s = \{s_0, \dots, s_{s_0}\}$ an exactly ω -large set:

$$g(s) = \begin{cases} 0 & \text{if } f(s \ominus 1) = s_i - 1 \text{ for some } i \leq s_0 \\ 1 - g(f(s \ominus 1) + 1, s_1, \dots, s_{f(s \ominus 1) + 1}) & \text{if } f(s \ominus 1) < s_0 - 1 \\ 1 - g(s_0, \dots, s_i, f(s \ominus 1) + 1, s_{i+2}, \dots, s_{s_0}) & \text{if } f(s \ominus 1) \in (s_i - 1, s_{i+1} - 1) \text{ for some } i < s_0 - 1 \\ 0 & \text{if } f(s \ominus 1) \in (s_{s_0-1} - 1, s_{s_0} - 1) \\ 1 & \text{otherwise (if } f(s \ominus 1) \geq s_{s_0}) \end{cases}$$

Let $H = \{x_0, x_1, \dots\} \subseteq \mathbb{N} \setminus \{0\}$ be an infinite g -homogeneous set as given by RT_2^{ω} . We claim that the set $G = \{x_0 - 1, x_1 - 1, \dots\}$ is f -free. Consider some exactly ω -large set $\{s_0 - 1, \dots, s_{s_0-1} - 1\} \subseteq G$. Then $\{s_0, \dots, s_{s_0-1}\} \subseteq H$. There are two cases:

Case 1: H is homogeneous for the color 0. Then, take s_{s_0} to be the next element of H after s_{s_0-1} and write $s = \{s_0, \dots, s_{s_0}\}$, then $g(s) = 0$. There are four subcases:

Subcase 1.1: $f(s \ominus 1) = s_i - 1$ for some $i \leq s_0$. In that case we are done.

Subcase 1.2: $f(s \ominus 1) \in (s_{s_0-1} - 1, s_{s_0} - 1)$. In that case, by definition of s_{s_0} , $f(s \ominus 1)$ is not in G .

Subcase 1.3: $f(s \ominus 1) < s_0 - 1$ and $g(f(s \ominus 1) + 1, s_1, \dots, s_{f(s \ominus 1)+1}) = 1$. If $f(s \ominus 1) \in G$ this contradicts the fact that H is g -homogeneous for the color 0.

Subcase 1.4: $f(s \ominus 1) \in (s_i - 1, s_{i+1} - 1)$ for some $i < s_0 - 1$ and $g(s_0, \dots, s_i, f(s \ominus 1) + 1, s_{i+2}, \dots, s_{s_0}) = 1$. If $f(s \ominus 1) \in G$ this contradicts the fact that H is g -homogeneous for the color 0.

Case 2: H is homogeneous for the color 1. Take $s_{s_0} \in H$ bigger than s_{s_0-1} and write $s = \{s_0, \dots, s_{s_0}\}$, then $g(s) = 1$. There are three subcases:

Subcase 2.1: $f(s \ominus 1) \geq s_{s_0}$. This case is impossible, indeed, as H is infinite, there exists some element $x \in H$ such that $x > f(s \ominus 1) + 1$ and therefore $f(s \ominus 1) \in (s_{s_0} - 1, x - 1)$, which leads to $g(s_0, \dots, s_{s_0-1}, x) = 0$ contradicting the fact that H is g -homogeneous for the color 1.

Subcase 2.2: $f(s \ominus 1) < s_0 - 1$ and $g(f(s \ominus 1) + 1, s_1, \dots, s_{f(s \ominus 1)+1}) = 0$. If $f(s \ominus 1) \in G$ this contradicts the fact that H is g -homogeneous for the color 1.

Subcase 2.3: $f(s \ominus 1) \in (s_i - 1, s_{i+1} - 1)$ for some $i < s_0 - 1$ and $g(s_0, \dots, s_i, f(s \ominus 1) + 1, s_{i+2}, \dots, s_{s_0}) = 0$. If $f(s \ominus 1) \in G$ this contradicts the fact that H is g -homogeneous for the color 1.

Notice that g is uniformly computable in f and that G is uniformly computable in H . Therefore, $\text{FS}^{\text{lw}} \leq_{\text{sW}} \text{RT}_2^{\text{lw}}$. □

The above proof uses the fact that the exactly ω -large sets are well-ordered under lexicographic ordering. Since the order type of this ordering is ω^ω and the statement “ ω^ω is well-ordered” implies the consistency of RCA_0 (by Theorem 3.6.3), the above proof is not formalizable in RCA_0 . Yet, since RT_2^{lw} implies ACA_0 and the latter proves that ω^ω is well-ordered we obtain the following corollary.

Corollary 8.3.6. $\text{RCA}_0 \vdash \text{RT}_2^{\text{lw}} \rightarrow \text{FS}^{\text{lw}}$.

Cholak et al. [18] have proved in RCA_0 that for all $k \geq 2$, FS^k implies TS^k (see Theorem 3.2 in [18]; the proof yields a strong Weihrauch reduction). A completely analogous argument establishes that TS^{lw} follows from (and is strongly Weihrauch reducible to) FS^{lw} .

To show that RRT^n holds, Galvin gave a reduction to Ramsey’s theorem. His argument is formalizable in RCA_0 and yields that for each $n, k \in \mathbb{N}$, $\text{RRT}_k^n \leq_{\text{sW}} \text{RT}_k^n$ (see Csima and Mileti [32, Theorem 5.2]). Csima and Mileti also showed that for each $n, k > 0$, RRT_k^n follows from RRT_k^{n+1} over RCA_0 (the proof of Theorem 5.3 in [32] yields a strong Weihrauch reduction).

We first observe that Galvin’s argument adapts to the case of colorings of exactly ω -large sets and establishes a strong Weihrauch reduction.

Theorem 8.3.7. For all $k > 0$, $\text{RRT}_k^{\text{lw}} \leq_{\text{sW}} \text{RT}_k^{\text{lw}}$. Moreover $\text{RCA}_0 \vdash \forall k (\text{RT}_k^{\text{lw}} \rightarrow \text{RRT}_k^{\text{lw}})$.

Proof. Let $f : [\mathbb{N}]^{\text{lw}} \rightarrow \mathbb{N}$ be a k -bounded coloring and fix a computable bijection $b(\cdot) : [\mathbb{N}]^{\text{lw}} \rightarrow \mathbb{N}$.

Define $g : [\mathbb{N}]^{\text{lw}} \rightarrow k$ as follows:

$$g(s) = \text{card}\{t \in [\mathbb{N}]^{\text{lw}} : b(t) < b(s) \text{ and } f(s) = f(t)\}.$$

The fact that g is a k -coloring depends on the hypothesis that f is k -bounded. Let H be an infinite set such that g is constant on $[H]^{\text{lw}}$, as given by RT_k^{lw} . Let $s, t \in [H]^{\text{lw}}$. Since $g(s) = g(t)$ and either $b(s) < b(t)$ or $b(t) < b(s)$ we have that $f(s) \neq f(t)$. Thus H is a rainbow for f . □

The following is an adaptation of a result by Wang [140], who showed that for every n , $\text{RRT}_2^n \leq_{\text{sW}} \text{FS}^n$.

Proposition 8.3.8. $\text{RRT}_2^{! \omega} \leq_{\text{sW}} \text{FS}^{! \omega}$.

Proof. Fix a computable bijection $b : [\mathbb{N}]^{! \omega} \rightarrow \mathbb{N}$. Let $f : [\mathbb{N}]^{! \omega} \rightarrow \mathbb{N}$ be 2-bounded. Define $g : [\mathbb{N}]^{! \omega} \rightarrow \mathbb{N}$ as follows:

$$g(s) = \begin{cases} \min(t \setminus s) & \text{if there is a } t \in [\mathbb{N}]^{! \omega} \text{ such that } b(t) < b(s) \text{ and } f(s) = f(t), \\ 0 & \text{otherwise.} \end{cases}$$

Since f is 2-bounded, if t exists in the definition of g then it is unique. If t and s are distinct exactly ω -large sets then $(t \setminus s) \neq \emptyset$, since $t \subseteq s$ is impossible. Let A be an infinite g -free set. We claim that A is a rainbow for f . Suppose otherwise, by way of contradiction, as witnessed by $s, t \in [A]^{! \omega}$ such that $f(s) = f(t)$. Without loss of generality we can assume $b(t) < b(s)$. Then $g(s) = \min(t \setminus s) \in A \setminus s$, contradicting that A is g -free. $\square$

Wang [140] proved that the free set, thin set and rainbow Ramsey theorems for fixed-sized sets satisfy cone avoidance. This entails that none of these principles codes the halting set or implies ACA_0 . A natural question is whether the same is true of their versions for exactly ω -large sets. We will show that $\text{FS}^{! \omega}$ and $\text{TS}^{! \omega}$ code $\emptyset^{(\omega)}$ and imply the much stronger system ACA_0^+ , while $\text{RRT}^{! \omega}$ admits cone avoidance.

Some first lower bounds on our principles can be obtained by adapting results from the finite-dimensional case. For example, the rainbow Ramsey theorem with internal quantification over all fixed finite dimensions $\forall n \text{RRT}_2^n$ follows from the rainbow Ramsey theorem for exactly ω -large sets.

Proposition 8.3.9. For each $k, n > 0$, $\text{RRT}_k^{! \omega} \geq_{\text{sW}} \text{RRT}_k^n$. Moreover, $\text{RCA}_0 \vdash \forall k (\text{RRT}_k^{! \omega} \rightarrow \forall n \text{RRT}_k^n)$.

Proof. Let $n > 0$ and let $f : [\mathbb{N}]^n \rightarrow \mathbb{N}$ be k -bounded. Define $g : [\mathbb{N}]^{! \omega} \rightarrow \mathbb{N}$ as follows. For an exactly ω -large set $s = \{s_0, \dots, s_{s_0}\}$,

$$g(s) = \begin{cases} \langle 0, s_0, f(s_1, \dots, s_n), s_{n+1}, \dots, s_{s_0} \rangle & \text{if } s_0 \geq n, \\ \langle 1, s \rangle & \text{if } s_0 < n, \end{cases}$$

where $\langle \cdot \rangle : \mathbb{N}^{< \mathbb{N}} \rightarrow \mathbb{N}$ is a fixed computable bijection.

Claim: g is k -bounded. Let $s = \{s_0, \dots, s_{s_0}\}$ be exactly ω -large. If $s_0 < n$ then there are no other exactly ω -large sets t such that $g(s) = g(t)$. If $s_0 \geq n$, suppose that some exactly ω -large set $t = \{t_0, \dots, t_{t_0}\}$ satisfies $g(s) = g(t)$. Then, necessarily $t_0 = s_0$, $s_i = t_i$ for $i > n$, and $f(s_1, \dots, s_n) = f(t_1, \dots, t_n)$, hence there are at most k such t by k -boundedness of f .

Now, let A be an infinite rainbow for g and define $B = A \cap [n, \infty)$, then $B \setminus \{\min B\}$ is an infinite rainbow for f . Indeed, let $s, t \in [B \setminus \{\min B\}]^n$ such that $f(s) = f(t)$. Since A is infinite, we can choose elements $x_{n+1} < \dots < x_{\min B}$ in A such that $x_{n+1} > \max\{\max s, \max t\}$. Define $s' = \{\min B\} \cup s \cup \{x_{n+1}, \dots, x_{\min B}\}$ and $t' = \{\min B\} \cup t \cup \{x_{n+1}, \dots, x_{\min B}\}$. Then s' and t' are exactly ω -large and $g(s') = g(t')$ by construction. Since A is a rainbow for g , this implies $s' = t'$, and therefore $s = t$. Hence, $B \setminus \{\min B\}$ is a rainbow for f . $\square$

Analogous results can be obtained for the large free set and the large thin set theorems, so that $\text{P}^{! \omega}$ implies $\forall n \text{P}^n$ over RCA_0 for $\text{P} \in \{\text{FS}, \text{TS}\}$. Both implications are also witnessed by strong Weihrauch reductions. We omit the proofs since these results are superseded by the results of the next section where we prove that $\text{FS}^{! \omega}$ and $\text{TS}^{! \omega}$ imply ACA_0^+ .

8.4 $\text{FS}^{! \omega}$ and $\text{TS}^{! \omega}$ code $\emptyset^{(\omega)}$

In this section we establish strong lower bounds on $\text{FS}^{! \omega}$ and $\text{TS}^{! \omega}$ showing that both these principles code $\emptyset^{(\omega)}$ and imply ACA_0^+ . This should be contrasted with the fact that neither $\forall n \text{FS}^n$ nor $\forall n \text{TS}^n$ implies ACA_0 .

The first goal is to prove the existence of a computable instance of $\text{TS}^{! \omega}$ such that every solution uniformly computes $\emptyset^{(\omega)}$. In particular, $\text{TS}^{! \omega}$ admits the same lower bound as $\text{RT}_2^{! \omega}$. Since $\text{TS}^{! \omega} \leq_{\text{SW}} \text{FS}^{! \omega}$, it follows that there is a computable instance of $\text{FS}^{! \omega}$ such that every solution computes $\emptyset^{(\omega)}$. By results in the previous section, this bound is optimal, since every computable instance of $\text{FS}^{! \omega}$ admits a solution computable in $\emptyset^{(\omega)}$ by Theorem 8.3.5 and Theorem 8.2.5.

One intuition behind this result is that, although TS^n admits strong cone avoidance for every $n \in \mathbb{N}$, this is not necessarily the case when, instead of trying to avoid one color in $\mathbb{N}$, one tries to avoid one out of a finite number of them. In fact, for every $k \in \mathbb{N}$, there exist finite colorings of $[\mathbb{N}]^n$ using increasingly many colors as n grows, such that one cannot avoid a single color without computing $\emptyset^{(k)}$. Our proof will combine these colorings into a single coloring $f : [\mathbb{N}]^{! \omega} \rightarrow \mathbb{N}$ such that every infinite f -thin set uniformly computes $\emptyset^{(k)}$, and thus computes $\emptyset^{(\omega)}$.

We therefore extend the definition of thinness to include finite codomains as follows:

Definition 8.4.1. Let C be any non-empty set. Given a coloring $f : [\mathbb{N}]^n \rightarrow C$, a set $H \subseteq \mathbb{N}$ is *f -thin for color $c \in C$* if $c \notin f([H]^n)$. A set $H \subseteq \mathbb{N}$ is *f -thin* (or *thin for f*) if H is f -thin for some color $c \in C$.

The definition of f -thin set depends on the choice of codomain of the function. It will always be clear from the context.

The following version of the thin set theorem for finite colorings was introduced in [36].

Statement 8.4.2 (Thin set theorem for finite colorings). Let $n, k > 0$ with $k \geq 2$. TS_k^n is the statement “For every coloring $f : [\mathbb{N}]^n \rightarrow k$, there exists an infinite set $H \subseteq \mathbb{N}$ such that H is thin for f .”

Dorais et al. [36, Proposition 5.5] proved the existence, for every $n > 0$, of a computable coloring $f : [\mathbb{N}]^{n+2} \rightarrow 2^n$ such that every infinite f -thin set computes $\emptyset'$. Wang [140] proved that for every $n \in \mathbb{N}$, and for large enough value of $k \in \mathbb{N}$, the statement TS_k^n admits strong cone avoidance. Cholak and Patey [17] were then able to give the exact threshold $k \in \mathbb{N}$ after which TS_k^n admits strong cone avoidance. Surprisingly, this threshold involved the following sequence of numbers, known as *Catalan numbers*, which are omnipresent in combinatorics, and inductively defined as follows:

$$C_0 = 1 \quad C_{n+1} = \sum_{i=0}^n C_i C_{n-i}$$

This sequence starts with 1, 1, 2, 5, 14, 42, ... (see sequence A000108 in the OEIS). The number C_n admits many characterizations, such as the number of ways of associating n applications of a binary operator.

Theorem 8.4.3 (Cholak and Patey [17]). *For every $n > 0$, TS_k^n admits strong cone avoidance if and only if $k > C_n$.*

Unfortunately, the results by Dorais et al. [36, Proposition 5.5] and Cholak and Patey [17] are not uniform: there is no single Turing functional capable of computing $\emptyset'$ from solutions to the instances of TS_k^n for the small values of k they constructed. Such uniformity is required for our construction in order to compute $\emptyset^{(\omega)}$. To achieve this, we instead work with even fewer colors and prove the existence, for every $n \geq 2$, of a computable coloring $f : [\mathbb{N}]^{n+1} \rightarrow n$ such that every infinite f -thin set uniformly computes $\emptyset'$.

Lemma 8.4.4. *There exist two computable arrays $(e_{n,k})_{n,k \in \mathbb{N}}$ and $(d_{n,k})_{n,k \in \mathbb{N}}$ of Turing indexes such that for every $n, k > 0$ with $n \geq 2$, $\Phi_{e_{n,k}}^{(k)}$ is a coloring $f_n^k : [\mathbb{N}]^n \rightarrow n$ such that for every infinite f_n^k -thin set H , $\Phi_{d_{n,k}}^H = \emptyset^{(k)}$.*

Proof. By generalizing the definition of the modulus of $\emptyset'$ from Chapter 4, we can show that for every $k > 0$, there exists a $\emptyset^{(k)}$ -computable function g_k such that any function dominating g_k computes $\emptyset^{(k)}$. Moreover, the Turing indices for the functionals computing g_k from the oracle $\emptyset^{(k)}$ can be found computably, uniformly in k .

For $n \geq 2$ let $f_n^k(x_0, \dots, x_{n-1}) = n-1$ if $g_k(x_0) \leq x_1$ and if $g_k(x_0) > x_1$, let $f_n^k(x_0, \dots, x_{n-1}) = i$ for $i < n-1$ the largest value such that $g_k(x_0) > x_{i+1}$.

Let H be an infinite f_n^k -thin set for some color c . The color c cannot be $n-1$ as for a given $x_0 \in H$ there exists $x_1 < \dots < x_{n-1} \in H \setminus \{0, \dots, g_k(x_0)\}$, hence $f_n^k(x_0, \dots, x_{n-1}) = n-1$.

We claim H is thin for color $n-2$. Indeed, assume by contradiction that there exists some tuple $x_0 < \dots < x_{n-1} \in H$ such that $g_k(x_0) > x_{n-1}$ so that f_n^k takes color $n-2$ on $\{x_0, \dots, x_{n-1}\}$. Since $c < n-2$ and H is infinite, there exists some $y_{c+2} < \dots < y_{n-1} \in H \setminus \{0, \dots, g_k(x_0)\}$, therefore $f_n^k(x_0, \dots, x_{c+1}, y_{c+2}, \dots, y_{n-1}) = c$, contradicting our assumption that H is f_n^k -thin for c .

Write $H = \{x_0 < x_1 < \dots\}$. For every $i \in \mathbb{N}$, $g_k(x_i) \leq x_{i+n-1}$, hence H computes a function dominating g_k and therefore computes $\emptyset^{(k)}$. This computation can be done uniformly in the set H , k and n . $\square$

Lemma 8.4.5. *There exists a functional Γ such that for every set X and every index e , if $\Phi_e^{X'}$ is a coloring $f : [\mathbb{N}]^n \rightarrow \ell$, then $\Gamma^X(e)$ is a coloring $g : [\mathbb{N}]^{n+1} \rightarrow \ell$ such that if an infinite set H is g -thin for a color c , then H is f -thin for the same color c .*

Proof. Consider a set X and an index e such that $\Phi_e^{X'}$ is a coloring $f : [\mathbb{N}]^n \rightarrow \ell$ for some $n, \ell > 0$. By Shoenfield's limit Lemma (see Proposition 4.1.7), we can consider an X -computable sequence $(f_s)_{s \in \mathbb{N}}$ of X -computable functions, such that, for every $F \in [\mathbb{N}]^n$, $\lim_{s \rightarrow \infty} f_s(F)$ exists and is equal to $f(F)$. Indices for such an approximation, where the f_s are seen as X -computable functions, can computably be found uniformly in X and e . Finally, consider the following X -computable coloring g defined by $g(x_0, \dots, x_n) = f_{x_n}(x_0, \dots, x_{n-1})$ (again, the construction is uniform).

Let $H \subseteq \mathbb{N}$ be an infinite set. If for some $x_0 < \dots < x_{n-1} \in H$ we have $f(x_0, \dots, x_{n-1}) = c$ for some color c , then $f_s(x_0, \dots, x_{n-1}) = c$ for every s bigger than a certain threshold. Thus, as H is infinite, there exists some $x_n \in H$ such that $g(x_0, \dots, x_n) = c$. $\square$

Combining these two lemmas, we get the following, where the array $(d_{n,k})_{n,k \in \mathbb{N}}$ is as in Lemma 8.4.4.

Lemma 8.4.6. *There exists a uniformly computable sequence of colorings $f_{n,k} : [\mathbb{N}]^{n+k} \rightarrow n$, for $n, k > 0$, such that for every infinite $f_{n,k}$ -thin set H , $\Phi_{d_{n,k}}^H = \emptyset^{(k)}$.*

The following theorem states that TS^{lw} does not admit cone avoidance in a strong sense: there exists a single computable instance of TS^{lw} that computes $\emptyset^{(\omega)}$.

Theorem 8.4.7. *There exists a computable coloring $f : [\mathbb{N}]^{\text{lw}} \rightarrow \mathbb{N}$ such that every infinite f -thin set H computes $\emptyset^{(\omega)}$. Moreover, this computation is uniform in H and an avoided color c .*

Proof. Let $f : [\mathbb{N}]^{\text{lw}} \rightarrow \mathbb{N}$ be defined as follows (where the functions $f_{n,k}$ are from Lemma 8.4.6): for $x_0 < \dots < x_{x_0}$ in $\mathbb{N}$ with $x_0 > 1$, let $f(x_0, \dots, x_{x_0}) = f_{x_0-k,k}(x_1, \dots, x_{x_0})$ for $k = \lceil \frac{x_0}{2} \rceil$; for $s \in [\mathbb{N}]^{\text{lw}}$ with $\min s \leq 1$ set $f(s) = 0$. By Lemma 8.4.6, f is indeed computable.

Let H be an infinite f -thin set for some color c and let $n \in \mathbb{N}$. Since H is infinite, there exists some $x_0 \in H$ such that, by letting $k = \lceil \frac{x_0}{2} \rceil$, we have $x_0 - k > c$ and

$k \geq n$. The set $G = H \setminus \{0, \dots, x_0\}$ is infinite and f -thin for c . By definition of f , G is $f_{x_0-k,k}$ -thin for the color c . Note that c is part of the range of $f_{x_0-k,k}$ as $x_0 - k > c$. By Lemma 8.4.6, since $k \geq n$, $G \geq_T \emptyset^{(n)}$. Since G is obviously H -computable, we have $H \geq_T \emptyset^{(n)}$. This computation can be done uniformly in H , c and n , thus $H \geq_T \emptyset^{(\omega)}$ uniformly in H and c . $\square$

Corollary 8.4.8. There exists a computable coloring $f : [\mathbb{N}]^{\text{!}\omega} \rightarrow \mathbb{N}$ such that every infinite f -free set computes $\emptyset^{(\omega)}$.

Proof. Immediate by Theorem 8.4.7 and the fact that $\text{TS}^{\text{!}\omega} \leq_{\text{SW}} \text{FS}^{\text{!}\omega}$. $\square$

Recall that ACA_0^+ is the theory consisting of RCA_0 together with the statement that, for every X , its ω -th jump $X^{(\omega)}$ exists. That statement is expressed in the language of second-order arithmetic by a formula of the form $\forall X \exists Y = \bigoplus_{n \in \mathbb{N}} Y_n (Y_0 = X) \wedge (\forall n Y_{n+1} = Y'_n)$ (see Chapter 9 for more details on the formalization of computability-theoretic notions in second-order arithmetic).

The uniformity in the proof of Lemma 8.4.4 makes it possible to instead prove the existence of two computable functions $e, d : \mathbb{N} \rightarrow \mathbb{N}$ such that, for every set X and every $n \geq 2$, $\Phi_{e(n)}^{X'}$ is a coloring $f_n : [\mathbb{N}]^n \rightarrow n$ with the property that for every infinite f_n -thin set H , $\Phi_{d(n)}^{H \oplus X} = X'$. Such a proof is formalizable in RCA_0 , and from it we can prove in RCA_0 a relativized version of Theorem 8.4.7, yielding the following reverse-mathematical corollary.

Corollary 8.4.9. Each of $\text{TS}^{\text{!}\omega}$ and $\text{FS}^{\text{!}\omega}$ implies ACA_0^+ over RCA_0 .

The above corollary coupled with Theorem 8.2.5 and Corollary 8.3.6 implies the equivalence over RCA_0 of $\text{RT}_2^{\text{!}\omega}$, $\text{TS}^{\text{!}\omega}$ and $\text{FS}^{\text{!}\omega}$. However, we do not know whether $\text{RT}_2^{\text{!}\omega}$ is strongly Weihrauch reducible to $\text{TS}^{\text{!}\omega}$ or $\text{FS}^{\text{!}\omega}$.

The remaining question is whether $\text{RRT}^{\text{!}\omega}$ codes $\emptyset^{(\omega)}$. In the next section, we shall answer this question negatively in a strong sense: $\text{RRT}^{\text{!}\omega}$ does not code any non-computable set. Some computability-theoretic weak anti-basis results for $\text{RRT}^{\text{!}\omega}$ can be obtained by streamlining results from the finite dimensional case. For instance, Csima and Mileti [32] proved the following theorem:

Theorem 8.4.10 (Csima-Mileti [32]). *For all $n, k \geq 2$ there exists a computable k -bounded $f : [\mathbb{N}]^n \rightarrow \mathbb{N}$ with no infinite Σ_n^0 rainbow.*

The proof of Theorem 8.4.10 is uniform in n . Using this uniformity we obtain the following, which implies that $\text{ACA}_0 \not\vdash \text{RRT}_2^{\text{!}\omega}$.

Theorem 8.4.11. *There is a computable instance of $\text{RRT}_2^{\text{!}\omega}$ with no arithmetical solution.*

Proof. For each $n \geq 2$ let $f_n : [\mathbb{N}]^n \rightarrow \mathbb{N}$ be the 2-bounded instance of RRT_2^n with no infinite Σ_n^0 rainbow given by Theorem 8.4.10. Let $g : [\mathbb{N}]^{\text{!}\omega} \rightarrow \mathbb{N}$ be defined as follows:

$$g(n, x_0, \dots, x_{n-1}) = \langle n, f_n(x_0, \dots, x_{n-1}) \rangle$$

The function g is clearly 2-bounded since each f_n is 2-bounded. Then for every g -rainbow H and every n in H , the set $H \setminus [0, n]$ is an f_n -rainbow, so H is not Σ_n^0 by Theorem 8.4.10. $\square$

Patey [110] proved that $\text{RCA}_0 \vdash (\forall n)[\text{RRT}_2^{n+1} \rightarrow \text{TS}^n]$ and for every $n > 0$, $\text{RCA}_0 \vdash \text{RRT}_2^{2n+1} \rightarrow \text{FS}^n$. The argument almost translates to the exactly ω -largeness setting as follows. A set $s \subseteq \mathbb{N}$ is *exactly* $(\omega + 1)$ -large if $s \setminus \{\min s\}$ is exactly ω -large. Let $\text{RRT}_2^{\text{!}(\omega+1)}$ be the rainbow Ramsey theorem for 2-bounded colorings of the exactly $(\omega + 1)$ -large sets.

Proposition 8.4.12. $\text{TS}^{\text{!}\omega} \leq_{\text{SW}} \text{RRT}_2^{\text{!}(\omega+1)}$.

Proof. Fix an instance $f : [\mathbb{N}]^{l\omega} \rightarrow \mathbb{N}$ of $\text{TS}^{l\omega}$. Then, consider the following f -computable instance g of $\text{RRT}_2^{l(\omega+1)}$: for every $s \in [\mathbb{N}]^{l\omega}$ and every $x \in \mathbb{N}$, if $f(s) = \langle x, y \rangle$ with $x < y < \min s$, let $g(x, s) = g(y, s)$ and otherwise assign to $g(x, s)$ a fresh color. The construction of g is uniform in f .

Let H be an infinite rainbow for g and let $x, y \in H$ with $x < y$. The set $H_1 = H \setminus [0, y]$ is f -thin for the color $\langle x, y \rangle$. Indeed, for every $s \in [H_1]^{l\omega}$, by definition of H_1 , $x < y < \min s$, thus, if $f(s) = \langle x, y \rangle$, then $g(x, s)$ would be equal to $g(y, s)$, contradicting the fact that H is a rainbow for g . The definition of H_1 is uniform in H , thus $\text{TS}^{l\omega} \leq_{\text{SW}} \text{RRT}_2^{l(\omega+1)}$. $\square$

Corollary 8.4.13. There exists a computable instance of $\text{RRT}_2^{l(\omega+1)}$ such that every solution computes $\emptyset^{(\omega)}$.

Proof. Immediate by Theorem 8.4.7 and Proposition 8.4.12. $\square$

8.5 $\text{RRT}^{l\omega}$ avoids cones

In this section, we prove that the rainbow Ramsey theorem for exactly ω -large sets admits strong cone avoidance. To this end, we introduce some variants of the free set theorem for large sets that themselves admit the strong cone avoidance property and are of independent interest. These variants will concern colorings where elements can have finitely many colors, defined as functions $f : [\mathbb{N}]^{l\omega} \rightarrow [\mathbb{N}]^{<\mathbb{N}}$. Accordingly, we can generalize the notion of an f -free set to these multi-colored assignments.

Definition 8.5.1. Let $\mathcal{C} \subseteq [\mathbb{N}]^{<\mathbb{N}}$ and $f : \mathcal{C} \rightarrow [\mathbb{N}]^{<\mathbb{N}}$ be a coloring. A set $H \subseteq \mathbb{N}$ is f -free if for every $s \in \mathcal{C} \cap [H]^{<\mathbb{N}}$, $f(s) \cap H \subseteq s$.

By the following result, we can transform every instance of $\text{RRT}^{l\omega}$ into an instance of a variant of the large free set theorem for these multi-colorings.

Proposition 8.5.2. Fix some $k \in \mathbb{N}$. For every k -bounded coloring $f : [\mathbb{N}]^{l\omega} \rightarrow \mathbb{N}$, there is an f' -computable coloring $g : [\mathbb{N}]^{l\omega} \rightarrow [\mathbb{N}]^{<\mathbb{N}}$ satisfying, for every $s \in [\mathbb{N}]^{l\omega}$:

- $\text{card } g(s) \leq k$.
- $\min g(s) \geq \min s$.

and such that every infinite g -free set is an f -rainbow.

Proof. Let $\leq_{\text{lex}}$ denote the lexicographic order on $[\mathbb{N}]^{l\omega}$ (seen as finite increasing sequences in $\mathbb{N}^{<\mathbb{N}}$). In particular, if $s <_{\text{lex}} t$, then $\min s \leq \min t$ and hence $\text{card } s \leq \text{card } t$ by definition of an exactly ω -large set. Since $s \neq t$, $t \setminus s$ is not empty and moreover $\min(t \setminus s) \geq \min s$.

Let $g(s) = \{\min(t \setminus s) : f(t) = f(s) \wedge s <_{\text{lex}} t\}$. For every $s \in [\mathbb{N}]^{l\omega}$, we have indeed $\text{card } g(s) \leq k$ and $\min g(s) \geq \min s$.

Note that g is computable in the Turing jump of f , as one needs to make an unbounded search for all $t >_{\text{lex}} s$.

We claim that every infinite g -free set H is an f -rainbow. Suppose by way of contradiction that there are some distinct $s, t \in [H]^{l\omega}$ such that $f(s) = f(t)$. One can suppose without loss of generality that $s <_{\text{lex}} t$. Then $\min(t \setminus s) \in (g(s) \cap H) \setminus s$, contradicting the g -freeness of H . $\square$

It may not be obvious at first what the purpose of this reduction is, as the large free set theorem was shown in Corollary 8.4.8 not to admit the strong cone avoidance property. Nevertheless, the requirements that, for every $s \in [\mathbb{N}]^{l\omega}$, $\text{card } g(s) \leq k$ and $\min g(s) \geq \min s$ severely restrict the coding power of our instances. We will refer to multi-colorings g satisfying these two requirements as being k -constrained and progressive respectively.

Definition 8.5.3. Let $\mathcal{C} \subseteq [\mathbb{N}]^{<\mathbb{N}}$. A coloring $f : \mathcal{C} \rightarrow [\mathbb{N}]^{<\mathbb{N}}$ is *b-constrained* for a bounding function $b : \mathbb{N} \rightarrow \mathbb{N}$ if for every $s \in \mathcal{C}$, $\text{card } f(s) \leq b(\min s)$. If b is the constant function $x \mapsto k$, then we say that f is *k-constrained*.

Definition 8.5.4. Fix $\mathcal{C} \subseteq [\mathbb{N}]^{<\mathbb{N}}$. A coloring $f : \mathcal{C} \rightarrow [\mathbb{N}]^{<\mathbb{N}}$ is *progressive* if for every $s \in \mathcal{C}$, either $f(s) = \emptyset$, or $\min f(s) \geq \min s$.

Thus, in light of Proposition 8.5.2, instead of working directly with $\text{RRT}^{\text{!}\omega}$, we will instead prove the strong cone avoidance of the following theorem.

Statement 8.5.5 (Progressive free set theorem). Fix some $k \in \mathbb{N}$. $\text{PFS}_k^{\text{!}\omega}$ is the statement “For every k -constrained progressive coloring $f : [\mathbb{N}]^{\text{!}\omega} \rightarrow [\mathbb{N}]^{<\mathbb{N}}$, there is an infinite f -free set”. $\text{PFS}^{\text{!}\omega}$ is the statement $\forall k \text{PFS}_k^{\text{!}\omega}$.

For our purpose, it is irrelevant that the instance of $\text{PFS}^{\text{!}\omega}$ constructed in Proposition 8.5.2 is computationally more complex than the instance of $\text{RRT}^{\text{!}\omega}$. Indeed, in a strong cone avoidance proof, the instances considered are already of arbitrarily high complexity.

The fact that the colorings we consider are progressive will greatly simplify the construction of solutions to our instances. Indeed, when one unfolds the reductions, the coloring witnessing that $\text{FS}^{\text{!}\omega}$ codes $\emptyset^{(\omega)}$ is not progressive. This is actually necessarily the case, as we shall now prove that $\text{PFS}^{\text{!}\omega}$ admits strong cone avoidance.

Theorem 8.5.6. Fix a set Z and a non- Z -computable set D . For every $k \in \mathbb{N}$ and every k -constrained progressive coloring $f : [\mathbb{N}]^{\text{!}\omega} \rightarrow [\mathbb{N}]^{<\mathbb{N}}$, there exists an infinite f -free set $G \subseteq \mathbb{N}$ such that $D \not\leq_T G \oplus Z$.

Proof. For simplicity, we prove the theorem in a non-relativized form (i.e., for $Z = \emptyset$). Relativization is straightforward. Fix a non-computable set D and let $f : [\mathbb{N}]^{\text{!}\omega} \rightarrow [\mathbb{N}]^{<\mathbb{N}}$ be a k -constrained, progressive coloring, for some $k \in \mathbb{N}$.

Let $\mathcal{P}$ be the collection of all progressive colorings of type $[\mathbb{N}]^{\leq \text{!}\omega} \rightarrow [\mathbb{N}]^{<\mathbb{N}}$, where $[A]^{\leq \text{!}\omega}$ denotes the subsets $s \subseteq A$ such that $\text{card } s \leq \min s + 1$. For a function $b : \mathbb{N} \rightarrow \mathbb{N}$ we write $\mathcal{P}_b$ for the class of all b -constrained colorings in $\mathcal{P}$.

Given two colorings $g, h \in \mathcal{P}$, we write $g \leq h$ if for every $s \in [\mathbb{N}]^{\leq \text{!}\omega}$, $g(s) \supseteq h(s)$. Note that if H is g -free and $g \leq h$, then H is h -free. Given two colorings $g, h \in \mathcal{P}$, let $g \cup h$ be the coloring defined by $(g \cup h)(s) = g(s) \cup h(s)$. The coloring $g \cup h$ is the greatest lower bound of g and h with respect to $\leq$.

We will build an infinite f -free set G such that $D \not\leq_T G$ by approximating it using an infinite sequence of good conditions:

Definition 8.5.7. A *condition* is a triple (g, σ, X) such that $g \in \mathcal{P}_{b_g}$ for some computable function $b_g : \mathbb{N} \rightarrow \mathbb{N}$, σ is some finite binary sequence called the *stem*, and $X \subseteq \mathbb{N}$ is an infinite set, called the *reservoir*, such that $\max \sigma < \min X$ and such that $D \not\leq_T X$.

Definition 8.5.8. A condition (g, σ, X) is *good* if it satisfies:

- (a) for every $s \in [\sigma \cup X]^{\leq \text{!}\omega}$ with $\min s \in \sigma$, $g(s) \cap X \subseteq s$.
- (b) for every $s \in [\sigma \cup X]^{\leq \text{!}\omega}$, $g(s) \cap \sigma \subseteq s$.

A good condition (g, σ, X) can be seen as a finite approximation of a g -free set. The stem σ corresponds to an initial segment of that set, while the elements outside the reservoir X are guaranteed not to be in the set. We write

$$[(g, \sigma, X)] = \{Y \subseteq \mathbb{N} : \sigma < Y \subseteq \sigma \cup X \text{ and } Y \text{ is } g\text{-free}\}$$

for the collection of all the sets compatible with the condition (g, σ, X) (i.e., potential candidates for the set being built).

Initially, our construction will start from the condition $(\hat{f}, \emptyset, \mathbb{N})$ where $\hat{f} : [\mathbb{N}]^{\leq \omega} \rightarrow [\mathbb{N}]^{< \mathbb{N}}$ is the k -constrained function extending f by letting $\hat{f}(s) = \emptyset$ for every $s \in [\mathbb{N}]^{\leq \omega} \setminus [\mathbb{N}]^{\omega}$. This condition is vacuously good.

We will then construct a sequence $(f_n, \sigma_n, X_n)_{n \in \mathbb{N}}$ of good conditions, with $(f_0, \sigma_0, X_0) = (\hat{f}, \emptyset, \mathbb{N})$ and such that the sequence $[(f_n, \sigma_n, X_n)]_{n \in \mathbb{N}}$ is decreasing, with the sets belonging to $[(f_n, \sigma_n, X_n)]$ satisfying stronger and stronger properties at each stage. By ensuring that the length of the σ_n is unbounded, the intersection $\bigcap_{n \in \mathbb{N}} [(f_n, \sigma_n, X_n)]$ will then be reduced to a singleton $\{G\}$, where $G = \bigcup_{n \in \mathbb{N}} \sigma_n$ will be our infinite f -free set such that $G \not\leq_T D$.

This leads to the following notion of extension between conditions:

Definition 8.5.9. A condition $d = (h, \tau, Y)$ *extends* $c = (g, \sigma, X)$ (written $d \leq c$) if $\tau \geq \sigma$, $Y \subseteq X$, $\tau \setminus \sigma \subseteq X$ and $h \leq g$.

The terminology may be confusing: if $d \leq c$, then $[d] \subseteq [c]$, which explains why the order is reversed.

To ensure infinitude, we will show that for every good condition $c = (g, \sigma, X)$, there exists a good condition $d = (h, \tau, Y) \leq c$ such that $\text{card } \tau > \text{card } \sigma$. To guarantee that the resulting set does not compute D , we will prove that for every Turing index $e \in \mathbb{N}$ and every good condition c , there exists some good condition $d \leq c$, such that $\Phi_e^H \neq D$ for every $H \in [d]$.

The following lemma states that property (a) of a good condition can be obtained “for free”, that is, by restricting the reservoir, and therefore does not impose any constraint on the stem.

Lemma 8.5.10. *Every condition (g, σ, X) can be extended into a condition (g, σ, Y) satisfying property (a) of Definition 8.5.8.*

Proof. For every $t \in [\sigma]^{\leq \omega}$ with $t \neq \emptyset$, every $i \leq \min t - \text{card } t + 1$ and every $j \leq b_g(\min t)$, let $g_{t,i,j} : [X]^i \rightarrow \mathbb{N}$ be defined for every $u \in [X]^i$ by letting $g_{t,i,j}(u)$ be the j -th element of $g(t \cdot u)$ if it exists and 0 otherwise.

As σ is a finite binary sequence, there are finitely many such colorings $g_{t,i,j}$, hence, by finitely many successive applications of the strong cone avoidance of $\mathbf{FS}^n$ for $n \in \mathbb{N}$, there exists an infinite subset $Y \subseteq X$ such that $D \not\leq_T Y$ and such that Y is simultaneously $g_{t,i,j}$ -free for every $t \in [\sigma]^{\leq \omega}$ with $t \neq \emptyset$, every $i \leq \min t - \text{card } t + 1$ and every $j \leq b_g(\min t)$.

We claim that Y is our desired set. Indeed, for every $s \in [\sigma \cup Y]^{\leq \omega}$ with $\min s \in \sigma$, letting $t = s \cap \sigma$ and $u = s \cap Y$, we have that every element of $g(s)$ is equal to $g_{t,i,j}(u)$ for $i = \text{card } u$ and some $j \leq b_g(\min t)$. So, by $g_{t,i,j}$ -freeness of Y , $g(s) \cap Y \subseteq u \subseteq s$. $\square$

The main difficulty lies in showing that elements can be added to the stem while preserving property (b). To this end, we will use the result of Cholak and Patey [17] that $\mathbf{TS}_k^n$ admits the strong cone avoidance property when $k > C_n$, for C_n the n -th Catalan number, to show that, for any good condition (g, σ, X) , any element of X can be added to σ , at the cost of restricting the reservoir X to a smaller set.

Definition 8.5.11. Let (g, σ, X) be a condition and $\ell \in \mathbb{N}$. We say that a set $Y \subseteq X$ *g-stabilizes* $\{0, \dots, \ell-1\}$ if for every $t \in [\ell]^{\leq \omega}$ with $t \neq \emptyset$ and every $n \leq \min t + 1 - \text{card } t$, there is a set $I_{t,n} \in [\ell]^{\leq b_g(\min t) \times C_n}$ such that for every $u \in [Y]^n$, $g(t \cdot u) \cap \{0, \dots, \ell-1\} \subseteq I_{t,n}$.

Lemma 8.5.12. *Let $c = (g, \sigma, X)$ be a condition and $\ell \in \mathbb{N}$, then c can be extended into a condition (g, σ, Y) , such that Y g-stabilizes $\{0, \dots, \ell-1\}$.*

Proof. For every $t \in [\ell]^{\leq \text{l}\omega}$ with $t \neq \emptyset$ and every $n \leq \min t + 1 - \text{card } t$, let $g_{t,n} : [X]^n \rightarrow [\ell]^{\leq b_g(\min t)}$ be defined by $g_{t,n}(u) = g(t \cdot u) \cap \{0, \dots, \ell - 1\}$. By finitely many successive applications of strong cone avoidance of TS_k^n for $k > C_n$, there is an infinite subset $Y \subseteq X$ such that $D \not\leq_T Y$ and for every $t \in [\ell]^{\leq \text{l}\omega}$ with $t \neq \emptyset$ and every $n \leq \min t + 1 - \text{card } t$, $\text{card } g_{t,n}([Y]^n) \leq C_n$. Note that here, a color is an element of $[\ell]^{\leq b_g(\min t)}$ instead of a natural number. Let $I_{t,n} = \bigcup g_{t,n}([Y]^n)$. Then $\text{card } I_{t,n} \leq b_g(\min t) \times C_n$. By definition, for every $u \in [Y]^n$, $g(t \cdot u) \cap \{0, \dots, \ell - 1\} = g_{t,n}(u) \subseteq I_{t,n}$. $\square$

Given a computable function $b : \mathbb{N} \rightarrow \mathbb{N}$, let $b^+ : \mathbb{N} \rightarrow \mathbb{N}$ be the computable function defined by $b^+(m) = \sum_{n \leq m} b(m) \times C_n$.

Definition 8.5.13. Let (g, σ, Y) be a condition and $\ell \in \mathbb{N}$ such that Y g -stabilizes $\{0, \dots, \ell - 1\}$. The *limit coloring* is the function $g_{\ell,Y} : [\ell]^{\leq \text{l}\omega} \rightarrow [\ell]^{< \mathbb{N}}$ defined by $g_{\ell,Y}(\emptyset) = \emptyset$ and $g_{\ell,Y}(t) = \bigcup_{n \leq \min t + 1 - \text{card } t} I_{t,n}$, for $I_{t,n}$ as in Definition 8.5.11, otherwise.

The limit function $g_{\ell,Y}$ is b_g^+ -constrained. Note that if $\rho \subseteq \{0, \dots, \ell - 1\}$ is $g_{\ell,Y}$ -free, then it is g -free. Indeed, for every $t \in [\ell]^{\leq \text{l}\omega}$, $I_{t,0} \subseteq g_{\ell,Y}(t)$ and by definition of g -stabilization for $n = 0$, $g(t) \cap \{0, \dots, \ell - 1\} \subseteq I_{t,0}$.

The following lemma is the core combinatorial lemma which specifies the conditions under which a block of elements ρ can be added to the stem while preserving the property (b).

Lemma 8.5.14. Let $(g, \sigma, X) \geq (g, \sigma, Y)$ be good conditions and $\ell \in \mathbb{N}$ such that Y g -stabilizes $\{0, \dots, \ell - 1\}$ with limit coloring $g_{\ell,Y} : [\ell]^{\leq \text{l}\omega} \rightarrow [\ell]^{< \mathbb{N}}$. Let $\rho \subseteq X \cap \{0, \dots, \ell - 1\}$ be a finite $g_{\ell,Y}$ -free set. Then $(g, \sigma \cup \rho, Y)$ satisfies property (b).

Proof. Fix some $s \in [\sigma \cup \rho \cup Y]^{\leq \text{l}\omega}$. We have multiple cases.

- Case 1: $s \cap \sigma \neq \emptyset$. Then by properties (a) and (b) of (g, σ, X) , $g(s) \cap (\sigma \cup X) \subseteq s$. Since $(\sigma \cup \rho) \subseteq (\sigma \cup X)$, then $g(s) \cap (\sigma \cup \rho) \subseteq s$.
- Case 2: $s \cap \sigma = \emptyset$ but $s \cap \rho \neq \emptyset$. By property (b) of (g, σ, X) , $g(s) \cap \sigma \subseteq s$. Let $t = s \cap \rho$ and $u = s \cap Y$. Since ρ is $g_{\ell,Y}$ -free, $g_{\ell,Y}(t) \cap \rho \subseteq t$. By definition of $g_{\ell,Y}(t)$, $I_{t, \text{card } u} \subseteq g_{\ell,Y}(t)$, so $I_{t, \text{card } u} \cap \rho \subseteq t$. In particular, $g(s) = g(t, u) \subseteq I_{t, \text{card } u}$, so $g(s) \cap \rho \subseteq t \subseteq s$. Thus, $g(s) \cap (\sigma \cup \rho) \subseteq s$.
- Case 3: $s \cap (\sigma \cup \rho) = \emptyset$. Then $s \subseteq Y$. Since $\min Y > \max(\sigma \cup \rho)$, then by progressiveness of g , $g(s) \cap (\sigma \cup \rho) = \emptyset$.

$\square$

One can combine Lemmas 8.5.10, 8.5.12 and 8.5.14 to obtain an extensibility lemma.

Lemma 8.5.15. Let (g, σ, X) be a good condition. There is a good extension $(g, \tau, Y) \leq (g, \sigma, X)$ such that $|\tau| > |\sigma|$.

Proof. Let $x = \min X$. By Lemma 8.5.12, there is an infinite subset $Y_0 \subseteq X$ g -stabilizing $\{0, \dots, x\}$ such that $D \not\leq_T Y_0$. Let g_{x+1, Y_0} be the limit function. By Lemma 8.5.10, there is an infinite subset $Y \subseteq Y_0$ such that $(g, \sigma \cup \{x\}, Y)$ satisfies property (a). By Lemma 8.5.14, $\{x\}$ being vacuously $g_{x+1, Y}$ -free, $(g, \sigma \cup \{x\}, Y)$ satisfies property (b). Thus, $(g, \sigma \cup \{x\}, Y)$ is a good extension. $\square$

Given a good condition $c = (g, \sigma, X)$ and a Turing index e , we will prove the existence of a good condition $d \leq c$ such that $\Phi_e^H \neq D$ for every $H \in [d]$. The main difficulty is that, for a given set H , there are two ways to ensure that $\Phi_e^H \neq D$: either there is some $x \in \mathbb{N}$ such that $\Phi_e^H(x) \downarrow \neq D(x)$, or there is some $x \in \mathbb{N}$ such that $\Phi_e^H(x) \uparrow$. And, depending on the index e and the condition c , we may not be able to

choose freely which of these two strategies to apply. The following *forcing question* $? \vdash$ helps us decide which case we are in.

Definition 8.5.16 (forcing question). Let (g, σ, X) be a condition and $\varphi(G)$ be a $\Sigma_1(G)$ -formula in Kleene normal form (see Proposition 5.1.2). Let $(g, \sigma, X) ? \vdash \varphi(G)$ hold if and only if for every $h \in \mathcal{P}_{b_g^+}$, there is a finite h -free set $\rho \subseteq X$ such that $\varphi(\sigma \cup \rho)$ holds.

The previous formulation of the forcing question is $\Pi_1^1(G)$ as it starts with a second-order universal quantification. However, by the following compactness result, it is equivalent to a $\Sigma_1(G)$ -formula:

Lemma 8.5.17. *Let (g, σ, X) be a condition and $\varphi(G)$ be a $\Sigma_1(G)$ -formula in Kleene normal form. Then $(g, \sigma, X) ? \vdash \varphi(G)$ if and only if there is some $k \in \mathbb{N}$ such that for every b_g^+ -constrained progressive function $h : [k]^{\leq \omega} \rightarrow [k]^{< \mathbb{N}}$, there is a finite h -free set $\rho \subseteq X \cap \{0, \dots, k-1\}$ such that $\varphi(\sigma \cup \rho)$ holds.*

Proof. Suppose first that there is some $k \in \mathbb{N}$ such that for every b_g^+ -constrained progressive function $h : [k]^{\leq \omega} \rightarrow [k]^{< \mathbb{N}}$, there is a finite h -free set $\rho \subseteq X \cap \{0, \dots, k-1\}$ such that $\varphi(\sigma \cup \rho)$ holds. Let $h \in \mathcal{P}_{b_g^+}$, and define $\hat{h} : [k]^{\leq \omega} \rightarrow [k]^{< \mathbb{N}}$ by $\hat{h}(s) = h(s) \cap \{0, \dots, k-1\}$, by assumption there is some finite $\hat{h}$ -free set $\rho \subseteq X \cap \{0, \dots, k-1\}$ such that $\varphi(\sigma \cup \rho)$ holds. In particular, ρ is h -free. Since there is such a ρ for every $h \in \mathcal{P}_{b_g^+}$, $(g, \sigma, X) ? \vdash \varphi(G)$ holds.

Suppose now that for every $k \in \mathbb{N}$, there is a b_g^+ -constrained progressive function $h_k : [k]^{\leq \omega} \rightarrow [k]^{< \mathbb{N}}$ such that for every h_k -free set $\rho \subseteq X \cap \{0, \dots, k-1\}$, $\varphi(\sigma \cup \rho)$ does not hold. Let T be the tree which, at level k , contains all such functions h_k , and which is ordered by $h_k < h_{k'}$ if for every $s \in [k]^{\leq \omega}$, $h_k(s) = h_{k'}(s) \cap \{0, \dots, k-1\}$. The tree T is finitely branching, so by König's lemma, there is an infinite path in T . This path corresponds to a function $h \in \mathcal{P}_{b_g^+}$ such that for every finite h -free set $\rho \subseteq X$, $\varphi(\sigma \cup \rho)$ does not hold, witnessing that $(g, \sigma, X) ? \vdash \varphi(G)$ does not hold. $\square$

For $\varphi(X)$ a $\Sigma_1(X)$ - or a $\Pi_1(X)$ -formula, we say that a good condition c *forces* $\varphi(X)$ if $\varphi(H)$ holds for every $H \in [c]$.

Lemma 8.5.18. *Let $c = (g, \sigma, X)$ be a good condition and $\varphi(G)$ be a $\Sigma_1(G)$ -formula.*

1. *If $c ? \vdash \varphi(G)$, then there is a good extension of c forcing $\varphi(G)$.*
2. *If $c ? \not\vdash \varphi(G)$, then there is a good extension of c forcing $\neg \varphi(G)$.*

Proof. Suppose first that $c ? \vdash \varphi(G)$ holds. By Lemma 8.5.17, there is some $k \in \mathbb{N}$ such that for every b_g^+ -constrained progressive function $h : [k]^{\leq \omega} \rightarrow [k]^{< \mathbb{N}}$, there is a finite h -free set $\rho \subseteq X$ such that $\varphi(\sigma \cup \rho)$ holds. By Lemma 8.5.12, there is an infinite subset $Y_0 \subseteq X$ h -stabilizing $\{0, \dots, k-1\}$ and such that $D \not\leq_T Y_0$. Let $h_{k, Y_0} : [k]^{\leq \omega} \rightarrow [k]^{< \mathbb{N}}$ be the limit function. Note that h_{k, Y_0} is b_g^+ -constrained and progressive, so there is a finite h_{k, Y_0} -free set $\rho \subseteq X$ such that $\varphi(\sigma \cup \rho)$ holds. By Lemma 8.5.14, $(g, \sigma \cup \rho, Y_0)$ satisfies (b). By Lemma 8.5.10, there is an infinite subset $Y \subseteq Y_0$ such that $(g, \sigma \cup \rho, Y)$ satisfies (a) and $D \not\leq_T Y$. Thus $(g, \sigma \cup \rho, Y)$ is a good extension of c . By choice of ρ , it forces $\varphi(G)$.

Suppose $c ? \not\vdash \varphi(G)$ holds. Then there is some $h \in \mathcal{P}_{b_g^+}$ such that for every finite h -free set $\rho \subseteq X$, $\varphi(\sigma \cup \rho)$ does not hold. Let $\hat{g} : [\mathbb{N}]^{\leq \omega} \rightarrow [\mathbb{N}]^{< \mathbb{N}}$ be defined by $\hat{g}(s) = g(s) \cup (h(s) \setminus \sigma)$. Note that $\hat{g} \leq g$, and every $\hat{g}$ -free subset $\rho \subseteq X$ is h -free. Moreover, $\hat{g}$ is $(b_g + b_g^+)$ -constrained and progressive. Since (g, σ, X) satisfies (b) and for every $s \in [\sigma \cup X]^{\leq \omega}$, $\hat{g}(s) \cap \sigma \subseteq g(s) \cap \sigma \subseteq s$, then $(\hat{g}, \sigma, X)$ satisfies (b). By Lemma 8.5.10, there is an infinite subset $Y \subseteq X$ such that $D \not\leq_T Y$ and $(\hat{g}, \sigma, Y)$ satisfies (a). Thus $d = (\hat{g}, \sigma, Y)$ is a good extension of c .

We claim that d forces $\neg \varphi(G)$. Indeed, suppose there is an extension $(\tilde{g}, \tau, Z) \leq d$ such that $\varphi(\tau)$ holds. Then by property (b) of $(\tilde{g}, \tau, Z)$, τ is $\tilde{g}$ -free, and since $\tilde{g} \leq \hat{g}$,

τ is $\hat{g}$ -free. Let $\rho = \tau \setminus \sigma$. By definition of $\hat{g}$, ρ is a h -free subset of X , contradicting our choice of h . $\square$

We can now prove our diagonalization lemma.

Lemma 8.5.19. *Let $c = (g, \sigma, X)$ be a good condition and Φ_e be a Turing functional. There is a good extension of c forcing $\Phi_e^G \neq D$.*

Proof. Let $U = \{(x, v) \in \mathbb{N} \times 2 : c \Vdash \Phi_e^G(x) \downarrow = v\}$. We have three cases:

- Case 1: $(x, 1 - D(x)) \in U$ for some $x \in \mathbb{N}$. By Lemma 8.5.18, there is a good extension of c forcing $\Phi_e^G(x) \downarrow = 1 - D(x)$, hence forcing $\Phi_e^G \neq D$.
- Case 2: $(x, D(x)) \notin U$ for some $x \in \mathbb{N}$. By Lemma 8.5.18, there is a good extension of c forcing $\neg(\Phi_e^G(x) \downarrow = D(x))$, hence forcing $\Phi_e^G \neq D$.
- Case 3: U is the graph of the characteristic function of D . By Lemma 8.5.17, the set U is $\Sigma_1(X)$, so $D \leq_T X$, contradiction. $\square$

We can now finish the proof of Theorem 8.5.6. By Lemma 8.5.15 and Lemma 8.5.19, there exists a decreasing sequence (f_n, σ_n, X_n) of good conditions such that:

- Every $G \in [(f_0, \sigma_0, X_0)]$ is f -free,
- $\text{card } \sigma_n \geq n$ for every $n \in \mathbb{N}$,
- Every $G \in [(f_{e+1}, \sigma_{e+1}, X_{e+1})]$ is such that $\Phi_e^G \neq D$ for every $e \in \mathbb{N}$.

Thus, for $H = \bigcup_{n \in \mathbb{N}} \sigma_n$, we have $H \in \bigcap_{n \in \mathbb{N}} [(f_n, \sigma_n, X_n)]$ and H is an infinite f -free set such that $H \not\leq_T D$. $\square$

Corollary 8.5.20. $\text{RRT}^{\text{!}\omega}$ admits strong cone avoidance.

Proof. Fix a set Z and let D be a non- Z -computable set.

Fix $k > 0$, let $f : [\mathbb{N}]^{\text{!}\omega} \rightarrow \mathbb{N}$ be an instance of $\text{RRT}_k^{\text{!}\omega}$ and let $g : [\mathbb{N}]^{\text{!}\omega} \rightarrow [\mathbb{N}]^{<\mathbb{N}}$ be the f' -computable k -constrained progressive coloring obtained in Proposition 8.5.2.

Then, by Theorem 8.5.6, there exists an infinite g -free set $G \subseteq \mathbb{N}$ such that $D \not\leq_T G \oplus Z$. By construction of g , G is also an infinite f -rainbow. $\square$

8.6 Conclusions and perspectives

The analysis of the exactly ω -large counterparts to the Ramsey's, free set and thin set theorems from a computable perspective, gives the exact same tight bound, namely, $\emptyset^{(\omega)}$, translating in reverse mathematical terms by an equivalence with ACA_0^+ . This equivalence is to be put in contrast with the finite-dimensional cases, where RT_2^n coincides with ACA_0 for $n \geq 3$, while FS^n and TS^n both admit strong cone avoidance. On the other hand, the rainbow Ramsey theorem for exactly ω -large sets ($\text{RRT}_k^{\text{!}\omega}$) still has no coding power, and admits strong cone avoidance. Figure 8.1 and Figure 8.2 summarize the relationship between the studied statements, in Reverse Mathematics and over strong Weihrauch reducibility, respectively.

Many questions remain open around the generalization of combinatorial statements to exactly ω -large sets and barriers.

By Theorem 8.3.7, $\text{RRT}_k^{\text{!}\omega} \leq_{\text{SW}} \text{RT}_k^{\text{!}\omega}$, so every computable instance of $\text{RRT}_k^{\text{!}\omega}$ admits a $\emptyset^{(\omega)}$ -computable solution. By Corollary 8.4.8, there exists a computable instance of $\text{FS}^{\text{!}\omega}$ such that every solution computes $\emptyset^{(\omega)}$. It follows that $\text{RRT}_k^{\text{!}\omega}$ is computably reducible to $\text{FS}^{\text{!}\omega}$ in the sense of Dzhafarov [38]. We give a direct combinatorial reduction in Proposition 8.3.8 in the case $k = 2$, and leave the general case open.

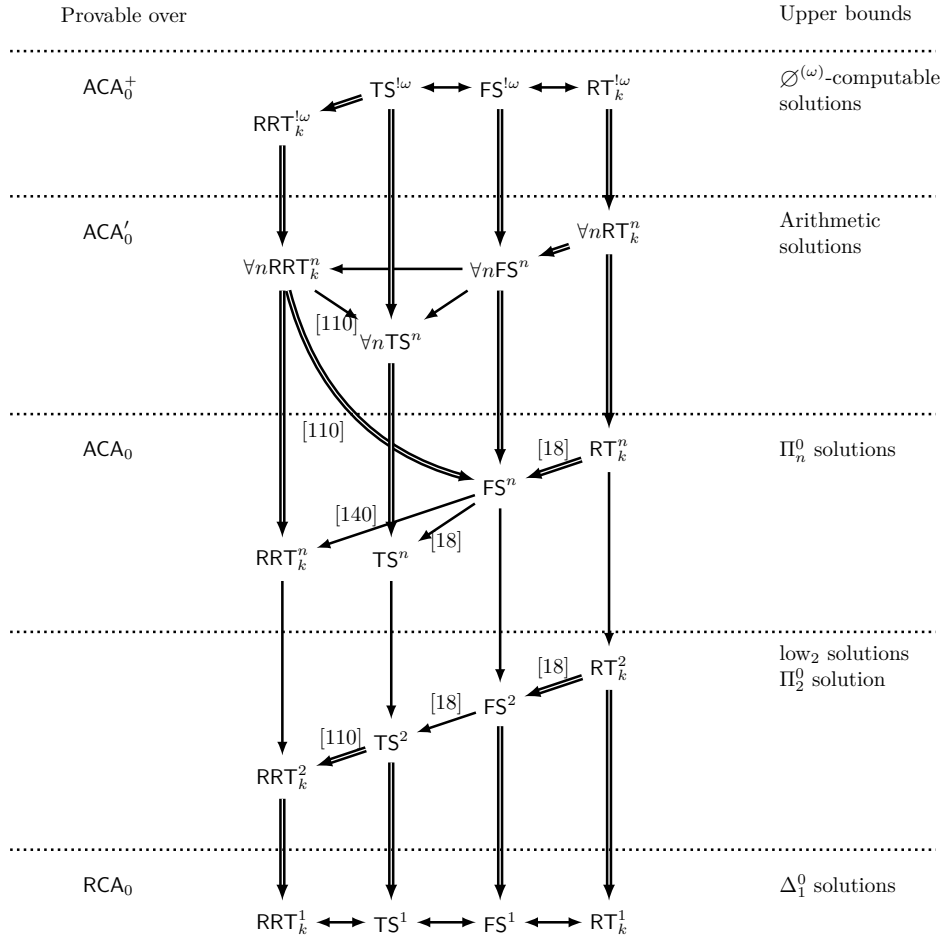


Figure 8.1: Studied statements from the perspectives of computability theory and reverse mathematics. A simple (double) arrow represents a (strict) implication over RCA_0 .

ACA_0' is the theory comprised of RCA_0 together with the statement $\forall n \forall X \exists Y (Y = X^{(n)})$, it sits strictly between ACA_0 and ACA_0^+ .

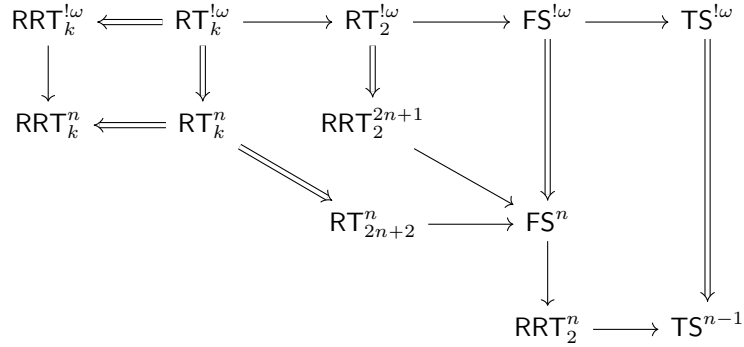


Figure 8.2: Studied statements from the perspective of Weihrauch analysis. A simple (double) arrow represents a (strict) strong Weihrauch reduction.

Question 8.6.1. Does $\text{RRT}_k^{! \omega} \leq_{\text{sW}} \text{FS}^{! \omega}$ for every $k > 0$?

While $\text{RT}^{! \omega}$, $\text{FS}^{! \omega}$ and $\text{TS}^{! \omega}$ code $\emptyset^{(\omega)}$ and are all equivalent to ACA_0^+ , we do not know of a direct proof of $\text{RT}_2^{! \omega}$ from either $\text{TS}^{! \omega}$ or $\text{FS}^{! \omega}$. More precisely, we do not know if $\text{RT}_2^{! \omega}$ is reducible to either $\text{TS}^{! \omega}$ or $\text{FS}^{! \omega}$.

Question 8.6.2. Does $\text{RT}_2^{! \omega} \leq_{\text{sW}} \text{TS}^{! \omega}$? Does $\text{RT}_2^{! \omega} \leq_{\text{sW}} \text{FS}^{! \omega}$?

Part III

First-order part

Chapter 9

Non-standard models

The remaining two parts of this thesis will involve non-standard models, that is, models $\mathcal{M} = (M, S)$ whose first-order part M is not the standard set of integers $\omega = \{0, 1, \dots\}$. Reasoning within such models can be quite counter-intuitive, since our aim is to study situations in which they closely resemble the standard integers while still differing in some crucial aspect.

As discussed in Section 5.2, the first-order part of all of these models will have a similar structure: each contains an initial segment isomorphic to the standard integers, followed by additional non-standard elements. Thus, from an external point of view, it is clear that these models are not the standard integers, since they contain elements with infinitely many predecessors.

However, from an internal perspective, the situation is less clear. Consider, for example, the case of geometry in which models that differ from the Euclidean plane $\mathbb{R}^2$ have been built. Within such models, what is referred to as a “straight line” is simply an object satisfying the given definitions and axioms, and need not coincide with a straight line in the usual Euclidean sense. For instance, in the Poincaré half-plane model, straight lines correspond to semicircles (orthogonal to the boundary) or vertical rays. But, from an internal point of view, what makes these “straight lines” different from those of $\mathbb{R}^2$ is not their unusual construction, but rather the fact that they fail to satisfy the parallel postulate: given a line L and a point P not on L , there are infinitely many lines L' passing through P that do not intersect L . Indeed, this property is formulated purely within the language of Euclidean geometry and makes no reference to $\mathbb{R}^2$.

Similarly, in non-standard models of arithmetic, a “finite” object will simply be an object that can be encoded by a first-order element of that model, even if from an external point of view this object is actually infinite.

Thus, when working with non-standard models, it is important to distinguish clearly between the two levels of description, internal and external, that is, between the object language and the meta-language. To avoid ambiguity, we will use the symbol ω to denote the set of standard integers, and $\mathbb{N}$ will be the symbol referring to the collection of all integers as expressed within the language of the theories under consideration. For example, we have that for every $n \in \omega$, $\text{ACA}_0 \vdash \forall a \in \mathbb{N} \forall f : [\mathbb{N}]^n \rightarrow a, \exists H \subseteq \mathbb{N} (H \text{ infinite} \wedge \text{card } f([H]^n) = 1)$, i.e., that for every $n \in \omega$, $\text{ACA}_0 \vdash \text{RT}^n$, but we don’t have $\text{ACA}_0 \vdash \forall b \in \mathbb{N} \text{RT}^b$.

An *external induction* can always be used to prove that the standard integers of a given model satisfy a given property, whereas an *internal induction* requires the model to satisfy some induction axioms but allows us to prove the property for every element of the model.

We will sometimes view non-standard integers as encoding ordinals, in that case, we will use the notations ω and ϵ_0 to refer to the ordinals ω and ϵ_0 in the model.

We will often make use of these two levels of description in our constructions and proofs. For example, in non-standard models of geometry, one may construct lines and points purely from the given postulates and axioms (for instance, from two points P_0 and P_1 , there exists a unique line passing through them), or instead define them externally (for example, in the Poincaré half-plane, by referring back to the equation of the corresponding semicircles). The external approach offers greater freedom in the kinds of objects that can be constructed, including some that may be meaningless within the model, whereas the internal approach makes it easier to prove properties of the constructed objects.

9.1 Computability and definability in non-standard models

To obtain conservation, equiconsistency or non-implication results, our proofs will typically follow a similar pattern: we begin with a (potentially non-standard) model $\mathcal{M} = (M, S)$ and transform it into a new model $\mathcal{N} = (N, T)$ satisfying a given second-order statement P . The goal will often be to make sure we preserve as many properties of $\mathcal{M}$ in $\mathcal{N}$ as possible. In this part, we shall restrict our attention to so-called ω -extensions¹ of $\mathcal{M}$.

Definition 9.1.1. A model $\mathcal{M} = (M, S)$ is an ω -submodel of $\mathcal{N} = (N, T)$ (or equivalently, $\mathcal{N}$ is an ω -extension of $\mathcal{M}$) if $N = M$ and $S \subseteq T$. We will often write $\mathcal{M} \subseteq \mathcal{N}$ when that is the case.

Building an ω -extension ensures that every Σ_1^1 property of $\mathcal{M}$ is preserved in $\mathcal{N}$ (and in particular every first-order property). In the case of conservation proofs, this yields conservation for all arithmetical formulas (and even for Π_1^1 -formulas).

The construction of these new models will often be similar to the proof of Proposition 8.1.5, in which solutions to instances of P are added one at a time to the second-order part of the model.

In the standard case, such constructions are made more intuitive thanks to the links between computability and definability. These links can still be used in non-standard models, although significant care is required to ensure that the various equivalences used do not require more induction than what is at our disposal.

For that, we will first see how to interpret the arithmetical hierarchy in non-standard models.

9.1.1 Non-standard arithmetical sets

When trying to define what should count as an arithmetical set within a non-standard model M of first-order arithmetic, one natural approach is to work externally and consider sets of the form $\{a \in M : M \models \varphi(a)\}$ for every arithmetical formula $\varphi(x)$. The problem with this approach is that even very simple sets, such as singletons, need not be arithmetically definable: for $a \in M$, there may not be any arithmetical formula $\varphi(x)$ satisfied uniquely by a . This issue is resolved by allowing $\varphi(x)$ to have first-order parameters from M (in which case the formula $x = a$ defines the singleton $\{a\}$).

This modification can be justified from an internal perspective. Using an encoding of formulas and terms by integers, we can view non-standard elements of a model as non-standard formulas and terms. Externally, those formulas can be viewed as having a non-standard depth: for instance, if c is a non-standard element, we may consider the non-standard term $t_c := \text{succ}(\dots(\text{succ}(0))\dots)$ of depth c , or the non-standard

¹This terminology is unfortunate, as the models under consideration are not necessarily ω -models (whose first-order part is ω).

formula

$$x \leq t_c \leftrightarrow (x = 0 \vee x = \text{succ}(0) \vee \dots \vee x = t_c).$$

For every $n \in \omega$, the existence of a well-behaved universal Σ_n -formula is provable in IS_1 (see [57, Corollary I.1.76]), and even in $\text{IS}_0 + \text{exp}$ (a fact that will be used in Section 9.1.3). Hence, from this proof, we obtain for every $n \in \omega$, a standard formula $\varphi_n(x, y)$ such that, if $M \models \text{IS}_0 + \text{exp}$, then every (possibly non-standard) Σ_n -formula $\psi(x)$ is equivalent in M to $\varphi_n(x, e)$ for some $e \in M$, even if externally ψ may be seen as having a non-standard depth.

Allowing first-order parameters is also consistent with our definition of the various schemes in Part I. For instance, the comprehension scheme (Definition 5.3.1) allowed formulas with free variables (first- and second-order) before taking the universal closure. As a result, in a model (M, S) of RCA_0 (hence a model of Δ_1^0 -comprehension), we have $\{a\} \in S$ for every $a \in M$.

In Part I, Σ_n - and Π_n -formulas were introduced, along with the corresponding Σ_n and Π_n sets. When relativizing these definitions, we defined the Σ_n^0 - and Π_n^0 -formulas by allowing free set parameters. However, we did not introduce the corresponding Σ_n^0 and Π_n^0 sets, since the notion is not meaningful without specifying the sets relative to which the definition is taken (every set X is $\Sigma_0(X)$).

For second-order models $\mathcal{M} = (M, S)$, we will define the sets that are Σ_n or Π_n relative to that model in order to measure the complexity of sets not belonging to S that will then be added to $\mathcal{M}$ to construct ω -extensions.

This leads to the following definitions of arithmetical sets relative to a model:

Definition 9.1.2. Let $n \in \omega$. Given a first-order structure M , a subset $A \subseteq M$ is said to be $\Sigma_n(M)$ (resp. $\Pi_n(M)$) if there exists a Σ_n - (resp. Π_n -) formula $\varphi(x, \vec{y})$ and parameters $\vec{b} \in M$ such that $A = \{a \in M : M \models \varphi(a, \vec{b})\}$. A set is $\Delta_n(M)$ if it is both $\Sigma_n(M)$ and $\Pi_n(M)$. As M will always be clear in the context, we will simply speak of Σ_n, Π_n and Δ_n sets.

These definitions can again be relativized: given a collection $C_0, C_1, \dots$ of subsets of M , one can define the $\Sigma_n(C_0, C_1, \dots)$ (resp. $\Pi_n(C_0, C_1, \dots)$ and $\Delta_n(C_0, C_1, \dots)$) sets accordingly.

For a second-order structure $\mathcal{M} = (M, S)$, a subset $A \subseteq M$ is said to be $\Sigma_n(\mathcal{M})$ (resp. $\Pi_n(\mathcal{M})$) if it is Σ_n (resp. Π_n) relative to the elements of S . In other words, if there exists some Σ_n^0 - (resp. Π_n^0 -) formula $\varphi(x, \vec{y}, \vec{Y})$ and parameters $\vec{b} \in M$ and $\vec{B} \in S$ such that $A = \{a \in M : M \models \varphi(a, \vec{b}, \vec{B})\}$. A set is $\Delta_n(\mathcal{M})$ if it is both $\Sigma_n(\mathcal{M})$ and $\Pi_n(\mathcal{M})$. If $\mathcal{M}$ is clear from the context, we instead speak of Σ_n^0, Π_n^0 and Δ_n^0 sets.

This allows us to mix the two notations: for $\mathcal{M} = (M, S)$ a second-order model, and $A, C_0, C_1, \dots \subseteq M$ sets (not necessarily belonging to S). We say that A is $\Sigma_n^0(C_0, C_1, \dots)$ (resp. $\Pi_n^0(C_0, C_1, \dots)$ and $\Delta_n^0(C_0, C_1, \dots)$) if it can be defined accordingly using parameters from M, S or from the C_i .

In a model of Δ_1^0 -comprehension, $\mathcal{M} = (M, S)$, the $\Sigma_n(\mathcal{M})$ and $\Pi_n(\mathcal{M})$ sets, for $n \geq 1$, need not belong to S . Nevertheless, the theory $\text{IS}_0^0 + \text{exp}$ is strong enough to prove the existence of a well-behaved universal Σ_n^0 -formula, making it possible to quantify over and reason about such sets internally.

The amount of induction satisfied by $\mathcal{M}$ determines how well-behaved these sets are.

9.1.2 When induction fails

In a non-standard model $\mathcal{M} = (M, S)$, some subsets $A \subseteq M$ satisfy properties that make it clear (internally) that they are not from a standard model.

This is the case for example with the set ω of standard integers, as it is a non-empty, bounded set with no maximal element. If we were to have $\omega \in S$, then $\mathcal{M} \not\models \text{IS}_0^0$,

as the formula $x \in \omega$ is Σ_0 with parameters from $\mathcal{M}$, and we have $\mathcal{M} \models 0 \in \omega$, $\mathcal{M} \models \forall a, a \in \omega \rightarrow \text{succ}(a) \in \omega$ but $\mathcal{M} \not\models \forall a, a \in \omega$. More generally, if ω were to be $\Sigma_n(\mathcal{M})$, then $\mathcal{M} \not\models \text{I}\Sigma_n^0$. This led to the definition of a cut.

Definition 9.1.3 (Cut). A *cut* in a model M of first-order arithmetic is a nonempty subset $I \subseteq M$ which is closed under successor, and is an initial segment of M , that is, if $a \in I$ and $b \leq a$, then $b \in I$.
A cut $I \subseteq M$ is *proper* if $I \neq M$.

Cuts will play an important role in Part IV of this thesis, where we modify the first-order part of the models in our constructions by restricting it to a cut. For now, we will use cuts as a way to witness the failure of induction in our models.

Proposition 9.1.4 (see e.g., [57, Observation IV.1.15]). *Let $n \in \omega$ and let M (resp. $\mathcal{M} = (M, S)$) be a first-order (resp. second-order) structure such that $M \models \text{I}\Sigma_n$ (resp. $\mathcal{M} \models \text{I}\Sigma_n^0$). Then, there is no Σ_n - (resp. Σ_n^0 -) definable proper cut. Conversely, if $M \models \text{B}\Sigma_n + \neg\text{I}\Sigma_n$ (resp. $\mathcal{M} \models \text{B}\Sigma_n^0 + \neg\text{I}\Sigma_n^0$), then there exists some Σ_n - (resp. Σ_n^0 -) definable proper cut I .*

Another property of non-standard models satisfying a certain amount of induction is the *overspill principle*. Since the standard integers ω cannot be defined by a formula for which induction is available, showing that a given property holds for every standard integer is enough to conclude that it must also hold for some non-standard element.

Proposition 9.1.5 (overspill). *Let $n \in \omega$ and let M (resp. $\mathcal{M} = (M, S)$) be a first-order (resp. second-order) non-standard structure such that $M \models \text{I}\Sigma_n$ (resp. $\mathcal{M} \models \text{I}\Sigma_n^0$). Then, for every Σ_n - (resp. Σ_n^0 -) formula $\varphi(x)$ such that $M \models \varphi(n)$ (resp. $\mathcal{M} \models \varphi(n)$) for every $n \in \omega$, there exists some non-standard element $c \in M \setminus \omega$ such that $M \models \varphi(c)$ (resp. $\mathcal{M} \models \varphi(c)$).*

More generally, for a subset $A \subseteq M$ to be well-behaved, one would like it to be *piecewise-coded* (also known in the literature as *amenable* or *regular*); that is, if for every $a \in M$, the set $A \upharpoonright a = \{x \in A : x < a\}$ is M -finite. In particular, in a model of $\text{I}\Sigma_0^0$, no proper cut is piecewise-coded.

Proposition 9.1.6. *Modulo $\text{I}\Sigma_0^0 + \text{exp}$, we have the following characterizations, for $n \geq 1$:*

- $\text{I}\Sigma_n^0$ is equivalent to the fact that every Σ_n^0 set is piecewise-coded (hence every Π_n^0 also is).
- $\text{I}\Delta_n^0$ (or equivalently $\text{B}\Sigma_n^0$ by Proposition 3.4.4) is equivalent to the fact that every Δ_n^0 set is piecewise-coded.

We will only work with piecewise-coded sets.

9.1.3 RCA_0^*

The theory RCA_0^* is a weakening of RCA_0 that is strong enough to formalize basic computability-theoretic constructions. RCA_0^* is sometimes more convenient to use as a base theory for reverse mathematics, as shown by Belanger [6].

Definition 9.1.7 (RCA_0^* (Simpson and Smith [129])). RCA_0^* denotes the theory $\text{I}\Delta_1^0$ together with the Δ_1^0 -comprehension axiom and exp , the statement asserting the totality of the exponential (as defined in Chapter 3).

In other words, RCA_0^* is obtained from RCA_0 by replacing $\text{I}\Sigma_1^0$ with $\text{I}\Delta_1^0 + \text{exp}$ (or equivalently by $\text{B}\Sigma_1^0 + \text{exp}$ by a relativization of Proposition 3.4.4).

By Simpson and Smith [129, Corollary 4.4], the first-order part of RCA_0^* is exactly $\text{B}\Sigma_1 + \text{exp}$.

RCA_0^* is sufficient to prove the existence of well-behaved universal Σ_n^0 -formulas for every $n \in \omega$. Furthermore, such formulas can be assumed to be in Kleene normal form. Let $\varphi(x, Y, y) := \exists s \psi(s, x, Y \upharpoonright s, y)$ be the universal Σ_1^0 -formula. Rather than choosing a specific model of computation (e.g., Turing machine, lambda calculus...) and formalizing it in the language of second-order arithmetic, we will directly define $\Phi_y^{Y \upharpoonright s}(x)[s] \downarrow := \psi(s, x, Y \upharpoonright s, y)$ and $\Phi_y^Y(x) \downarrow := \varphi(x, Y, y)$. The other computability-theoretic notations (Turing jump, reduction, join...) are then defined accordingly.

Thus, the first level of the (relativized) Post theorem holds essentially by definition: $A \leq_T B$ if and only if A is $\Delta_1(B)$. The remaining task is to determine the amount of induction required to establish the basic properties of $\leq_T$ and of the Turing jump.

Most of these properties can be understood in terms of the behavior of the levels of the arithmetical hierarchy under composition of formulas. Namely, if a set A is arithmetical and a set B is arithmetical relative to A , what can be said about the complexity of B after substituting a definition of A into the definition of B ?

By a relativization of Proposition 3.4.2, $\text{B}\Sigma_1^0$ proves that the classes of Σ_1^0 - and Π_1^0 -formulas are closed under bounded quantifications. This can then be used to prove the transitivity of $\leq_T$ for the elements inside a model of RCA_0^* (the Δ_1^0 sets). Indeed, if A is $\Delta_1(B)$ and B is $\Delta_1(C)$ then we can obtain a $\Sigma_1(C)$ and a $\Pi_1(C)$ description of A by composing the other descriptions and moving bounded quantifiers across.

Proposition 9.1.8 (See Belanger [6, Lemma 2.2]). *Let $\mathcal{M} = (M, S) \models \text{RCA}_0^*$, and let $A, B, C \in S$. If $A \leq_T B$ and $B \leq_T C$ then $A \leq_T C$.*

More generally, the other basic properties of $\leq_T$ and of the join $\oplus$ hold between the elements of S , but not necessarily between the elements arithmetically definable from $\mathcal{M}$ (Groszek and Slaman [55, Proposition 4.1] contains a proof that even in RCA_0 , $\leq_T$ cannot be proven to be transitive for Σ_1^0 sets).

In order to go higher in the arithmetical hierarchy, we need the following result, which can be seen as a relativization of Post's theorem.

Proposition 9.1.9 (Post's theorem (see Belanger [6, Lemma 2.4])). *Let $n > 0$, $\mathcal{M} = (M, S) \models \text{RCA}_0$ and $A \in S$. Then the $\Sigma_{n+1}(A)$ sets are exactly the $\Sigma_n(A')$ sets. More generally, if $\mathcal{M} = (M, S) \models \text{RCA}_0^* + \text{I}\Sigma_k^0$ and $A \in S$. Then the $\Sigma_{n+k}(A)$ sets are exactly the $\Sigma_n(A^{(k)})$ sets.*

We can use this result to transform a model of $\text{RCA}_0^* + \text{B}\Sigma_n^0$ into one of RCA_0^* .

Proposition 9.1.10. *Let $n > 0$, and let $\mathcal{M} = (M, S) \models \text{RCA}_0^*$. Let T be the collection of the $\Delta_n^0(\mathcal{M})$ sets, and set $\mathcal{N} = (M, T)$. Then $\mathcal{N} \models \text{RCA}_0^*$ if and only if $\mathcal{M} \models \text{RCA}_0^* + \text{B}\Sigma_n^0$.*

In particular, the Turing jump and reduction are well-behaved in a model of $\text{RCA}_0 + \text{B}\Sigma_n^0$ as long as we stay below Δ_n^0 sets.

9.2 Π_1^1 -conservation

From now on, we will only consider second-order models satisfying Δ_1^0 -comprehension. Thus, when adding a solution G of an instance of a problem P in the second-order part of some model $\mathcal{M} = (M, S)$, we also want to add the sets that are computable from G . This leads to the following definition:

Definition 9.2.1. For a second-order structure $\mathcal{M} = (M, S)$ and a subset $G \subseteq M$, we write $\mathcal{M}[G]$ for the second-order structure (M, T) , where T consists of all $\Delta_1^0(G)$ sets.

In particular, $\mathcal{M}[G]$ is an ω -extension of $\mathcal{M}$.

The notation $\mathcal{M}[G]$ parallels the notation $V[G]$ from set-theoretic forcing, which describes the extension of a universe of sets by the addition of a generic set G . When

adding a new element to a model, one must ensure that all objects that can be constructed from it are included as well. In set-theoretic forcing, this is handled by the use of $\mathbb{P}$ -names, which control all the sets definable from G . In our setting, since we are primarily interested in building models satisfying Δ_1^0 -comprehension, we use Turing indexes e to control all the sets Δ_1 -definable from G together with the sets already in S .

A major difference from set-theoretic forcing is that the set G may be definable in the ground model $\mathcal{M} = (M, S)$, as is the case with internal constructions. Indeed, since the theories under consideration only satisfy a limited amount of set comprehension (e.g., Δ_1^0 -comprehension for RCA_0), there may exist sets definable in $\mathcal{M}$ that are not elements of S . By contrast, for models of ZF , this phenomenon does not occur: a generic set G cannot be definable from the ground model, otherwise it would already belong to it by the axiom schema of separation.

For simplicity, we will often work in models $\mathcal{M} = (M, S)$ where the second-order part is composed of exactly the sets that are $\Delta_1(A)$ for some $A \in S$. Such models are said to be *topped*.

Definition 9.2.2 (Topped). A model $\mathcal{M} = (M, S)$ is *topped* if there is some set $A \in S$ such that $Y \leq_T A$ for every $Y \in S$.

If $\mathcal{M} = (M, S)$ is topped by some set $A \in S$, then $\mathcal{M}[G]$ is topped by $A \oplus G$. Being topped is an important restriction, as models $\mathcal{M} = (M, S)$ of principles such as WKL , ADS , EM and RT_2^2 are never topped. Indeed, for every $A \in S$ there exist $\Delta_1(A)$ -instances of these principles with no $\Delta_1(A)$ -solution.

Nevertheless, by successively adding solutions to instances of a Π_2^1 problem P , one can build an infinite sequence of ω -extensions of topped models, whose union yields a model of P in the limit, as shown by the following theorem:

Theorem 9.2.3. *Let $\mathcal{T}$ be a Π_2^1 -axiomatizable theory and let P be a Π_2^1 problem. Suppose that for every countable topped model $\mathcal{M} = (M, S) \models \text{RCA}_0^* + \mathcal{T}$ and every P -instance $I \in S$, there exists a P -solution $G \subseteq M$ to I such that $\mathcal{M}[G] \models \text{RCA}_0^* + \mathcal{T}$. Then $\text{RCA}_0^* + \mathcal{T} + P$ is Π_1^1 -conservative over $\text{RCA}_0^* + \mathcal{T}$.*

Proof. Assume $\text{RCA}_0^* + \mathcal{T} \not\models \forall X \varphi(X)$ for φ an arithmetic formula. Then by the completeness theorem and the downward Löwenheim–Skolem theorem (see Theorem 5.2.6 and Theorem 5.2.8), there exists a countable model $\mathcal{M} = (M, S) \models \text{RCA}_0^* + \mathcal{T} + \neg \varphi(A)$ for some $A \in S$. By only keeping the $\Delta_1(A)$ -definable sets, we may further assume that $\mathcal{M}$ is topped by A .

From our assumption, we can build a sequence (indexed by the standard integers)

$$\mathcal{M} = \mathcal{M}_0 \subseteq \mathcal{M}_1 = (M, S_1) \subseteq \mathcal{M}_2 = (M, S_2) \subseteq \dots$$

of countable topped models of $\text{RCA}_0^* + \mathcal{T}$ such that, fixing a bijection $\langle \cdot, \cdot \rangle : \omega^2 \rightarrow \omega$, if $s = \langle i, k \rangle$ then $\mathcal{M}_{s+1}$ contains a solution to the i -th instance of P in $\mathcal{M}_k$ (for some fixed bijection between S_k and ω).

Let $\mathcal{N} := (M, \bigcup_{n < \omega} S_n)$, then $\mathcal{N} \models P$ as every instance of P in $\mathcal{N}$ belongs to some S_n , and hence has a solution in some S_m by construction. Since $A \in S_0 \subseteq \bigcup_{n < \omega} S_n$, $A \in \mathcal{N}$. Moreover, since $\mathcal{M} \models \neg \varphi(A)$ and the first-order part of $\mathcal{N}$ is the same as that of $\mathcal{M}$, we also have $\mathcal{N} \models \neg \varphi(A)$.

Let $\theta(X, Y)$ be an arithmetical formula with two free set parameters X, Y . If $\mathcal{N} \not\models \forall X \exists Y \theta(X, Y)$, then there exists $B \in \bigcup_{n < \omega} S_n$ such that $\mathcal{N} \models \forall Y \neg \theta(B, Y)$. Hence, for sufficiently large $n \in \omega$ such that $B \in S_n$, we have $\mathcal{M}_n \models \forall Y \neg \theta(B, Y)$. Taking the contrapositive, any Π_2^1 sentence true in every $\mathcal{M}_n$ is also true in $\mathcal{N}$. Since both RCA_0^* and $\mathcal{T}$ are Π_2^1 -axiomatizable, and as $\mathcal{M}_n \models \text{RCA}_0^* + \mathcal{T}$ for every $n \in \omega$, it follows that $\mathcal{N} \models \text{RCA}_0^* + \mathcal{T}$.

Thus, $\mathcal{N} \models \text{RCA}_0^* + \mathcal{T} + \text{P} + \neg \forall X \varphi(X)$. So $\text{RCA}_0^* + \mathcal{T} + \text{P} \not\models \forall X \varphi(X)$, yielding the Π_1^1 -conservation. $\square$

It is with similar constructions that the Π_1^1 -conservation results mentioned in Part I have been obtained. Notably, this is the case for: the Π_1^1 -conservation of WKL_0 over RCA_0 (Theorem 5.3.10), of RT_2^2 over $\text{RCA}_0 + \text{IS}_2^0$ ([19]), of ADS over $\text{RCA}_0 + \text{BS}_2^0$ (Chong, Slaman, and Yang [22]).

Theorem 9.2.3 provides an example of an external construction. Indeed, the sequence $(\mathcal{M}_i)_{i \in \omega}$ of ω -extensions is indexed by the standard integers, and no assumptions are made about the definability of the sequence.

It is when proving the assumption that every countable topped model $\mathcal{M} = (M, S) \models \text{RCA}_0 + \mathcal{T}$ can be extended to a model $\mathcal{M}[G] \models \text{RCA}_0 + \mathcal{T}$ containing a solution to an instance of P that internal constructions will come into play.

In the search for the first-order part of RT_2^2 , we will mostly work inside models of BS_2^0 (and not necessarily of IS_2^0) as $\text{RCA}_0 + \text{RT}_2^2 \vdash \text{BS}_2^0$. By the results of Section 9.1.3, the Δ_2^0 sets of such models are well-behaved.

In the case of WKL , we can make a purely internal construction thanks to the low basis theorem (Theorem 4.2.7).

Proposition 9.2.4 (Clote et al. [27]). *WKL is Π_1^1 -conservative over $\text{RCA}_0 + \text{BS}_2^0$.*

Proof. We want to apply Theorem 9.2.3.

Let $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{BS}_2^0$ be topped by some A and let $T \in S$ be an instance of WKL . By a relativized form of the low basis theorem (Theorem 4.2.7), which is provable in RCA_0 , there exists some path $P \subseteq M$ in T such that $A' \geq_T (P \oplus A)'$. Then, as $\mathcal{M} \models \text{RCA}_0 + \text{BS}_2^0$, we have $\mathcal{M}[A'] \models \text{RCA}_0^*$ (by the direct implication in Proposition 9.1.10), hence $\mathcal{M}[P \oplus A] = \mathcal{M}[P] \models \text{RCA}_0 + \text{BS}_2^0$ (by the reverse implication in Proposition 9.1.10).

Thus, by Theorem 9.2.3, WKL is Π_1^1 -conservative over $\text{RCA}_0 + \text{BS}_2^0$. $\square$

In order to obtain the Π_1^1 -conservation of WKL over RCA_0 (as stated in Theorem 5.3.10), it is possible to work with a weaker notion than Turing reducibility called weak truth-table reducibility (see Belanger [6, Corollary A.2]) which is well-behaved for Δ_2^0 sets in RCA_0 .

For RT_2^2 and many other Ramsey-like statements, no such direct construction is possible. Indeed, there exist computable instances of RT_2^2 with no Σ_2 (and in particular no Δ_2) solutions (see Proposition 6.0.9). In particular, no such low basis theorem can be proven for RT_2^2 . Thus, our constructions will not be purely internal.

Fortunately, by a relativization of Theorem 6.0.10, for every instance X of RT_2^2 , any set of PA degree over X' computes a solution for X . And, thanks to the following result, sets of PA degree can be added for free in a model of RCA_0^* .

Just as WKL_0 is the theory $\text{RCA}_0 + \text{WKL}$, we consider the theory $\text{WKL}_0^* := \text{RCA}_0^* + \text{WKL}$. Then, Harrington's result (Theorem 5.3.10) that WKL_0 is Π_1^1 -conservative over RCA_0 also holds between RCA_0^* and WKL_0^* .

Theorem 9.2.5 (Simpson–Smith [129, Theorem 4.6]). *Every countable model of RCA_0^* can be ω -extended into a model of WKL_0^* . Hence, WKL is Π_1^1 -conservative over RCA_0^* .*

In particular, for every countable model $\mathcal{M} = (M, S) \models \text{RCA}_0^*$ and every $A \in S$, there exists some $B \subseteq M$ such that $\mathcal{M}[B] \models \text{RCA}_0^*$ and such that every infinite A -computable binary tree has a B -computable path. Thus, by Proposition 9.1.10, for every $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{BS}_2^0$ and every instance X of RT_2^2 in S , there exists some $B \subseteq M$ such that $\mathcal{M}[B] \models \text{RCA}_0^*$ and such that every infinite X' -computable binary tree has a B -computable path (i.e., B is of PA degree over X').

Thus, to establish the Π_1^1 -conservation of RT_2^2 over $\text{RCA}_0 + \text{B}\Sigma_2^0$, it suffices to prove (a relativization of) Theorem 6.0.10 within $\text{RCA}_0 + \text{B}\Sigma_2^0$ and then conclude the argument as in Proposition 9.2.4. However, no such proof is currently known, and the question remains open.

In fact, a result of Fiori-Carones et al. [46] shows that this condition is also necessary: RT_2^2 is Π_1^1 -conservative of RT_2^2 over $\text{RCA}_0 + \text{B}\Sigma_2^0$ if and only if $\text{RCA}_0 + \text{B}\Sigma_2^0$ proves (a relativization of) Theorem 6.0.10.

The relativization of Theorem 6.0.10 is a $\forall\Pi_5^0$ -sentence (i.e., of the form $\forall X\varphi(X)$ with $\varphi \in \Pi_5^0$) and is provable in $\text{RCA}_0 + \text{RT}_2^2$, giving us the following result:

Theorem 9.2.6 (Fiori-Carones et al. [46, Corollary 5.2]). *RT_2^2 is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$ if and only if it is $\forall\Pi_5^0$ -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$.*

In Chapter 14, we will prove RT_2^2 to be $\forall\Pi_4^0$ -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$, thereby bringing us one step closer to resolving this question.

Before turning to this, we will work in Chapter 10 with theories stronger than $\text{RCA}_0 + \text{B}\Sigma_2^0$ to obtain Π_1^1 -conservation results yielding upper bounds on the first-order part of $\text{RCA}_0 + \text{B}\Sigma_2^0$.

Chapter 10

Π_1^1 -conservation of RT_2^2 and well-foundedness

In this chapter, we give an upper bound for the first-order part of $\text{RCA}_0 + \text{RT}_2^2$ by proving that Ramsey's theorem for pairs and two colors is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$ and over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \bigcup_{n \in \omega} \text{WF}^0(\omega_n)$. Here $\bigcup_{n \in \omega} \text{WF}^0(\omega_n)$ denotes the theory consisting of the axioms $\text{WF}^0(\omega_n)$ for each integer $n \in \omega$. It is strictly weaker than $\text{WF}^0(\epsilon_0)$ (which, over RCA_0 , is equivalent to $\forall n \text{WF}^0(\omega_n)$), since there exist non-standard models of $\text{RCA}_0 + \text{B}\Sigma_2^0 + \bigcup_{n \in \omega} \text{WF}^0(\omega_n)$ in which $\text{WF}^0(\omega_c)$ fails for some non-standard integer c .

Nevertheless, proving that an axiom P is conservative over a given theory doesn't necessarily imply conservation over a stronger theory, so we present both results.

10.1 Introduction

The well-foundedness of ordinals naturally occurs in the proof-theoretic analysis of Ramsey's theorem for pairs. Chong, Slaman, and Yang's proof [23] that RT_2^2 does not imply $\text{I}\Sigma_2^0$ over RCA_0 uses a model of $\text{RCA}_0 + \text{B}\Sigma_2^0 + \neg \text{I}\Sigma_2^0$ which furthermore satisfies a principle called bounded monotone enumeration (BME).

Definition 10.1.1 ([23]). Let E be a procedure that computably enumerates a finitely branching tree. We say that E is a *monotone enumeration* if and only if the following conditions apply to its stage-by-stage behavior.

- (1) The empty sequence is enumerated by E during stage 0.
- (2) Only finitely many sequences are enumerated by E during any stage.
- (3) Suppose $T[s]$ is the tree enumerated by E at the end of stage s . Then any sequence enumerated by E during stage $s + 1$ must extend some terminal node of $T[s]$.

A monotone enumeration E is *b-bounded* if every node enumerated by E has depth at most b . E is said to be *bounded* if it is b -bounded for some $b \in \mathbb{N}$.

Statement 10.1.2 ($\text{BME}_\star$). $\text{BME}_\star$ is the statement "Every bounded monotone enumeration is finite".

The statement BME is an iterated version of $\text{BME}_\star$. Its formal definition is rather technical and will not be used within this thesis. Kreuzer and Yokoyama [84] proved that $\text{BME}_\star$ is equivalent to $\text{WF}^0(\omega^\omega)$ over $\text{I}\Sigma_1^0$, and that BME is equivalent to $\bigcup_{n \in \omega} \text{WF}^0(\omega_n)$.

On the other hand, Patey and Yokoyama [115] proved that RT_2^2 is $\forall \Pi_3^0$ -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$. It follows that RT_2^2 does not imply $\text{BME}_\star$ over RCA_0 .

(see, for instance, the discussion surrounding Theorem 3.6.3), and therefore that $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$ and $\text{RCA}_0 + \text{B}\Sigma_2^0 + \bigcup_{n \in \omega} \text{WF}^0(\omega_n)$ cannot be the first-order part of $\text{RCA}_0 + \text{RT}_2^2$, but are only strict upper bounds.

In this chapter, we show that the separation of Chong, Slaman, and Yang [23] of RT_2^2 from $\text{I}\Sigma_2^0$ depends only on the assumption that the original model satisfies $\text{B}\Sigma_2^0 + \text{BME}$, by proving the following conservation theorems.

Main Theorem 10.1.3. $\text{RT}_2^2 + \text{WKL}$ is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$.

Main Theorem 10.1.4. $\text{RT}_2^2 + \text{WKL}$ is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \bigcup_{k \in \omega} \text{WF}^0(\omega_k)$.

These theorems strengthen a result from Kołodziejczyk and Yokoyama [80] who showed that RT_2^2 is Π_4^0 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{BME}$. By Cholak, Jockusch, and Slaman [19], the result holds when $\text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$ is replaced by $\text{I}\Sigma_2^0$. Thus, conservations over $\text{RCA}_0 + (\text{WF}^0(\epsilon_0) \vee \text{I}\Sigma_2^0)$ and over $\text{RCA}_0 + (\bigcup_{n \in \omega} \text{WF}^0(\omega_n) \vee \text{I}\Sigma_2^0)$ are the best known results.

We will make use of the decomposition of RT_2^2 into $\text{COH} + \text{SEM} + \text{SADS}$ and prove conservation results about WKL , COH , SEM and SADS separately to obtain our main theorems. By Theorem 9.2.3, it will be sufficient to show that for every problem $P \in \{\text{WKL}, \text{COH}, \text{SEM}, \text{SADS}\}$, every countable topped model $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$ (resp. $\bigcup_{n \in \omega} \text{WF}^0(\omega_n)$), and every instance of P in S , there exists some solution $G \subseteq M$ to that instance so that $\mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$ (resp. $\bigcup_{n \in \omega} \text{WF}^0(\omega_n)$).

10.2 Conservation over $\text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$

Consider a countable model $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$ for some $\alpha \leq \epsilon_0$ (or even for larger ordinals, provided that a sufficiently well-behaved system of ordinal notations is available), topped by a set $Y \in S$ and a P -instance $I \in S$ of some Π_1^1 problem P .

The construction of a solution $G \subseteq M$ to I will be similar to that in the proof of Theorem 8.5.6, but inside a non-standard model. Each problem P will have its own notion of a condition (like Definition 8.5.8 for PFS^{low}) that corresponds to an approximation of G at a given stage of the construction. We will then construct an infinite decreasing sequence of conditions $c_0 \geq c_1 \geq \dots$ that will uniquely specify G and make it so that it satisfies additional constraints (such as not computing an infinite decreasing sequence of ordinals smaller than α) ensuring that $\mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$.

The two main differences compared to the proof of Theorem 8.5.6 are, first, that the c_s are now indexed by nonstandard integers (as we are doing an internal construction), and second, that we must keep track of the computational strength needed to obtain c_{s+1} from c_s , so that the set G will not be too complex and can be added to $\mathcal{M}$. To do so, we require the sequence (c_s) to be P -computable for some set $P \subseteq M$ such that $\mathcal{M}[P] \models \text{RCA}_0^*$ (in the case of COH and SEM , P will be the set of PA degree over Y' given by Theorem 9.2.5, and for WKL and SADS , it will be Y'). We will also P -compute $(G \oplus Y)'$ simultaneously, so that every $\Delta_2(G \oplus Y)$ property is $\Delta_1(P)$ and hence that $\mathcal{M}[G \oplus Y] = \mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0$ by Proposition 9.1.10.

10.2.1 Shore blocking

One difficulty with this approach is that, even if we succeed in proving the existence of a P -computable procedure producing c_{s+1} from c_s for every $s \in M$, there is still no guarantee that the construction can actually be carried out for all $s \in M$.

Indeed, the formula $\varphi(s)$ expressing that the construction can be pursued for s steps is $\Sigma_1(P)$: “there exists a finite sequence $c_0, \dots, c_s$ of conditions such that c_0 is our initial condition and, for every $i < s$, c_{i+1} is obtained from c_i via the given P -computable procedure”. However, as $\mathcal{M}[P]$ does not necessarily satisfy $\mathsf{I}\Sigma_1^0$, the set $\{s : \mathcal{M} \models \varphi(s)\}$ may form a proper cut.

If we look at the values of the (elements of M encoding the) c_s for s ranging over that cut, then these values must necessarily be M -unbounded (i.e., there is no $k \in M$ such that the code of c_s is smaller than k for all such s). Indeed, if such a bound existed, then the property $\varphi(s)$ would be $\Delta_1(P)$, since the unbounded existential quantifier could be replaced by a bounded one. And, as $\mathcal{M}[P] \models \mathsf{I}\Delta_1^0$, $\varphi(s)$ could not define a proper cut.

In our constructions, having the c_s of M -unbounded sizes guarantees that they define a unique set G . The issue, however, is that G may then fail to satisfy all the desired requirements. For example, suppose we had a list $(\mathcal{R}_s)_{s \in M}$ of requirements and intended to satisfy $\mathcal{R}_s$ at stage s of the construction, then the requirements $\mathcal{R}_s$ for s not in the cut will not be satisfied.

A way around this problem is to use a technique known as *Shore blocking* (where “blocking” is meant in the sense of grouping requirements into blocks). Rather than searching at stage s for some c_{s+1} that satisfies only one additional requirement, we instead search for a condition c_{s+1} satisfying every requirement whose indices are smaller than the code of c_s . This makes the construction more complex, as satisfying a block consisting of a potentially nonstandard number of requirements is more difficult than satisfying just one.

In this way, it is no longer essential that the construction proceeds along all of M : the conditions c_s that are defined are of sizes unbounded in M , hence all the requirements are eventually satisfied.

10.2.2 Well-foundedness and Shore blocking

For $\mathcal{M}[G]$ to be a model of $\mathsf{WF}^0(\alpha)$, we want, for every $e \in M$, to satisfy the requirement that $\Phi_e^{G \oplus Y}$ is not an infinite decreasing sequence of ordinals smaller than α . In this section we will show that satisfying a block of such requirements is not harder than satisfying just one, by grouping several decreasing sequences of ordinals into one.

Instead of using the standard sums and products over ordinals, the natural (Hessenberg) sums and products will be more convenient.

Definition 10.2.1 (Natural sum and product). Let $\alpha = \omega^{\gamma_{\ell-1}}n_{\ell-1} + \dots + \omega^{\gamma_0}n_0$ and $\beta = \omega^{\gamma_{\ell-1}}m_{\ell-1} + \dots + \omega^{\gamma_0}m_0$ be two ordinals smaller than ϵ_0 , written in their Cantor normal forms (except that we allow the n_i and m_i to be equal to 0 in order to write α and β using the same exponents γ_i).

We define the natural sum of α and β (denoted by $\alpha \oplus \beta^a$) to be the following ordinal (that does not depend on the choice of the family (γ_i)):

$$\omega^{\gamma_{\ell-1}}(n_{\ell-1} + m_{\ell-1}) + \dots + \omega^{\gamma_0}(n_0 + m_0)$$

It is easy to check that, contrary to the standard addition of ordinals, $\oplus$ is associative and commutative. For these reasons, we can unambiguously write $\bigoplus_{i \in I} \alpha_i$ for the natural sum of a finite amount of ordinals.

We also define the natural product of α and β (denoted $\alpha \otimes \beta$) to be equal to:

$$\bigoplus_{\substack{i < \ell \\ j < \ell}} \omega^{\gamma_i \oplus \gamma_j} n_i m_j$$

^aThe symbol $\oplus$ is also used to denote the Turing join; however, since its meaning is unambiguous from the context, we use the same notation.

We will only encounter the case where $\beta = k = \omega^0 k$ for some $k \in M$, in which case $\alpha \otimes k = \omega^{\gamma_{\ell-1}}(n_{\ell-1}k) + \dots + \omega^{\gamma_0}(n_0k)$. Note that $\alpha \otimes k \leq \alpha \times \omega$ for every $k \in M$.

The natural sum and product of ordinals enjoy several desirable properties that the standard ordinal sum and product lack, such as commutativity, associativity, and distributivity of $\otimes$ over $\oplus$. The property relevant for our purposes is the following: $\oplus$ is strictly increasing in both arguments (which is not the case for $+$, as $1 + \omega = \omega$).

Proposition 10.2.2. *For all $\beta < \epsilon_0$, and $\alpha' < \alpha < \epsilon_0$: $\alpha' \oplus \beta < \alpha \oplus \beta$ and $\beta \oplus \alpha' < \beta \oplus \alpha$.*

Proof. Clear from the definition of $\oplus$. $\square$

Lemma 10.2.3. *Let $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{WF}^0(\alpha)$ where α is an ordinal smaller than ϵ_0 . Then $\mathcal{M} \models \forall k \text{WF}^0(\alpha \otimes k)$.*

Proof. RCA_0 proves that the product of two well-orders is a well-order, hence $\mathcal{M} \models \text{WF}^0(\alpha \times \omega)$. Since $\alpha \otimes k \leq \alpha \times \omega$ for every $k \in M$, we get the desired result. $\square$

We can now use these results to group a block of k decreasing sequence of ordinals smaller than α into one of ordinals smaller than $\alpha \otimes k$.

Lemma 10.2.4. *Consider a model $\mathcal{M} = (M, S)$ and $\alpha \in M$ an ordinal smaller than ϵ_0 . For every $k \in M$, there is a Turing functional Γ_k such that, letting $\beta < \alpha$ be the largest ordinal less than α of index at most k (for our encoding of the ordinals smaller than ϵ_0), for every $X \in 2^M$ such that $\mathcal{M}[X] \models \text{RCA}_0^*$, if there is some $e < k$ such that Φ_e^X is an M -infinite decreasing sequence of ordinals smaller than β , then Γ_k^X is an M -infinite decreasing sequence of ordinals smaller than $\beta \otimes k$. Moreover, an index of Γ_k can be found computably in k and in an index for α .*

Proof. Let Γ_k^X be defined as follows: on input n , find the smallest value $s \in M$ such that there exist $n_0, \dots, n_{k-1} \in M$ with $n_0 + \dots + n_{k-1} = n$ and, such that, for every $e < k$, $\Phi_e^X(0)[s] \downarrow, \dots, \Phi_e^X(n_e - 1)[s] \downarrow$ with $\Phi_e^X(0), \dots, \Phi_e^X(n_e - 1)$ a decreasing sequence of ordinals smaller than β . Then, output $\Phi_0^X(n_0 - 1) \oplus \dots \oplus \Phi_{k-1}^X(n_{k-1} - 1)$ (with the convention that $\Phi_e^X(-1) = \beta$). If several such tuples $(n_0, \dots, n_{k-1})$ are found, pick the one for which the quantity $\Phi_0^X(n_0 - 1) \oplus \dots \oplus \Phi_{k-1}^X(n_{k-1} - 1)$ is minimal.

Suppose that X is such that $\mathcal{M}[X] \models \text{RCA}_0^*$ and that there is an $e < k$ such that Φ_e^X is an infinite decreasing sequence of ordinals smaller than β , then Γ_k^X is also total as we will always find such tuples $(n_0, \dots, n_{k-1})$.

Then, by definition of Γ_k^X , $\Gamma_k^X(0), \Gamma_k^X(1), \dots$ is an M -infinite decreasing sequence of ordinals smaller than $\beta \otimes k$. $\square$

10.2.3 Π_1^1 -conservation of COH

We now turn to the proof that $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha) + \text{COH}$ is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$ for every ordinal $\alpha \leq \epsilon_0$. In particular, letting $\alpha = \omega$, we reprove the theorem of Chong, Slaman, and Yang [22] stating that COH is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$. Note that the original proof by Chong, Slaman, and Yang uses a case analysis, depending on whether the model satisfies $\text{I}\Sigma_2^0$ or not. Belanger [6] gave a more direct proof using the characterization of p-cohesive sets as those whose jump is PA over $0'$ (see Jockusch and Stephan [69]). Our proof follows the ideas from Belanger, but with an even more direct construction based on an instance-wise correspondence between Y -computable instances of COH and Y' -computable instances of WKL .

Proposition 10.2.5. *Consider a countable model $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$ topped by a set $Y \in S$, where $\alpha \leq \epsilon_0$ is an ordinal. Then for every countable sequence of sets $\vec{R} \in S$ and every $P \gg Y'$ (i.e., P is PA over Y') such that $\mathcal{M}[P] \models \text{RCA}_0^*$, there exists $G \subseteq M$ such that*

1. G is $\vec{R}$ -cohesive;
2. $P \geq_T (G \oplus Y)'$;
3. $\mathcal{M}[G] \models \mathbf{RCA}_0 + \mathbf{B}\Sigma_2^0 + \mathbf{WF}^0(\alpha)$.

Proof. Given $\vec{R} = R_0, R_1, \dots$ and $\sigma \in 2^{<M}$, let $\vec{R}_\sigma = \bigcap_{\sigma(i)=0} \mathbb{N} \setminus R_i \cap \bigcap_{\sigma(i)=1} R_i$. Let T be the $\Sigma_1(Y)$ tree $\{\sigma \in 2^{<M} : \exists x(x > |\sigma| \wedge x \in \vec{R}_\sigma)\}$. Note that a node σ is extendible to an infinite path in T if and only if $\vec{R}_\sigma$ is infinite. The tree T is $\Delta_1(Y')$, so since $P \gg Y'$, there is a $\Delta_1(P)$ set $Q \in [T]$.

Definition 10.2.6. A *condition* is a pair of M -finite binary strings (σ, ρ) such that $\rho < Q$. A condition (σ_2, ρ_2) *extends* a condition (σ_1, ρ_1) , and we write $(\sigma_2, \rho_2) \leq (\sigma_1, \rho_1)$, if $\sigma_2 \geq \sigma_1$, $\rho_2 \geq \rho_1$ and $\sigma_2 - \sigma_1 \subseteq \vec{R}_{\rho_1}$ (where $\sigma_2 - \sigma_1$ is the difference between the sets corresponding to the two sequences).

A set $X \subseteq M$ is *compatible* with a condition (σ, ρ) , if $\sigma < X \subseteq \sigma \cup \vec{R}_\rho$.

One can think of a condition (σ, ρ) as a so-called Mathias condition $(\sigma, \vec{R}_\rho)$, where σ is an initial segment of the set G under construction, and $\vec{R}_\rho$ serves as a reservoir restricting the possible future elements that may be added to G . Requiring that $\rho < Q$ ensures that the reservoir $\vec{R}_\rho$ is M -infinite.

When increasing the length of ρ , we make progress in cohesiveness, as every infinite subset of $\vec{R}_\rho$ is cohesive for $R_0, \dots, R_{|\rho|}$. Note that being a condition is a $\Delta_1(Q)$ predicate, as it simply requires checking that $\sigma < Q$, while the extension relation is $\Delta_1(Y)$. In particular, both are $\Delta_1(Q \oplus Y)$, so $\Delta_1(P)$.

In the particular case of finite sets, we say that a binary sequence τ is compatible with a condition (σ, ρ) if $\sigma \leq \tau$ and $\tau \subseteq \sigma \cup \vec{R}_\rho$. The compatibility relation for such sequences is a $\Delta_1(Y)$ predicate.

To control various properties of G throughout the construction, we introduce the following forcing relations, which allow us to control $\Sigma_1^0(G)$ and $\Pi_1^0(G)$ properties.

Definition 10.2.7. Let (σ, ρ) be a condition, and $e, x \in M$, we write:

1. $(\sigma, \rho) \Vdash \Phi_e^{G \oplus Y}(x) \downarrow$ if $\Phi_e^{\sigma \oplus Y}(x) \downarrow$.
2. $(\sigma, \rho) \Vdash \Phi_e^{G \oplus Y}(x) \uparrow$ if for every τ compatible with (σ, ρ) , $\Phi_e^{\tau \oplus Y}(x) \uparrow$.

We then say that the property $\Phi_e^{G \oplus Y}(x) \downarrow$ or $\Phi_e^{G \oplus Y}(x) \uparrow$ is *forced* by the condition (σ, ρ) .

By definition of $\Vdash$, it is clear that if (σ, ρ) forces some property, then this property holds for every set X compatible with (σ, ρ) .

Note that the relations $(\sigma, \rho) \Vdash \Phi_e^{G \oplus Y}(x) \downarrow$ and $(\sigma, \rho) \Vdash \Phi_e^{G \oplus Y}(x) \uparrow$ are $\Delta_1(Y)$ and $\Pi_1(Y)$, respectively (using the convention that, when the oracle of a Turing functional is only partially specified, the computation is required to halt within a number of steps bounded by the index of the first undefined bit of the oracle, and otherwise diverges by running into an infinite loop).

We need to satisfy three kinds of requirements for every $k \in M$:

- $\mathcal{R}_k$: $(G \oplus Y)' \upharpoonright k$ is decided
- $\mathcal{S}_k$: for every $e < k$, $\Phi_e^{G \oplus Y}$ is not an M -infinite α -decreasing sequence
- $\mathcal{T}_k$: for every $e < k$, $G \subseteq^* R_e$ or $G \subseteq^* \mathbb{N} \setminus R_e$ and there exists some $x > k$ such that $x \in G$.

The requirements $\mathcal{T}_k$ can be satisfied easily by extending a condition (σ, ρ) to a condition $(\sigma \cup \{x\}, \rho')$ where $\rho' < Q$ has length at least k and $x \in \vec{R}_\rho$ is bigger than k . This can be done Q -computably and guarantees that $G \subseteq \sigma \cup \vec{R}_{\rho'}$, and hence that $G \subseteq^* R_e$ or $G \subseteq^* \mathbb{N} \setminus R_e$ for every $e < k$, by definition of $\vec{R}_{\rho'}$. Thus, we get the following:

Lemma 10.2.8. *For every condition (σ, ρ) and every $k \in M$, there exists some extension $(\tau, \rho') \leq (\sigma, \rho)$ such that $(\tau, \rho') \Vdash \mathcal{T}_k$.*

Proof. Since $\rho < Q$, the set $\vec{R}_\rho$ is $\mathcal{M}$ -infinite and there exists some τ compatible with (σ, ρ) and some $x > k$ such that $x \in \tau$, making $(\tau, \rho) \Vdash \exists x \in G, x \geq k$. Then taking $\rho' = Q \upharpoonright k$ ensures that every G compatible with (τ, ρ') will be such that $G \setminus \tau \subseteq \vec{R}_{Q \upharpoonright k}$, hence that $(\tau, \rho') \Vdash \mathcal{T}_k$. $\square$

We will need the following two technical lemmas, which are Shore blocking arguments for satisfying the first two kinds of requirements, namely, deciding the jump and preserving $\text{WF}^0(\alpha)$. In what follows, $(\tau, \rho) \Vdash (G \oplus Y)' \upharpoonright k = \sigma'$ means that $|\sigma'| = k$ and that for every $e < k$, if $\sigma'(e) = 1$ then $(\tau, \rho) \Vdash \Phi_e^{G \oplus Y}(e) \downarrow$ and if $\sigma'(e) = 0$, then $(\tau, \rho) \Vdash \Phi_e^{G \oplus Y}(e) \uparrow$. Thus, this relation is $\Pi_1(Y)$.

Lemma 10.2.9. *Let (σ, ρ) be a condition. For every $k \in M$, there exists some M -finite $\sigma' \in 2^k$ and an extension $(\tau, \rho) \leq (\sigma, \rho)$ such that $(\tau, \rho) \Vdash (G \oplus Y)' \upharpoonright k = \sigma'$.*

Proof. Let W be the set of all $\sigma' \in 2^k$ such that $(\exists \tau > \sigma)(\exists t)(\forall e < k)(\tau - \sigma \subseteq \vec{R}_\rho \wedge |\tau - \sigma| \geq 1 \wedge \sigma'(e) = 1 \rightarrow \Phi_e^{\tau \oplus Y}(e)[t] \downarrow)$. Note that W is $\Sigma_1(Y)$ and hence M -finite as $\mathcal{M} \models \text{BC}\Sigma_1^0$ by a relativization of Proposition 3.5.2. Moreover, W is non-empty, as it contains the string $00 \dots 0$. Let $\sigma' \in W$ be its lexicographically maximal element, and let $\tau \geq \sigma$ witness that $\sigma' \in W$.

We claim that (τ, ρ) forces $(G \oplus Y)' \upharpoonright k = \sigma'$. Fix some $e < k$.

Case (a): $\sigma'(e) = 1$. Then $\Phi_e^{\tau \oplus Y}(e) \downarrow$ (hence for every X compatible with (τ, ρ) , $e \in (X \oplus Y)'$).

Case (b): $\sigma'(e) = 0$. The maximality of σ' ensures that for every μ compatible with (τ, ρ) , $\Phi_e^{\mu \oplus Y}(e) \uparrow$ (hence that for every X compatible with (τ, ρ) , $\Phi_e^{X \oplus Y}(e) \uparrow$ and $e \notin (X \oplus Y)'$). $\square$

Lemma 10.2.10. *Let (σ, ρ) be a condition. For every $k \in M$, there exists some $n \in M$ and an extension $(\tau, \rho) \leq (\sigma, \rho)$ forcing $\Phi_e^{G \oplus Y}$ not to be an α -decreasing sequence of length more than n for every $e < k$.*

Proof. By twisting the Turing functionals, we can assume that for every $e, n \in M$, and every $\rho \in 2^{<M}$, if $\Phi_e^\rho(n) \downarrow$, then

- $\Phi_e^\rho(m) \downarrow$ for every $m < n$
- $\Phi_e^\rho(0), \Phi_e^\rho(1), \dots, \Phi_e^\rho(n)$ is a strictly α -decreasing sequence

Suppose the lemma does not hold. Then in particular, for every M -finite string τ compatible with (σ, ρ) and every $n \in M$, there is some $\tau_1 \geq \tau$ compatible with (σ, ρ) and some $e < k$ such that $\Phi_e^{\tau_1 \oplus Y}(n) \downarrow$.

Build Y -computably an infinite sequence $\tau_0 \leq \tau_1 \leq \tau_2 \leq \dots$ such that for every n ,

- τ_n is compatible with (σ, ρ)
- $\Phi_e^{\tau_n \oplus Y}(n) \downarrow$ for some $e < k$ (which depends on n)

Since $\mathcal{M} \models \text{B}\Sigma_2^0$, $\mathcal{M} \models \forall k \text{RT}_k^1$ (see Theorem 6.0.5), there is an M -infinite set $Z \in \mathcal{M}$ and some $e < k$ such that for every $n \in Z$, $\Phi_e^{\tau_n \oplus Y}(n) \downarrow$.

Then, letting $f(n) = \Phi_e^{\tau_n \oplus Y}(n)$, we have an M -infinite α -decreasing sequence, contradicting $\mathcal{M} \models \text{WF}^0(\alpha)$. $\square$

Note that, as formulated, the requirement $\mathcal{S}_k$ is a $\Pi_2^0(G)$ property. However, once a value of n is fixed, the statement that $\Phi_e^{G \oplus Y}(n) \uparrow$ for every $e < k$ is $\Pi_1^0(G)$. Thus, our formalism allows us to force Π_2^0 properties of G .

Construction. We P -computably construct a decreasing sequence (σ_s, ρ_s) of conditions. In parallel, the set $(G \oplus Y)'$ is approximated by a sequence (σ'_s) of finite binary sequences.

Let $\sigma_0 = \rho_0 = \sigma'_0 = \varepsilon$. Assume that σ_s, ρ_s , and σ'_s have been defined for some $s \in M$ and let

$$k_s = 1 + \max\{|\sigma_s|, |\rho_s|, |\sigma'_s|\}.$$

Using Lemma 10.2.8, Lemma 10.2.9 and Lemma 10.2.10, there exists some extension $(\tau, \rho) \leq (\sigma_s, \rho_s)$ forcing $\mathcal{R}_{k_s}, \mathcal{S}_{k_s}$ and $\mathcal{T}_{k_s}$. Such an extension can be found P -computably. We then let $\sigma_{s+1} = \tau$, $\rho_{s+1} = \rho$ and $\sigma'_{s+1} = \sigma'$ where σ' is the sequence obtained from the fact that $\mathcal{R}_{k_s}$ is forced.

As explained in Section 10.2.1, the σ_s, σ'_s and ρ_s may only be defined on some cut I but are nonetheless of unbounded length in M .

Let $G = \bigcup_{s \in I} \sigma_s$. By construction G is P -computable and, as the $\mathcal{R}_{k_s}$ are satisfied, $(G \oplus Y)'$ is equal to $\bigcup_{s \in I} \sigma'_s$ and is also P -computable, hence $\mathcal{M}[G \oplus Y] = \mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0$. Since the requirements $\mathcal{T}_k$ are satisfied, the set G is $\vec{R}$ -cohesive. Then, as the requirements $\mathcal{S}_k$ are satisfied, $\mathcal{M}[G] \models \text{WF}^0(\alpha)$.

This completes the proof of Proposition 10.2.5. $\square$

Corollary 10.2.11. Consider a countable model $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$ topped by a set $Y \in S$, where $\alpha \leq \epsilon_0$ is an ordinal. Then for every $\vec{R} \in S$, there exists $G \subseteq M$ such that

1. G is $\vec{R}$ -cohesive;
2. $\mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$.

Proof. By Theorem 9.2.5, there exists a set $P \subseteq M$ such that $P \gg Y'$ and $\mathcal{M}[P] \models \text{RCA}_0^*$. The conclusion then follows from Proposition 10.2.5. $\square$

We therefore obtain our first conservation theorem:

Theorem 10.2.12. Fix an ordinal $\alpha \leq \epsilon_0$. COH is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$.

Proof. Immediate by Proposition 10.2.5 and Theorem 9.2.3. $\square$

10.2.4 Π_1^1 -conservation of WKL

We now turn to the proof of Π_1^1 -conservation of WKL. As for COH, in the case $\alpha = \omega$, we recover the well-known conservation theorem of Hájek [56] stating that WKL is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$.

Proposition 10.2.13. Consider a countable model $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$ topped by a set $Y \in S$, where $\alpha \in M$ is an ordinal $\leq \epsilon_0$. Then for every M -infinite tree $T \in S$, there exists $G \subseteq M$ such that

1. G is a path in T ;
2. $Y' \geq_T (G \oplus Y)'$;
3. $\mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$

Proof. We will build the path G using forcing on trees. A condition will be a pair (σ, U) with $U \in S$ an M -infinite Y -computable subtree of T with root σ (i.e., its only elements of length $\leq |\sigma|$ are the prefixes of σ) and we write $(\sigma, U_1) \leq (\tau, U_2)$ if $\tau \leq \sigma$

and $U_1 \subseteq U_2$. In a condition, the tree U is given by a Turing index e such that Φ_e^Y computes it.

Again, a condition (σ, U) is an approximation of the path G being constructed where σ is an initial segment of G and G will be one of the paths of U . A set $X \subseteq M$ is said to be *compatible* with (σ, U) if it belongs to $[U]$.

Definition 10.2.14. Let (σ, U) be a condition and $e, x \in M$, we write:

1. $(\sigma, U) \Vdash \Phi_e^{G \oplus Y}(x) \downarrow$ if $\Phi_e^{\sigma \oplus Y}(x) \downarrow$
2. $(\sigma, U) \Vdash \Phi_e^{G \oplus Y}(x) \uparrow$ if for every $\tau \in U$, $\Phi_e^{\tau \oplus Y}(x) \uparrow$.

Note that the relations $(\sigma, U) \Vdash \Phi_e^{G \oplus Y}(x) \downarrow$ and $(\sigma, U) \Vdash \Phi_e^{G \oplus Y}(x) \uparrow$ are $\Delta_1(Y)$ and $\Pi_1(Y)$, respectively. As before, the definition above induces a forcing relation for all $\Sigma_1^0(G)$ - and $\Pi_1^0(G)$ -formulas in that every set X compatible with a condition (σ, U) forcing some property $\varphi(G)$ will be such that $\varphi(X)$ holds.

For our construction, we want to satisfy the following requirements :

- $\mathcal{R}_k$: $(G \oplus Y)' \upharpoonright k$ is decided.
- $\mathcal{S}_k$: Let $\beta < \alpha$ be the largest ordinal of index less than k , for every $e < k$, $\Phi_e^{G \oplus Y}$ is not an M -infinite decreasing sequence of ordinals smaller than β .
- $\mathcal{T}_k$: $|\sigma| \geq k$.

We first prove a simple technical lemma stating the existence of arbitrarily large extendible nodes. This lemma even holds over RCA_0^* and can be found in various places in the literature.

Lemma 10.2.15 (Fernandes et al [44]). *Let $\mathcal{M} = (M, S) \models \text{RCA}_0^*$ and $T \subseteq 2^{<M}$ be an M -infinite tree in S . Then for every $\ell \in M$, there is some $\sigma \in T$ of length ℓ such that $\{\tau \in T : \tau \text{ is comparable with } \sigma\}$ is M -infinite.*

Proof. Assume by contradiction that for every $\sigma \in 2^\ell$ the tree

$$\{\tau \in T : \tau \text{ is comparable with } \sigma\}$$

is M -bounded, this means that $\mathcal{M} \models \forall \sigma \in 2^\ell \exists b_\sigma \forall \tau \in 2^{b_\sigma}, \sigma < \tau \Rightarrow \tau \notin T$.

The formula $\forall \tau \in 2^{b_\sigma}, \sigma < \tau \Rightarrow \tau \notin T$ is $\Sigma_0(T)$. Therefore, we can use BS_1^0 (which holds in RCA_0^*) to get :

$$\mathcal{M} \models \exists b \forall \sigma \in 2^\ell \exists b_\sigma < b \forall \tau \in 2^{b_\sigma}, \sigma < \tau \Rightarrow \tau \notin T$$

This yields that T is bounded by such a b , contradicting our assumption that T is M -infinite. So there is some $\sigma \in T$ of length ℓ such that $\{\tau \in T : \tau \text{ is comparable with } \sigma\}$ is M -infinite. $\square$

We will need the following two technical lemmas, which are Shore blocking arguments for satisfying the first two kinds of requirements, namely, deciding the jump and preserving $\text{WF}^0(\alpha)$. In what follows, $(\tau, U) \Vdash (G \oplus Y)' \upharpoonright k = \sigma'$ means that for every $e < |\sigma'|$, if $\sigma'(e) = 1$ then $(\tau, U) \Vdash \Phi_e^{G \oplus Y}(e) \downarrow$ and if $\sigma'(e) = 0$, then $(\tau, U) \Vdash \Phi_e^{G \oplus Y}(e) \uparrow$. Thus, this relation is $\Pi_1(Y)$.

Lemma 10.2.16. *Let (σ, T_1) be a condition. For every $k \in M$, there exists some M -finite σ' and an extension $(\tau, T_2) \leq (\sigma, T_1)$ forcing $(G \oplus Y)' \upharpoonright k = \sigma'$.*

Proof. Let W be the set of all $\sigma' \in 2^k$ such that the tree $\{\rho \in T_1 : \sigma'(e) = 0 \rightarrow \Phi_e^{\rho \oplus Y}(e) \uparrow\}$ is M -infinite. In general, the predicate “ X is infinite” is $\Pi_2(X)$ ($\forall x \exists y > x, y \in X$), but in the case of trees U , their closure under prefixes makes this predicate $\Pi_1(X)$ ($\forall \ell \exists \rho \in 2^\ell, \rho \in U$). Thus, W is $\Pi_1(Y)$ and hence M -finite as $\mathcal{M} \models \text{BC}\Pi_1^0$ by a relativization of Proposition 3.5.2, and it is non-empty as it contains the string $111 \dots 1$.

We can consider $\sigma' \in W$ its lexicographically minimal element, and the corresponding tree T_3 . For every $e < k$ such that $\sigma'(e) = 1$, the minimality of σ' gives us that the set of $\tau \in T_3$ such that $\Phi_e^{\tau \oplus Y}(e) \uparrow$ is M -finite, so we have a level h_e such that for every $\tau \in T_3 \cap 2^{h_e}$, $\Phi_e^{\tau \oplus Y}(e) \downarrow$. The set $\{e < k : \sigma'(e) = 1\}$ is M -finite, so we can consider an upper bound h of all the h_e . Finally, by Lemma 10.2.15, there is a $\tau \in T_3 \cap 2^h$ such that $T_2 := \{\rho \in T_3 : \rho \text{ is comparable with } \tau\}$ is M -infinite (since T_3 is M -infinite).

We claim that (τ, T_2) forces $(G \oplus Y)' \upharpoonright k = \sigma'$. Fix some $e < k$.

Case (a): $\sigma'(e) = 0$. Then $\Phi_e^{\tau \oplus Y}(e) \uparrow$ for every $\tau \in T_2$ (since $T_2 \subseteq T_3$), hence for every G compatible with (τ, T_2) , so $e \notin (G \oplus Y)'$.

Case (b): $\sigma'(e) = 1$. The definition of τ ensures that $\Phi_e^{\tau \oplus Y}(e) \downarrow$, so $e \in (G \oplus Y)'$. $\square$

Lemma 10.2.17. *Let (σ, T_1) be a condition. For every $k \in M$, letting Γ_k be the functional of Lemma 10.2.4, there is an extension $(\sigma, T_2) \leq (\sigma, T_1)$ forcing $\Gamma_k^{G \oplus Y}$ to be partial.*

Proof. Let $\beta < \alpha$ be the largest ordinal smaller than α of index at most k . By twisting Γ_k , we can assume that for every $e, n \in M$, and every binary sequence $\rho \in 2^{<M}$, that if $\Gamma_k^{\rho \oplus Y}(n) \downarrow$, then

- $\Gamma_k^{\rho \oplus Y}(m) \downarrow$ for every $m < n$;
- $\Gamma_k^{\rho \oplus Y}(0), \Gamma_k^{\rho \oplus Y}(1), \dots, \Gamma_k^{\rho \oplus Y}(n)$ is a strictly decreasing sequence of ordinals smaller than $\beta \otimes k$.

We have two cases.

Case 1: there exists some $n \in M$ such that for every $\ell \in M$, there is some $\tau \in T_1$ of length ℓ such that $\Gamma_k^{\tau \oplus Y}(n) \uparrow$. Let $T_2 = \{\tau \in T_1 : \Gamma_k^{\tau \oplus Y}(n) \uparrow\}$. The set T_2 is closed-downward, and by assumption, is M -infinite. Then $(\sigma, T_2) \leq (\sigma, T_1)$. Moreover, by construction, $(\sigma, T_2) \Vdash \Gamma_k^{G \oplus Y}(n) \uparrow$.

Case 2: for every $n \in M$, there is some $\ell_n \in M$ such that for every $\tau \in T_1$ of length ℓ_n , $\Gamma_k^{\tau \oplus Y}(n) \downarrow$.

Then, for every n , let $\alpha_n = \max \{\Gamma_k^{\tau \oplus Y}(n) : \tau \in T_1 \wedge |\tau| = \ell_n\}$. We claim that for every $n \in M$, $\alpha_{n+1} < \alpha_n$. Indeed, for every $\tau \in T_1$ such that $|\tau| = \ell_{n+1}$, $\Gamma_k^{\tau \oplus Y}(n+1) < \Gamma_k^{(\tau \upharpoonright \ell_n) \oplus Y}(n)$, so

$$\max \{\Gamma_k^{\tau \oplus Y}(n+1) : \tau \in T_1 \wedge |\tau| = \ell_{n+1}\} < \max \{\Gamma_k^{\tau \oplus Y}(n) : \tau \in T_1 \wedge |\tau| = \ell_n\}$$

So $\mathcal{M} \not\models \mathbf{WF}^0(\alpha \otimes k)$. However, by Lemma 10.2.3, since $\mathcal{M} \models \mathbf{RCA}_0 + \mathbf{B}\Sigma_2^0 + \mathbf{WF}^0(\alpha)$, $\mathcal{M} \models \mathbf{WF}^0(\alpha \otimes k)$. Contradiction. $\square$

Note that the statement “ (σ, T_1) forces $\Gamma_k^{G \oplus Y}(n) \uparrow$ ” is $\Pi_1(Y)$.

Construction We Y' -computably construct a decreasing sequence (σ_s, T_s) of conditions. In parallel, the set $(G \oplus Y)'$ is approximated by a sequence (σ'_s) of finite binary sequences.

Let $\sigma_0 = \sigma'_0 = \varepsilon$ and $T_0 = T$. Assume that σ_s, T_s , and σ'_s have been defined for some $s \in M$. Let e_s be the Turing index associated with T_s and let

$$k_s = 1 + \max\{|\sigma_s|, e_s, |\sigma'_s|\}.$$

Using Lemma 10.2.16 and Lemma 10.2.17, there exists some extension $(\tau, U) \leq (\sigma_s, T_s)$ forcing $\mathcal{R}_{k_s}$ and $\mathcal{S}_{k_s}$. Using Lemma 10.2.15, we can assume without loss of generality that $|\tau| \geq k_s$ and force $\mathcal{T}_{k_s}$. We then let $\sigma_{s+1} = \tau$, $T_{s+1} = U$ and $\sigma'_{s+1} = \sigma'$ where σ' is the sequence witnessing that $\mathcal{R}_{k_s}$ is forced.

The σ_s and σ'_s may only be defined on some cut I but are nonetheless of unbounded length in M .

Let $G = \bigcup_{s \in I} \sigma_s$. By construction G is Y' -computable and, as the $\mathcal{R}_k$ are satisfied, $(G \oplus Y)'$ is equal to $\bigcup_{s \in I} \sigma'_s$ and is also Y' -computable, hence $\mathcal{M}[G \oplus Y] = \mathcal{M}[G] =$

$\text{RCA}_0 + \text{B}\Sigma_2^0$. Since the requirements $\mathcal{T}_k$ are satisfied, the set G is an infinite path in T . Then, as the requirements $\mathcal{S}_k$ are satisfied, $\mathcal{M}[G] \models \text{WF}^0(\alpha)$.

This completes the proof of Proposition 10.2.13. $\square$

Theorem 10.2.18. *Fix an ordinal $\alpha \leq \epsilon_0$. $\text{RCA}_0 + \text{WKL} + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$ is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$.*

Proof. Immediate by Proposition 10.2.13 and Theorem 9.2.3. $\square$

10.3 Conservation over $\text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$

The cases of SEM and SADS are more complicated than those of WKL and COH, as $\text{WF}^0(\omega^{\alpha \times \omega})$ seems to be necessary to preserve $\text{WF}^0(\alpha)$ where $\alpha \leq \epsilon_0$. The ordinal ϵ_0 is not a fixed point of $\alpha \mapsto \omega^{\alpha \times \omega}$. Fortunately, to preserve $\text{WF}^0(\alpha)$, it is sufficient to preserve $\text{WF}^0(\beta)$ for every $\beta < \alpha$. This is why, in Lemma 10.2.4, we considered the largest ordinal $\beta < \alpha$ of index at most k , rather than α itself. Thus, we obtain that $\text{WF}^0(\sup_{\beta < \alpha} \omega^{\beta \times \omega})$ suffices to preserve $\text{WF}^0(\sup_{\beta < \alpha} \beta)$, and therefore that $\text{WF}^0(\epsilon_0)$ suffices to preserve itself.

10.3.1 Π_1^1 -conservation of SEM

As mentioned in the introduction, the bounded monotone enumeration principle naturally shows up when formalizing the first-jump control of Ramsey's theorem for pairs. Within the decomposition of RT_2^2 into $\text{COH} + \text{SADS} + \text{SEM}$ given in Proposition 6.2.11, SEM is the only statement not known to be Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$. It is therefore natural to expect to involve $\text{BME}_\star$ in the proof below. Since $\text{BME}_\star$ is equivalent to $\text{WF}^0(\omega^\omega)$ and we suppose $\mathcal{M} \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\sup_{\beta < \alpha} \omega^{\beta \times \omega})$, we will be able to use $\text{BME}_\star$.

Proposition 10.3.1. *Consider a countable model $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\sup_{\beta < \alpha} \omega^{\beta \times \omega})$ topped by a set $Y \in S$, where $\alpha \in M$ is an ordinal $\leq \epsilon_0$. Then for every infinite stable tournament R in S and every set $P \gg Y'$ such that $\mathcal{M}[P] \models \text{RCA}_0^*$, there exists $G \subseteq M$ such that*

1. G is an M -infinite transitive subtournament;
2. $P \geq_T (G \oplus Y)'$;
3. $\mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$.

Proof. In Proposition 10.2.5, the reservoirs $\vec{R}_p$ were all computable from the set Y . In the present construction, however, extending a condition requires computing a path through a tree that is itself computable from the previous stage. As a result, the construction involves an infinite sequence $Y \ll X_0 \ll X_1 \ll \dots$ of sets, each of PA degree over its predecessor.

By Proposition 10.2.13, we can add paths through trees to our model $\mathcal{M}$. Since we wish to remain within a topped model, we cannot directly apply Theorem 10.2.18 to obtain a model of WKL_0 containing such a sequence (and we also want the sequence to be uniformly computable in P , which Theorem 10.2.18 doesn't provide). However, it can be shown that there exists a single Y -computable tree whose paths encode such sequences, hence, a single application of Proposition 10.2.13 suffices:

Indeed, consider the class $\mathcal{C}$ equal to

$$\left\{ \bigoplus_a X_a : \forall a \forall e, \Phi_e^{Y \oplus X_0 \oplus \dots \oplus X_a} \text{ is an infinite binary tree} \rightarrow X_{\langle a, e \rangle} \in [\Phi_e^{Y \oplus X_0 \oplus \dots \oplus X_a}] \right\}$$

$\mathcal{C}$ is a non-empty Π_1^0 class relative to Y . By a relativization of Proposition 4.2.2, there exists a Y -computable tree whose paths are exactly the elements of $\mathcal{C}$. Thus,

Proposition 10.2.13 gives us a set $\vec{X} = \bigoplus_a X_a \in \mathcal{C}$ such that $Y' \geq_T (\vec{X} \oplus Y)'$ and $\mathcal{M}[\vec{X}] \models \mathbf{RCA}_0 + \mathbf{B}\Sigma_2^0 + \mathbf{WF}^0(\sup_{\beta < \alpha} \omega^{\beta \times \omega})$.

Furthermore, Proposition 10.2.5 applied to $\mathcal{M}[\vec{X}]$ gives us a set $C \subseteq M$ satisfying:

- C is $\vec{X}$ -cohesive;
- $P \geq_T (C \oplus \vec{X} \oplus Y)'$;
- $\mathcal{M}[\vec{X} \oplus C] \models \mathbf{RCA}_0 + \mathbf{B}\Sigma_2^0 + \mathbf{WF}^0(\sup_{\beta < \alpha} \omega^{\beta \times \omega})$

Let $A_0(R) := \{x : \forall^\infty y R(x, y)\}$ and $A_1(R) := \{x : \forall^\infty y R(y, x)\}$, where $\forall^\infty$ means “for all but finitely many”, so $A_0(R)$ and $A_1(R)$ are $\Sigma_2(Y)$. The stability of R gives us $A_0(R) \sqcup A_1(R) = M$, so $A_0(R)$ and $A_1(R)$ are $\Delta_2(Y)$.

Definition 10.3.2. For two sets F and G , the notation $F \rightarrow_R G$ denotes the formula $(\forall x \in F)(\forall y \in G)R(x, y)$.

Definition 10.3.3 (Minimal interval). Let $a, b \in M$, the interval (a, b) is the set of all x such that $R(a, x)$ and $R(x, b)$.

Let F be an M -finite transitive subtournament of R , for $a, b \in F \cup \{-\infty, +\infty\}$ we say that (a, b) is a minimal interval of F if $F \cap (a, b) = \emptyset$, where we consider that $R(-\infty, x)$ and $R(x, +\infty)$ hold for every $x \in M$.

Definition 10.3.4. A condition is a pair (σ, a) where σ is an M -finite binary string corresponding to an R -transitive subtournament, $a \in M$ is such that $C \subseteq^* X_a$ and X_a is included in a minimal interval of σ .

A set $X \subseteq M$ is *compatible* with a condition (σ, a) if $\sigma < X \subseteq \sigma \cup X_a$.

Being a condition is a $\Delta_1((C \oplus \vec{X} \oplus Y)')$ -formula. Indeed, being a finite R -transitive subtournament is Y -computable, checking that X_a is included in a minimal interval of σ is $(X_a \oplus Y)'$ -computable, and checking if $C \subseteq^* X_a$ can be decided using $(C \oplus X_a)'$ (since C is $\vec{X}$ -cohesive, we know that $C \subseteq^* X_a$ or $C \subseteq^* \mathbb{N} \setminus X_a$ and these events are $\Sigma_2(C \oplus X_a)$). Since $P \geq_T (C \oplus \vec{X} \oplus Y)'$, it is a $\Delta_1(P)$ predicate.

Definition 10.3.5. Let (σ, a) be a condition and $e, x \in M$, we write:

1. $(\sigma, a) \Vdash \Phi_e^{G \oplus Y}(x) \downarrow$ if $\Phi_e^{\sigma \oplus Y}(x) \downarrow$
2. $(\sigma, a) \Vdash \Phi_e^{G \oplus Y}(x) \uparrow$ if for every $\sigma \leq \tau \subseteq \sigma \cup X_a$, $\Phi_e^{\tau \oplus Y}(x) \uparrow$.

The relations $(\sigma, a) \Vdash \Phi_e^{G \oplus Y}(x) \downarrow$ and $(\sigma, a) \Vdash \Phi_e^{G \oplus Y}(x) \uparrow$ are $\Delta_1(Y)$ and $\Pi_1(Y \oplus \vec{X})$ respectively. Again, by definition of the forcing relation $\Vdash$, it is clear that if a condition (σ, a) forces some $\Sigma_1^0(G)$ or $\Pi_1^0(G)$ property, then this property holds for every set compatible with (σ, a) .

Definition 10.3.6. A condition (τ, b) extends a condition (σ, a) (and we write $(\tau, b) \leq (\sigma, a)$) if $\sigma \leq \tau$, $X_b \subseteq X_a$ and $\tau - \sigma \subseteq X_a$.

For our construction, we want to satisfy the following requirements :

- $\mathcal{R}_k$: $(G \oplus Y)' \upharpoonright k$ is decided.
- $\mathcal{S}_k$: Let $\beta < \alpha$ be the biggest ordinal of index less than k , for every $e < k$, $\Phi_e^{G \oplus Y}$ is not an M -infinite decreasing sequence of ordinals smaller than β .
- $\mathcal{T}_k$: There exists some $x > k$ such that $x \in G$.

We will need two technical lemmas, which are Shore blocking arguments for satisfying the two kinds of requirements, namely, deciding the jump and preserving $\mathbf{WF}^0(\alpha)$. Before this, we introduce some basic combinatorial concepts which will be useful for both lemmas.

Definition 10.3.7. Fix a condition (σ, a) .

- (1) A (σ, a) -node is an M -finite sequence of the form

$$\nu = \langle (i_0, \rho_0), (i_1, \rho_1), \dots, (i_{q-1}, \rho_{q-1}) \rangle$$

such that for every $s < q$, $i_s < 2$, $\rho_s \subseteq X_a$ is a non-empty M -finite R -transitive set with $\min \rho_s > \max \rho_{s-1}$ (letting $\rho_{-1} = \sigma$), and for every $s < t < q$, if $i_s = 0$ then $\rho_s \rightarrow_R \rho_t$, and if $i_s = 1$ then $\rho_t \rightarrow_R \rho_s$. We then write

$$\pi(\nu) = \sigma \cup \rho_0 \cup \dots \cup \rho_{q-1}$$

- (2) This (σ, a) -node is *valid* if for every $s < q$, $\rho_s \subseteq A_{i_s}(R)$
- (3) A (σ, a) -node ν satisfies a $\Sigma_1^0(\mathcal{M})$ -formula $\varphi(G, x) \equiv \exists i \psi(G \upharpoonright i, x)$ if $\psi(\pi(\nu), |\nu| - 1)$ holds.

Lemma 10.3.8. Fix a (σ, a) -node ν . Then $\pi(\nu)$ is R -transitive.

Proof. Say $\nu = \langle (i_0, \rho_0), (i_1, \rho_1), \dots, (i_{q-1}, \rho_{q-1}) \rangle$. We prove by Σ_0^0 -induction that for every $\ell < t$, ρ_ℓ is an R -transitive tournament included in a minimal interval of $\sigma \cup \rho_0 \cup \dots \cup \rho_{\ell-1}$ and that $\sigma \cup \rho_0 \cup \dots \cup \rho_\ell$ is R -transitive.

This is true for $\ell = 0$ since (σ, a) is a condition. And assuming the property to be true at a rank $\ell < t - 1$, the construction of $\rho_{\ell+1}$ ensures that it is included in the minimal interval of (σ, a) and that $\rho_{\ell+1} \rightarrow_R \rho_j$ if $i_j = 1$ and $\rho_j \rightarrow_R \rho_{\ell+1}$ if $i_j = 0$, so no element of $\sigma \cup \rho_0 \cup \dots \cup \rho_\ell$ is between two elements of $\rho_{\ell+1}$. Since $\sigma \cup \rho_0 \cup \dots \cup \rho_\ell$ is R -transitive (by the induction hypothesis), this means that $\rho_{\ell+1}$ is included in one of its minimal intervals, and since it is itself R -transitive, $\sigma \cup \rho_0 \cup \dots \cup \rho_\ell \cup \rho_{\ell+1}$ is R -transitive. $\square$

Lemma 10.3.9. Fix a valid (σ, a) -node ν . Then $(\pi(\nu), b)$ is an extension of (σ, a) for some $b \in M$.

Proof. Say $\nu = \langle (i_0, \rho_0), (i_1, \rho_1), \dots, (i_{q-1}, \rho_{q-1}) \rangle$. By definition of validity and stability of R , for every $s < q$, if $i_s = 0$ then $\forall x \in \rho_s \forall^\infty y \in X_a, R(x, y)$, and if $i_s = 1$ then $\forall x \in \rho_s \forall^\infty y \in X_a, R(y, x)$. By $\mathbf{B}\Sigma_2^0$, there exists some $r \in M$ such that for every $s < q$, if $i_s = 0$ then $\rho_s \rightarrow_R X_a \setminus \{0, \dots, r\}$, and if $i_s = 1$ then $X_a \setminus \{0, \dots, r\} \rightarrow_R \rho_s$. Let $b \in M$ be such that $X_b = X_a \setminus \{0, \dots, r\}$ and let $\tau = \sigma \cup \rho_0 \cup \dots \cup \rho_{q-1}$. By Lemma 10.3.8, τ is R -transitive and by construction, $X_b \subseteq X_a$ is included in a minimal interval of τ . Thus (τ, b) is an extension of (σ, a) . $\square$

Definition 10.3.10. Fix a condition (σ, a) .

- (1) A (σ, a) -tree is a set of (σ, a) -nodes closed downward under the prefix relation.
- (2) A (σ, a) -tree E is *universal* if for every (σ, a) -node ν that is not a leaf, there is some $r \in M$ such that for every M -finite 2-partition $A_0 \sqcup A_1 = \{0, \dots, r-1\}$, there is a side $i < 2$ and an M -finite R -transitive $\rho \subseteq A_i \cap X_a$ such that $\nu \cdot \langle i, \rho \rangle \in E$.
- (3) A (σ, a) -tree E satisfies a $\Sigma_1^0(\mathcal{M})$ -formula if all its nodes satisfy it.
- (4) A universal (σ, a) -tree E satisfying a $\Sigma_1^0(\mathcal{M})$ -formula $\varphi(G, x)$ is *maximal* for that formula if we cannot add any descendants to a leaf of E while still obtaining a universal (σ, a) -tree satisfying $\varphi(G, x)$.

Lemma 10.3.11. For every $\Sigma_1^0(\mathcal{M})$ -formula $\varphi(G, x)$ and every condition (σ, a) , there is a monotone enumeration of a universal (σ, a) -tree E satisfying $\varphi(G, x)$ that is maximal for $\varphi(G, x)$.

Proof. At stage 0 enumerate the empty sequence $\langle \rangle$. At stage s , for every leaf $\nu \in E$, if for every 2-partition $A_0 \sqcup A_1 = \{0, \dots, s\}$, there is some $i < 2$ and some $\rho \subseteq A_i$ for which

- (i) $\varphi(\pi(\nu) \cup \rho, |\nu|)$ holds
- (ii) $\nu \cdot \langle (i, \rho) \rangle$ is a (σ, a) -node.

then enumerate $\nu \cdot \langle (i, \rho) \rangle$ for all such partitions, and go to the next stage. If nothing is enumerated, simply go to the next stage. $\square$

Such a universal (σ, a) -tree is not necessarily M -finite.

Lemma 10.3.12. *Every M -finite universal (σ, a) -tree E has a valid leaf.*

Proof. We prove it by Δ_2^0 -induction. The root $\langle \rangle$ is vacuously valid. Assume $\nu \in E$ is a valid non-leaf. By universality of E , there is some $r \in M$ such that for every M -finite 2-partition $A_0 \sqcup A_1 = \{0, \dots, r-1\}$, there is a side $i < 2$ and an M -finite R -transitive $\rho \subseteq A_i \cap X_a$ such that $\nu \cdot \langle (i, \rho) \rangle \in E$. By $\mathbf{BC}\Delta_2^0(Y)$ (which is equivalent to $\mathbf{B}\Sigma_2^0(Y)$ modulo $\mathbf{RCA}_0^*$), the following partition is M -finite:

$$A_0(R) \upharpoonright r \sqcup A_1(R) \upharpoonright r = \{0, \dots, r-1\}$$

Thus, there exists some $i < 2$ and some $\rho \subseteq A_i(R)$ such that $\nu \cdot \langle (i, \rho) \rangle \in E$. By choice of i and ρ , the node $\nu \cdot \langle (i, \rho) \rangle$ is valid. $\square$

Lemma 10.3.13. *Let (σ, a) be a condition, $\varphi(G, x)$ be a $\Sigma_1^0(\mathcal{M})$ -formula, and E be an M -finite universal (σ, a) -tree satisfying $\varphi(G, x)$, which is maximal for $\varphi(G, x)$. Then there is an extension $(\tau, c) \leq (\sigma, a)$ and some $q \in M$ such that $(\tau, c) \Vdash \varphi(G, q-1)$ and $(\tau, c) \Vdash \neg\varphi(G, q)$.*

Proof. By Lemma 10.3.12, there is a valid leaf $\nu \in E$. Since E satisfies $\varphi(G, x)$, $\varphi(\pi(\nu), |\nu| - 1)$ holds. By Lemma 10.3.9, there is some $b \in M$ such that $(\pi(\nu), b) \leq (\sigma, a)$.

By maximality of the tree, for every $s \in M$, there exists a 2-partition $A_0 \sqcup A_1 = \{0, \dots, s\}$, such that for every side $i < 2$ and every M -finite $\rho \subseteq A_i \cap X_a$ such that $\nu \cdot \langle (i, \rho) \rangle$ is a (σ, a) -node, $\varphi(\pi(\nu) \cup \rho, |\nu|)$ does not hold.

Note that by choice of X_b , for every non-empty M -finite R -transitive $\rho \subseteq A_i \cap X_b$, $\nu \cdot \langle (i, \rho) \rangle$ is a (σ, a) -node. Thus, for every $s \in M$, there exists a 2-partition $A_0 \sqcup A_1 = \{0, \dots, s\}$, such that the following property holds:

(P): For every side $i < 2$ and every non-empty M -finite R -transitive $\rho \subseteq A_i \cap X_b$, $\varphi(\pi(\nu) \cup \rho, |\nu|)$ does not hold.

Consider the binary tree of all such partitions satisfying (P), where a partition $B_0 \sqcup B_1 = \{0, \dots, s+1\}$ is a direct descendant of $A_0 \sqcup A_1 = \{0, \dots, s\}$ when $B_i \cap \{0, \dots, s\} = A_i$ for every $i < 2$. This tree is therefore infinite, and there exists a 2-partition $A_0 \sqcup A_1 = X_b$ such that (P) holds. By definition of $\vec{X}$, there exists some $c, d \in M$ such that X_c, X_d is a 2-partition of X_b in $\mathcal{C}$. Moreover, since $C \subseteq^* X_b$, either $C \subseteq^* X_c$ or $C \subseteq^* X_d$. Since they play a symmetric role, say $C \subseteq^* X_c$. Then (τ, c) is an extension of (τ, b) forcing $\varphi(G, |\nu| - 1)$ and $\neg\varphi(G, |\nu|)$. $\square$

We are now ready to prove our two Shore blocking lemmas.

Lemma 10.3.14. *Let (σ, a) be a condition. For every $k \in M$, there exists some M -finite $\sigma' \in 2^k$ and an extension $(\tau, b) \leq (\sigma, a)$ such that $(\tau, b) \Vdash (G \oplus Y)' \upharpoonright k = \sigma'$*

Proof. Let $\varphi(G, x)$ be the $\Sigma_1^0(Y)$ -formula which holds if there exists some $e_0 < \dots < e_{x-1} < k$ such that for every $s < x$, $\Phi_{e_s}^{G \oplus Y}(e_s) \downarrow$.

By Lemma 10.3.11, there exists a monotone enumeration of a universal (σ, a) -tree E satisfying $\varphi(G, x)$ such that no descendants can be added to any leaf while preserving the property of being a universal (σ, a) -tree satisfying $\varphi(G, x)$. Moreover, since $\varphi(G, k+1)$ never holds, the enumeration E is k -bounded. Since $\alpha > 0$, $\mathcal{M} \models \mathbf{WF}^0(\omega^\omega)$, hence $\mathcal{M} \models \mathbf{BME}_*$, so E is M -finite. By Lemma 10.3.13, there is an extension $(\tau, b) \leq (\sigma, a)$ and some $q \in M$ such that $(\tau, b) \Vdash \varphi(G, q-1)$ and $(\tau, b) \Vdash \neg\varphi(G, q)$. Let $\sigma' = \{e < k : \Phi_e^{\tau \oplus Y}(e) \downarrow\}$. Then $(\tau, b) \Vdash (G \oplus Y)' \upharpoonright k = \sigma'$. $\square$

Lemma 10.3.15. *Let (σ, a) be a condition. For every $k \in M$, letting Γ_k be the functional of Lemma 10.2.4 for the order α , there is an extension $(\tau, b) \leq (\sigma, a)$ forcing $\Gamma_k^{G \oplus Y}$ to be partial.*

Proof. Let $\beta < \alpha$ be the largest ordinal smaller than α of index at most k . By twisting Γ_k , we can assume that for every $e, n \in M$, and every binary sequence $\rho \in 2^{<M}$ that if $\Gamma_k^{\rho \oplus Y}(n) \downarrow$, then

- $\Gamma_k^{\rho \oplus Y}(m) \downarrow$ for every $m < n$;
- $\Gamma_k^{\rho \oplus Y}(0), \Gamma_k^{\rho \oplus Y}(1), \dots, \Gamma_k^{\rho \oplus Y}(n)$ is a strictly decreasing sequence of ordinals smaller than $\beta \otimes k$.

Let $\varphi(G, x)$ be the $\Sigma_1^0(Y)$ -formula which holds if $\Gamma_k^{G \oplus Y}(x) \downarrow$. By Lemma 10.3.11, there exists a monotone enumeration of a universal (σ, a) -tree E satisfying $\varphi(G, x)$ that is maximal for $\varphi(G, x)$. We have two cases.

Case 1: E is M -finite. Then by Lemma 10.3.13, there is an extension $(\tau, c) \leq (\sigma, a)$ and some $q \in M$ such that $(\tau, c) \Vdash \varphi(G, q-1)$ and $(\tau, c) \Vdash \neg\varphi(G, q)$. In particular, $(\tau, c) \Vdash \Gamma_k^{G \oplus Y}(q) \uparrow$.

Case 2: E is not M -finite. Then there exist M -infinitely many stages $s \in M$ such that $E[s] \neq E[s+1]$. By speeding up the enumeration, we can assume without loss of generality that $E[s+1] \neq E[s] \neq \{\diamond\}$ for every $s \in M$.

For every $s \in M$ and $\nu \in E[s]$, let

$$\zeta_s(\nu) = \begin{cases} \omega^{\Gamma_k^{\pi(\nu) \oplus Y}(|\nu|-1)} & \text{if } \nu \text{ is a leaf of } E[s] \\ \bigoplus_{\mu \sqsubset E[s] \nu} \zeta_s(\mu) & \text{otherwise} \end{cases}$$

Finally, let $\zeta_s = \zeta_s(\diamond)$.

Claim: $\zeta_{s+1} < \zeta_s$ for all $s \in M$. Since $E[s+1] \neq E[s]$, there is some leaf $\nu \in E[s]$ which is not a leaf in $E[s+1]$. By assumption on Γ_k , for every $\mu \sqsubset_{E[s+1]} \nu$,

$$\Gamma_k^{\pi(\mu) \oplus Y}(|\mu|-1) < \Gamma_k^{\pi(\nu) \oplus Y}(|\nu|-1)$$

Hence

$$\bigoplus_{\mu \sqsubset_{E[s+1]} \nu} \omega^{\Gamma_k^{\pi(\mu) \oplus Y}(|\mu|-1)} < \omega^{\Gamma_k^{\pi(\nu) \oplus Y}(|\nu|-1)}$$

So $\zeta_{s+1}(\nu) < \zeta_s(\nu)$. By Σ_0^0 -induction along the ancestors of ν up to the root, we obtain $\zeta_{s+1}(\diamond) < \zeta_s(\diamond)$. This concludes the proof of our claim.

As $\zeta_s \leq \zeta_0 \leq \omega^{\beta \otimes k} \leq \omega^{\beta \times \omega}$ for all $s \in M$. This contradicts the fact that $\mathcal{M} \models \mathbf{WF}^0(\omega^{\beta \times \omega})$. $\square$

Construction We P -computably construct a decreasing sequence (σ_s, a_s) of conditions. In parallel, the set $(G \oplus Y)'$ is approximated by a sequence (σ'_s) of finite binary sequences.

Let $\sigma_0 = \sigma'_0 = \varepsilon$ and a_0 be an index corresponding to M . Assume that σ_s, a_s , and σ'_s have been defined for some $s \in M$ and let

$$k_s = 1 + \max\{|\sigma_s|, a_s, |\sigma'_s|\}.$$

Using Lemma 10.3.14 and Lemma 10.3.15, there exists some extension $(\tau, b) \leq (\sigma_s, a_s)$ forcing $\mathcal{R}_{k_s}$ and $\mathcal{S}_{k_s}$. Such an extension can be P -computably found. We can then add an element of X_b bigger than k_s to τ and then extract an infinite minimal interval from X_b using P to obtain some condition (ρ, c) forcing $\mathcal{T}_{k_s}$. We then let $\sigma_{s+1} = \rho$, $a_{s+1} = c$ and $\sigma'_{s+1} = \sigma'$ where σ' is the sequence witnessing that $\mathcal{R}_{k_s}$ is forced.

The σ_s and σ'_s may only be defined on some cut I but are nonetheless of unbounded length in M .

Let $G = \bigcup_{s \in I} \sigma_s$. By construction G is P -computable and, as the $\mathcal{R}_k$ are satisfied, $(G \oplus Y)'$ is equal to $\bigcup_{s \in I} \sigma'_s$ and is also P -computable, hence $\mathcal{M}[G \oplus Y] = \mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0$. Since the requirements $\mathcal{T}_k$ are satisfied, the set G is an M -infinite transitive subtournament of R . Then, as the requirements $\mathcal{S}_k$ are satisfied, $\mathcal{M}[G] \models \text{WF}^0(\alpha)$.

This completes the proof of Proposition 10.3.1. $\square$

Corollary 10.3.16. Consider a countable model $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\sup_{\beta < \alpha} \omega^{\beta \times \omega})$ topped by a set $Y \in S$, where $\alpha \in M$ is an ordinal $\leq \epsilon_0$. Then for every infinite stable tournament R in S , there exists $G \subseteq M$ such that

1. G is an M -infinite transitive subtournament;
2. $\mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$.

Proof. By Theorem 9.2.5, there exists a set $P \subseteq M$ such that $P \gg Y'$ and $\mathcal{M}[P] \models \text{RCA}_0^*$. The conclusion follows from Proposition 10.3.1. $\square$

Theorem 10.3.17. $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0) + \text{SEM}$ is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$.

Proof. Immediate by Corollary 10.3.16 for the case $\alpha = \epsilon_0$ and Theorem 9.2.3. $\square$

10.3.2 Π_1^1 -conservation of SADS

We now turn to the proof that SADS is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$. There exist two main constructions of solutions to instances of SADS: an asymmetric one, which assumes that there is no infinite computable descending sequence, and constructs an ascending one, as in Hirschfeldt and Shore [63, Theorem 2.11], and a symmetric one using split pairs, as in Lerman, Solomon, and Towsner [94] or Patey [114, Theorem 26]. Our proof follows the second construction.

Proposition 10.3.18. Consider a countable model $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\sup_{\beta < \alpha} \omega^{\beta \times \omega})$ topped by a set $Y \in S$, where $\alpha \in M$ is an ordinal $\leq \epsilon_0$. Then for every linear order $\mathcal{L} = (M, \leq_{\mathcal{L}}) \in S$ of order type $\omega + \omega^*$, there exists $G \subseteq M$ such that

1. G is an M -unbounded ascending or descending sequence for $\leq_{\mathcal{L}}$;
2. $Y' \geq_T (G \oplus Y)'$;
3. $\mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\alpha)$.

Proof. Let $\mathcal{L} = (M, \leq_{\mathcal{L}})$ be such an order. If $\mathcal{M}$ already contains an M -unbounded ascending or descending sequence for $\leq_{\mathcal{L}}$, then we are done, so assume otherwise. Let $U^0 \subseteq M$ and $U^1 \subseteq M$ be the sets of elements that have a finite amount of predecessors

and successors for $\leq_{\mathcal{L}}$, respectively. By assumption, $U^0 \sqcup U^1 = M$, and U^0 and U^1 are both $\Delta_2(Y)$.

Contrary to the cases of COH, WKL, and SEM, we will be building two sets in parallel instead of one: an ascending sequence G^0 and a descending sequence G^1 , and prove by a pairing argument that there exists some side $i < 2$ where all the requirements are forced.

A condition is a pair (σ^0, σ^1) with σ^0 an M -finite $\mathcal{L}$ -ascending sequence included in U^0 and σ^1 an M -finite $\mathcal{L}$ -descending sequence included in U^1 . We say that (τ^0, τ^1) extends (σ^0, σ^1) (and we write $(\tau^0, \tau^1) \leq (\sigma^0, \sigma^1)$) if $\tau^0 \geq \sigma^0$ and $\tau^1 \geq \sigma^1$. A set G^i is compatible with a condition (σ^0, σ^1) on side $i < 2$ when $\sigma^i < G^i$.

Being a condition is a $\Delta_2(Y)$ -predicate, while the extension relation is Δ_1 . We define a weaker notion of pair which is $\Delta_1(Y)$:

Definition 10.3.19. A *split pair* is a pair (σ^0, σ^1) of strings such that σ^0 is an M -finite $\mathcal{L}$ -ascending sequence, σ^1 is an M -finite $\mathcal{L}$ -descending sequence, and $\max_{\mathcal{L}} \sigma^0 <_{\mathcal{L}} \min_{\mathcal{L}} \sigma^1$.

The core property of split pairs (σ^0, σ^1) comes from the fact that if $\sigma^0 \not\subseteq U^0$, then $\sigma^1 \subseteq U^1$. In particular, if (σ^0, σ^1) is a condition and (τ^0, τ^1) is a split pair such that $\sigma^0 \leq \tau^0$ and $\sigma^1 \leq \tau^1$, then either (τ^0, σ^1) or (σ^0, τ^1) is a condition extending (σ^0, σ^1) .

Definition 10.3.20. Let (σ^0, σ^1) be a condition, $e, x \in M$ and $i < 2$, we write:

- $(\sigma^0, \sigma^1) \Vdash_i \Phi_e^{G^i \oplus Y}(x) \downarrow$ if $\Phi_e^{\sigma^i \oplus Y}(x) \downarrow$.
- $(\sigma^0, \sigma^1) \Vdash_i \Phi_e^{G^i \oplus Y}(x) \uparrow$ if $\forall \tau^i \geq \sigma^i$ such that the pair obtained by replacing the i -th component in (σ^0, σ^1) by τ^i is a split pair, $\Phi_e^{\tau^i \oplus Y}(x) \uparrow$.

Note that the relations $(\sigma^0, \sigma^1) \Vdash_i \Phi_e^{G^i \oplus Y}(x) \downarrow$ and $(\sigma^0, \sigma^1) \Vdash_i \Phi_e^{G^i \oplus Y}(x) \uparrow$ are $\Delta_1(Y)$ and $\Pi_1(Y)$, respectively.

As in the proof of Theorem 8.5.6, we introduce a forcing question which, thanks to Lemma 10.3.22, allows us to Y' -computably determine a side $i < 2$ and construct extensions forcing either a given property or its negation on side i .

Definition 10.3.21. Let (σ^0, σ^1) be a condition and $e_0, e_1, x_0, x_1 \in M$. We write $(\sigma^0, \sigma^1) ? \vdash \Phi_{e_0}^{G^0 \oplus Y}(x_0) \downarrow \vee \Phi_{e_1}^{G^1 \oplus Y}(x_1) \downarrow$ if there exists some split pair (τ^0, τ^1) such that $\sigma^i \leq \tau^i$ and $\Phi_{e_i}^{\tau^i \oplus Y}(x_i) \downarrow$ for every $i < 2$.

The relation $(\sigma^0, \sigma^1) ? \vdash \Phi_{e_0}^{G^0 \oplus Y}(x_0) \downarrow \vee \Phi_{e_1}^{G^1 \oplus Y}(x_1) \downarrow$ is $\Sigma_1(Y)$.

Lemma 10.3.22. Let (σ^0, σ^1) be a condition, and let $e_0, e_1, x_0, x_1 \in M$. Then:

- If $(\sigma^0, \sigma^1) ? \vdash \Phi_{e_0}^{G^0 \oplus Y}(x_0) \downarrow \vee \Phi_{e_1}^{G^1 \oplus Y}(x_1) \downarrow$, then there exist an extension $(\tau^0, \tau^1) \leq (\sigma^0, \sigma^1)$ and some $i < 2$ such that $(\tau^0, \tau^1) \Vdash_i \Phi_{e_i}^{G^i \oplus Y}(x_i) \downarrow$.
- If $(\sigma^0, \sigma^1) ? \not\vdash \Phi_{e_0}^{G^0 \oplus Y}(x_0) \downarrow \vee \Phi_{e_1}^{G^1 \oplus Y}(x_1) \downarrow$, then there exist an extension $(\tau^0, \tau^1) \leq (\sigma^0, \sigma^1)$ and some $i < 2$ such that $(\tau^0, \tau^1) \Vdash_i \Phi_{e_i}^{G^i \oplus Y}(x_i) \uparrow$.

Furthermore, deciding which case holds and computing the extension (τ^0, τ^1) can be done Y' -computably.

Proof. Assume first that $(\sigma^0, \sigma^1) ? \vdash \Phi_{e_0}^{G^0 \oplus Y}(x_0) \downarrow \vee \Phi_{e_1}^{G^1 \oplus Y}(x_1) \downarrow$. Then there exists a split pair (ρ^0, ρ^1) such that, for every $i < 2$, $\sigma^i \leq \rho^i$ and $\Phi_{e_i}^{\rho^i \oplus Y}(x_i) \downarrow$. Using Y , we can find such a split pair. Since (ρ^0, ρ^1) is a split pair, there is some side $i < 2$ such that $\rho^i \subseteq U^i$, and such a side can be found using Y' . Replacing σ^i by ρ^i in the condition

(σ^0, σ^1) yields an extension $(\tau^0, \tau^1) \leq (\sigma^0, \sigma^1)$ such that $(\tau^0, \tau^1) \Vdash_i \Phi_{e_i}^{G^i \oplus Y}(x_i) \downarrow$.

Now suppose that $(\sigma^0, \sigma^1) \nVdash \Phi_{e_0}^{G^0 \oplus Y}(x_0) \downarrow \vee \Phi_{e_1}^{G^1 \oplus Y}(x_1) \downarrow$. Consider the program which runs through all M -finite $\mathcal{L}$ -ascending sequences ρ^0 such that $\sigma^0 \leq \rho^0$, $\max_{\mathcal{L}} \rho^0 <_{\mathcal{L}} \min_{\mathcal{L}} \sigma^1$ and $\Phi_{e_0}^{\rho^0 \oplus Y}(x_0) \downarrow$. Whenever a new ρ^0 is found whose value of $\max_{\mathcal{L}} \rho^0$ is smaller than all previously found ones, enumerate $\max_{\mathcal{L}} \rho^0$. By construction, this program yields a Y -computable $\mathcal{L}$ -descending sequence. By our assumption, $\mathcal{M}$ does not contain any solution to the given instance of SADS, this enumeration must therefore be finite. Hence, either no such ρ^0 is enumerated, in which case $(\sigma^0, \sigma^1) \Vdash_0 \Phi_{e_0}^{G^0 \oplus Y}(x_0) \uparrow$ and we are done, or else we can Y' -computably find some ρ^0 such that $\sigma^0 \leq \rho^0$, $\max_{\mathcal{L}} \rho^0 <_{\mathcal{L}} \min_{\mathcal{L}} \sigma^1$ and $\Phi_{e_0}^{\rho^0 \oplus Y}(x_0) \downarrow$ and for which $\max_{\mathcal{L}} \rho^0$ is minimal.

Similarly, either we already have $(\sigma^0, \sigma^1) \Vdash_1 \Phi_{e_1}^{G^1 \oplus Y}(x_1) \uparrow$ or we can Y' -computably find some M -finite $\mathcal{L}$ -descending sequence ρ^1 such that $\sigma^1 \leq \rho^1$, $\min_{\mathcal{L}} \rho^1 >_{\mathcal{L}} \max_{\mathcal{L}} \sigma^0$ and $\Phi_{e_1}^{\rho^1 \oplus Y}(x_1) \downarrow$ and such that $\min_{\mathcal{L}} \rho^1$ is maximal among all such sequences.

In the case where such ρ^0 and ρ^1 are found, we cannot have $\rho^i \subseteq U^i$ for both $i < 2$, since then (ρ^0, ρ^1) would form a split pair, contradicting our assumption that $(\sigma^0, \sigma^1) \nVdash \Phi_{e_0}^{G^0 \oplus Y}(x_0) \downarrow \vee \Phi_{e_1}^{G^1 \oplus Y}(x_1) \downarrow$, and we can Y' -computably find some side $i < 2$ such that $\rho^i \not\subseteq U^i$. Without loss of generality, assume $i = 0$. Define $\tau^0 = \sigma^0$ and $\tau^1 = \sigma^1 \cup \{\max_{\mathcal{L}} \rho^0\}$, then (τ^0, τ^1) is a condition extending (σ^0, σ^1) . Indeed, since $\rho^0 \not\subseteq U^0$, we have $\max_{\mathcal{L}} \rho^0 \in U^1$, and by construction $\max_{\mathcal{L}} \rho^0 <_{\mathcal{L}} \min_{\mathcal{L}} \sigma^1$. Finally, we have $(\tau^0, \tau^1) \Vdash_0 \Phi_{e_0}^{G^0 \oplus Y}(x_0) \uparrow$. Indeed, by minimality of $\max_{\mathcal{L}} \rho^0$, there is no M -finite $\mathcal{L}$ -ascending sequence $\rho^{0'}$ such that $\tau^0 = \sigma^0 \leq \rho^{0'}$, $\max_{\mathcal{L}} \rho^{0'} <_{\mathcal{L}} \min_{\mathcal{L}} \tau^1 = \max_{\mathcal{L}} \rho^0$ and $\Phi_{e_0}^{\rho^{0'} \oplus Y}(x_0) \downarrow$. $\square$

We want our construction to satisfy for some $i < 2$ the following requirements for every $k \in M$:

- $\mathcal{R}_k^i$: $(G^i \oplus Y)' \upharpoonright k$ is decided.
- $\mathcal{S}_k^i$: Let $\beta < \alpha$ be the biggest ordinal of index less than k , for every $e < k$, $\Phi_e^{G^i \oplus Y}$ is not an M -infinite decreasing sequence of ordinals smaller than β .
- $\mathcal{T}_k^i$: There exists some $x > k$ such that $x \in G^i$.

For this, we will use a pairing argument, and ensure that for every $k \in M$, the following requirements are met: $\mathcal{R}_k^0 \vee \mathcal{R}_k^1$, $\mathcal{R}_k^0 \vee \mathcal{S}_k^1$, $\mathcal{S}_k^0 \vee \mathcal{R}_k^1$, $\mathcal{S}_k^0 \vee \mathcal{S}_k^1$, and $\mathcal{T}_k^0 \wedge \mathcal{T}_k^1$, and therefore, that $(\mathcal{R}_k^0 \wedge \mathcal{S}_k^0 \wedge \mathcal{T}_k^0) \vee (\mathcal{R}_k^1 \wedge \mathcal{S}_k^1 \wedge \mathcal{T}_k^1)$ is met for every $k \in M$. Then, there will be one side $i < 2$ such that $\mathcal{R}_k^i \wedge \mathcal{S}_k^i \wedge \mathcal{T}_k^i$ is satisfied for M -unboundedly many values of k , hence that all requirements are satisfied for this side.

Definition 10.3.23. Fix a condition (σ^0, σ^1) .

A (σ^0, σ^1) -tree is a binary tree $T \subseteq 2^{<M}$ labeled by a family of split pairs $(\sigma_\rho^0, \sigma_\rho^1)_{\rho \in T}$ such that for all $\rho \in T$:

- $\sigma^0 \leq \sigma_\rho^0$ and $\sigma^1 \leq \sigma_\rho^1$;
- For every $\mu \in T$ and every $i < 2$, if $\rho \cdot i \leq \mu$ then $\sigma_\rho^i \leq \sigma_\mu^i$;

A node ρ is said to be *valid* if for every $i < 2$ and every $\mu \in T$ such that $\mu \cdot i \leq \rho$, $\sigma_\mu^i \subseteq U^i$. Being valid is a $\Delta_2(Y)$ condition.

Definition 10.3.24. Let T be a (σ^0, σ^1) -tree

- A node $(\sigma_\rho^0, \sigma_\rho^1)$ in T satisfies a pair of $\Sigma_1^0(Y)$ -formulas $(\varphi^0(G^0, x), \varphi^1(G^1, x))$ if $\varphi^0(\sigma_\rho^0, |\rho| - \text{card } \rho)$ and $\varphi^1(\sigma_\rho^1, \text{card } \rho)$ holds (where $\text{card } \rho$ is the cardinal of ρ seen as a set, to be distinguished from the length $|\rho|$ of ρ as a string).
- T satisfies $(\varphi^0(G^0, x), \varphi^1(G^1, x))$ if every node of T satisfies it.

Lemma 10.3.25. *For every pair $(\varphi^0(G^0, x), \varphi^1(G^1, x))$ of $\Sigma_1^0(Y)$ -formulas and every condition (σ^0, σ^1) , there is a maximal $\Sigma_1(Y)$ (σ^0, σ^1) -tree T satisfying it.*

Proof. T will be the limit of the following increasing sequence of Y -computable (σ^0, σ^1) -trees $(T_s)_{s \in M}$ that all satisfy $(\varphi^0(G^0, x), \varphi^1(G^1, x))$ (and therefore T will also be a (σ^0, σ^1) -tree satisfying $(\varphi^0(G^0, x), \varphi^1(G^1, x))$).

At stage 0, let T_0 be the empty tree. It vacuously satisfies $(\varphi^0(G^0, x), \varphi^1(G^1, x))$.

At stage $s+1$, if T_s is empty and there exists a split pair (τ^0, τ^1) with $|\tau^0|, |\tau^1| \leq s$ such that $\varphi^i(\tau^i, 0)$ holds for every $i < 2$, then let T_{s+1} consist solely of the root indexed by (τ^0, τ^1) ; it therefore satisfies $(\varphi^0(G^0, x), \varphi^1(G^1, x))$. If T_s is non-empty and there exists a split pair (τ^0, τ^1) with $|\tau^0|, |\tau^1| \leq s$, a $\rho \in T_s$ and an $i < 2$ with $\rho \cdot i \notin T_s$ such that adding $\rho \cdot i$ to T_s with the index (τ^0, τ^1) still yields a (σ^0, σ^1) -tree satisfying $(\varphi^0(G^0, x), \varphi^1(G^1, x))$, then let T_{s+1} be that tree. If, for several values of ρ, i , the node $\rho \cdot i$ can be added to T_s , choose the one for which $\rho \cdot i$ is minimal in the lexicographic ordering, so that every node in T that can potentially be extended eventually will be. If no such element is found, simply go to the next stage. $\square$

Lemma 10.3.26. *Let (σ^0, σ^1) be a condition, $(\varphi^0(G^0, x), \varphi^1(G^1, x))$ be a pair of $\Sigma_1^0(Y)$ -formulas, and T be an M -finite maximal (σ^0, σ^1) -tree satisfying $(\varphi^0(G^0, x), \varphi^1(G^1, x))$. Then there is an extension $(\hat{\sigma}^0, \hat{\sigma}^1) \leq (\sigma^0, \sigma^1)$ and some $p_0, p_1 \in M \cup \{-1\}$ such that the following holds:*

- $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_0 \varphi^0(G^0, p_0)$ and $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_1 \varphi^1(G^1, p_1)$
- $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_0 \neg \varphi^0(G^0, p_0 + 1)$ or $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_1 \neg \varphi^1(G^1, p_1 + 1)$.

With the convention that $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \varphi^i(G^i, -1)$ always holds.

Proof. If T is empty, then $(\sigma^0, \sigma^1) \not\models \varphi^0(G^0, 0) \vee \varphi^1(G^1, 0)$ as otherwise the corresponding split pair could be added in position ε in T , contradicting its maximality. Thus, by Lemma 10.3.22, there exists some extension $(\hat{\sigma}^0, \hat{\sigma}^1) \leq (\sigma^0, \sigma^1)$ and some side $i < 2$ such that $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \neg \varphi^i(G^i, 0)$. Thus, we can assume that T is not empty.

Consider a valid $\rho \in T$ having no valid descendant in T , this is possible since the set of valid elements of T is non empty because $(\sigma_\varepsilon^0, \sigma_\varepsilon^1)$ is a valid condition, so we can just take a maximal valid element of T using ID_2^0 (which we have by BS_2^0).

Since $(\sigma_\rho^0, \sigma_\rho^1)$ is a split pair, we have two cases:

Case 1: $\sigma_\rho^0 \subseteq U^0$. In that case $\rho \cdot 0 \notin T$ by maximality of ρ . Then take $p_0 = |\rho| - \text{card } \rho$, $p_1 = \text{card } \rho - 1$ and $(\tau^0, \tau^1) = (\sigma_\rho^0, \sigma_{\rho \upharpoonright (i_0+1)}^1)$ with i_0 the biggest index $i < |\rho|$ such that $\rho(i) = 1$ (if no such index exists we take σ^1 instead of $\sigma_{\rho \upharpoonright (i_0+1)}^1$). The pair (τ^0, τ^1) is a condition by definition of validity and the hypothesis $\sigma_\rho^0 \subseteq U^0$. $(\tau^0, \tau^1) \Vdash_0 \varphi^0(G^0, p_0)$ and $(\tau^0, \tau^1) \Vdash_1 \varphi^1(G^1, p_1)$ because T satisfies $(\varphi^0(G^0, x), \varphi^1(G^1, x))$.

We have $(\tau^0, \tau^1) \not\models \varphi^0(G^0, p_0 + 1) \vee \varphi^1(G^1, p_1 + 1)$ as otherwise the corresponding split pair could be added in position $\rho \cdot 0$ in T , contradicting its maximality. Hence, by Lemma 10.3.22, there exists some extension $(\hat{\sigma}^0, \hat{\sigma}^1)$ and some side $i < 2$ such that $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \neg \varphi^i(G^i, p_i + 1)$.

Case 2: The case $\sigma_\rho^1 \subseteq U^1$ can be treated similarly. $\square$

Lemma 10.3.27. *Let $k \in M$, let $\Phi_k(G, x)$ be the formula: $\exists e_0 < e_1 < \dots < e_x < k$ such that $\Phi_{e_i}^{G \oplus Y}(e_i) \downarrow$ for all $i < x$, let $\Psi_k(G, x)$ be the formula: $\Gamma_k^{G \oplus Y}(x) \downarrow$ (with Γ_k the functional of Lemma 10.2.4 for the order α).*

For any combination $\varphi^0(G, x), \varphi^1(G, x) \in \{\Phi_k(G, x), \Psi_k(G, x)\}$ and every condition (σ^0, σ^1) , there exists an extension $(\hat{\sigma}^0, \hat{\sigma}^1) \leq (\sigma^0, \sigma^1)$ and some $p_0, p_1 \in M \cup \{-1\}$ such that the following holds:

- $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_0 \varphi^0(G^0, p_0)$ and $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_1 \varphi^1(G^1, p_1)$
- $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_0 \neg \varphi^0(G^0, p_0 + 1)$ or $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_1 \neg \varphi^1(G^1, p_1 + 1)$.

With the convention that $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \varphi^i(G^i, -1)$ always holds.

Proof. By Lemma 10.3.25 there exists a maximal $\Sigma_1(\sigma^0, \sigma^1)$ -tree T satisfying

$$(\varphi^0(G, x), \varphi^1(G, x)).$$

To apply Lemma 10.3.26 and conclude, we just need to prove the M -finiteness of T .

Let $\beta < \alpha$ be the largest ordinal smaller than α of index at most k . By twisting Γ_k , we can assume that for every $e, n \in M$, and every binary sequence $\rho \in 2^{<M}$, that if $\Gamma_k^{\rho \oplus Y}(n) \downarrow$, then

- $\Gamma_k^{\rho \oplus Y}(m) \downarrow$ for every $m < n$;
- $\Gamma_k^{\rho \oplus Y}(0), \Gamma_k^{\rho \oplus Y}(1), \dots, \Gamma_k^{\rho \oplus Y}(n)$ is a strictly decreasing sequence of ordinals smaller than $\beta \otimes k$.

Assume by contradiction that T is M -infinite and let $(T_s)_{s \in M}$ be an increasing sequence of computable tree approaching T . Without loss of generality we can assume that T_{s+1} contains exactly one more element than T_s for all $s \in M$.

For every $s \in M$ and ρ node of T_s indexed by $(\sigma_\rho^0, \sigma_\rho^1)$ with set of immediate successors S_ρ ($\text{card } S_\rho \leq 2$ since T_s is a binary tree), define inductively :

$$\zeta_s(\rho) = \omega^{f_0(\rho) \oplus f_1(\rho)} (2 - \text{card } S_\rho) \oplus \bigoplus_{\mu \in S_\rho} \zeta_s(\mu)$$

where $f_0(\rho) = \Gamma_k^{\sigma_\rho^0 \oplus Y}(|\rho| - \text{card } \rho)$ if $\varphi^0(G, x) = \Psi_k(G, x)$ and $f_0(\rho) = k - (|\rho| - \text{card } \rho)$ if $\varphi^0(G, x) = \Phi_k(G, x)$. And $f_1(\rho) = \Gamma_k^{\sigma_\rho^1 \oplus Y}(\text{card } \rho)$ if $\varphi^1(G, x) = \Psi_k(G, x)$ and $f_1(\rho) = k - \text{card } \rho$ if $\varphi^1(G, x) = \Phi_k(G, x)$.

And let $\zeta_s = \zeta_s(\varepsilon)$.

Claim. *The sequence ζ_s is an M -infinitely decreasing sequence of ordinals smaller than $\omega^{\beta \times \omega}$.*

Let $s \in M$, since $T_{s+1} \neq T_s$, there is a $\rho \in T_s$ and an $i < 2$ such that $T_{s+1} = T_s \cup \rho \cdot i$. Therefore:

$$\zeta_s(\rho) \oplus (\omega^{f_0(\rho \cdot i) \oplus f_1(\rho \cdot i)} \otimes 2) = \zeta_{s+1}(\rho) \oplus \omega^{f_0(\rho) \oplus f_1(\rho)}$$

But for every possible values of φ^0 and φ^1 , we have $f_0(\rho) > f_0(\rho \cdot 0)$, $f_0(\rho) = f_0(\rho \cdot 1)$, $f_1(\rho) = f_1(\rho \cdot 0)$ and $f_1(\rho) > f_1(\rho \cdot 1)$. So for every possible value of i , we have

$$\omega^{f_0(\rho) \oplus f_1(\rho)} > \omega^{f_0(\rho \cdot i) \oplus f_1(\rho \cdot i)} \times 2$$

and therefore $\zeta_{s+1}(\rho) < \zeta_s(\rho)$ which gives us $\zeta_{s+1} < \zeta_s$ (as $\rho \cdot i$ is the only difference between T_s and T_{s+1}).

Therefore, for all $s \in M$, $\zeta_s \leq \zeta_0 \leq \omega^{\beta \otimes k} \leq \omega^{\beta \times \omega}$.

This contradicts $\mathbf{WF}^0(\omega^{\beta \times \omega})$, so T is M -finite and the result follows. $\square$

Remark 10.3.28. If $(\sigma^0, \sigma^1) \Vdash_i \Phi_k(G^i, p-1)$ and $(\sigma^0, \sigma^1) \Vdash_i \neg \Phi_k(G^i, p)$, then there exists a $\rho' \in 2^k$ such that $(\sigma^0, \sigma^1) \Vdash_i (G^i \oplus Y)' \upharpoonright k = \rho'$. Such a ρ' can be found computably in Y .

Moreover, if $(\sigma^0, \sigma^1) \Vdash_i \Psi_k(G^i, p-1)$ and $(\sigma^0, \sigma^1) \Vdash_i \neg \Psi_k(G^i, p)$ then $(\sigma^0, \sigma^1) \Vdash_i \Gamma_k^{G^i \oplus Y}(p) \uparrow$, so (σ^0, σ^1) forces $\Gamma_k^{G^i \oplus Y}$ to be partial.

Remark 10.3.29. Because of the disjunctive requirements, the preservation of $\mathbf{WF}^0(\epsilon_0)$ and the control of $(Y \oplus G)'$ must be combined. Thus, at first sight, $\mathbf{WF}^0(\epsilon_0)$ is used to control $(Y \oplus G)'$. However, if one wants only to control $(Y \oplus G)'$, no well-

foundedness assumption is necessary. Indeed, let $\mathcal{P}$ be the set of pairs $(p, q) \in k^2$ such that there is a split pair (τ^0, τ^1) with $\tau^0 \geq \sigma^0$ and $\tau^1 \geq \sigma^1$ satisfying $\Psi_k(\sigma^1, p)$ and $\Psi_k(\tau^1, q)$. Then $\mathcal{P}$ exists by bounded Σ_1^0 comprehension. Using this set $\mathcal{P}$, one can find an extension with the desired property. Therefore, one can prove that SADS is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$, as already proved by Chong, Slaman, and Yang [22] using the first construction.

Construction We Y' -computably construct a decreasing sequence (σ_s^0, σ_s^1) of conditions. In parallel, the sets $(G^0 \oplus Y)'$ and $(G^1 \oplus Y)'$ are approximated by sequences $(\sigma_s^{0'})$ and $(\sigma_s^{1'})$ of finite binary sequences.

Let $\sigma_0^0 = \sigma_0^1 = \sigma_0^{0'} = \sigma_0^{1'} = \varepsilon$. Assume that $\sigma_s^0, \sigma_s^1, \sigma_s^{0'}$, and $\sigma_s^{1'}$ have been defined for some $s \in M$ and let

$$k_s = 1 + \max\{|\sigma_s^0|, |\sigma_s^1|, |\sigma_s^{0'}|, |\sigma_s^{1'}|\}.$$

Using Lemma 10.3.27 and the fact that U^0 and U^1 are M -infinite, there exists some extension $(\sigma^0, \sigma^1) \leq (\sigma_s^0, \sigma_s^1)$ forcing $\mathcal{R}_{k_s}^0 \vee \mathcal{R}_{k_s}^1, \mathcal{R}_{k_s}^0 \vee \mathcal{S}_{k_s}^1, \mathcal{S}_{k_s}^0 \vee \mathcal{R}_{k_s}^1, \mathcal{S}_{k_s}^0 \vee \mathcal{S}_{k_s}^1$, and $\mathcal{T}_{k_s}^0 \wedge \mathcal{T}_{k_s}^1$. Such an extension can be found Y' -computably. Thus, there exists some side $i < 2$ such that (σ^0, σ^1) forces $\mathcal{R}_{k_s}^i \wedge \mathcal{S}_{k_s}^i \wedge \mathcal{T}_{k_s}^i$. We then let $\sigma_{s+1}^i = \sigma^i$, $\sigma_{s+1}^{1-i} = \sigma_s^{1-i}$, $\sigma_{s+1}^{i'} = \sigma^{i'}$, where $\sigma^{i'}$ is the sequence witnessing that $\mathcal{R}_{k_s}^i$ is forced and $\sigma_{s+1}^{1-i'} = \sigma_s^{1-i'}$.

Even if the $\sigma_s^0, \sigma_s^1, \sigma_s^{0'}$ and $\sigma_s^{1'}$ may only be defined on some cut I , the values taken by the k_s are unbounded in M , hence for every $k \in M$, there exists some side $i < 2$ such that $\mathcal{R}_k^i \wedge \mathcal{S}_k^i \wedge \mathcal{T}_k^i$ is satisfied. Therefore, there exists some $i < 2$ such that $\mathcal{R}_k^i \wedge \mathcal{S}_k^i \wedge \mathcal{T}_k^i$ is satisfied for every $k \in M$.

Let $G = \bigcup_{s \in I} \sigma_s^i$. By construction G is Y' -computable and, as the $\mathcal{R}_k^i$ are satisfied, $(G \oplus Y)'$ is equal to $\bigcup_{s \in I} \sigma_s^{i'}$ and is also Y' -computable, hence $\mathcal{M}[G \oplus Y] = \mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0$. Since the requirements $\mathcal{T}_k^i$ are satisfied, the set G is an M -infinite $\mathcal{L}$ -ascending sequence if $i = 0$ and an M -infinite $\mathcal{L}$ -descending sequence if $i = 1$. Then, as the requirements $\mathcal{S}_k^i$ are satisfied, $\mathcal{M}[G] \models \text{WF}^0(\alpha)$.

This completes the proof of Proposition 10.3.18. $\square$

Theorem 10.3.30. SADS is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$.

Proof. Immediate by Proposition 10.3.18 for the case $\alpha = \epsilon_0$ and Theorem 9.2.3. $\square$

10.4 Π_1^1 -conservation of RT_2^2

Thanks to the decomposition of Proposition 6.2.11, we can combine the conservation theorems developed in Section 10.2 and Section 10.3 to prove our main theorems. Yokoyama [142] proved the following amalgamation theorem:

Theorem 10.4.1 (Yokoyama [142]). *Let $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2$ be Π_2^1 theories such that $\mathcal{T}_0 \supseteq \text{RCA}_0$ and $\mathcal{T}_1$ and $\mathcal{T}_2$ are both Π_1^1 -conservative over $\mathcal{T}_0$. Then $\mathcal{T}_1 + \mathcal{T}_2$ is Π_1^1 -conservative over $\mathcal{T}_0$.*

In particular, letting $\mathcal{T}_0$ be $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$, one can combine Theorem 10.2.12, Theorem 10.2.18, Theorem 10.3.17 and Theorem 10.3.30 to deduce that $\text{RT}_2^2 + \text{WKL}$ is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$, thereby proving our first main theorem of the chapter.

Main Theorem 10.1.3. $\text{RT}_2^2 + \text{WKL}$ is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{WF}^0(\epsilon_0)$.

A more direct proof of that result can be obtained by noticing that, in the proof of Theorem 9.2.3, it is possible to alternate between several problems when adding solutions as is the case in the proof of Theorem 10.4.2 below.

10.4.1 Preserving $\bigcup_n \text{WF}^0(\omega_n)$

In their proof that RT_2^2 does not imply IS_2^0 over RCA_0 , Chong, Slaman, and Yang [23] started with a model satisfying a slightly weaker well-foundedness axiom than $\text{WF}^0(\epsilon_0)$, namely, $\bigcup_{n \in \omega} \text{WF}^0(\omega_n)$. We show that our assumption can also be weakened to $\bigcup_{n \in \omega} \text{WF}^0(\omega_n)$ with a little bit of extra work.

Theorem 10.4.2. *Let $\mathcal{T}$ be a Π_2^1 -axiomatizable theory, $n \in \omega$ and $P_0, \dots, P_{n-1}$ be Π_2^1 problems. Suppose that there exists some $\ell \in \omega$ such that, for every countable topped model $\mathcal{M} = (M, S) \models \text{RCA}_0^* + \mathcal{T} + \text{WF}^0(\omega_\ell^\alpha)$ with $\alpha < \epsilon_0$ and every instance of P_i in $\mathcal{M}$ for some $i < n$, there exists a solution $G \subseteq M$ of that instance such that $\mathcal{M}[G] \models \text{RCA}_0^* + \mathcal{T} + \text{WF}^0(\alpha)$. Then $P_0 + \dots + P_{n-1}$ is Π_1^1 -conservative over $\text{RCA}_0^* + \mathcal{T} + \bigcup_{k \in \omega} \text{WF}^0(\omega_k)$.*

Proof. Assume $\text{RCA}_0^* + \mathcal{T} + \bigcup_{k \in \omega} \text{WF}^0(\omega_k) \not\models \forall X \varphi(X)$ for φ an arithmetic formula. By completeness, there exists a model $\mathcal{N} = (N, R) \models \text{RCA}_0^* + \mathcal{T} + \bigcup_{k \in \omega} \text{WF}^0(\omega_k) + \exists X \neg \varphi(X)$.

Extend the language of second order arithmetic with a first-order constant symbol c . Every finite fragment of the theory $\text{RCA}_0^* + \mathcal{T} + \text{WF}^0(\omega_c) + \bigcup_{k \in \omega} (c > k) + \exists X \neg \varphi(X)$ also has a model: take $k \in \omega$ such that no axiom of the form $c > k'$ for some $k' \geq k$ is in the fragment and consider the model (N, R, k) (i.e., the model $\mathcal{N}$ in which c is interpreted by k). So, by compactness (Theorem 5.2.7) and the Löwenheim–Skolem theorem (Theorem 5.2.8), there exists a countable model $\mathcal{M} = (M, S) \models \text{RCA}_0^* + \mathcal{T} + \text{WF}^0(\omega_c) + \bigcup_{k \in \omega} (c > k) + \exists X \neg \varphi(X)$ which can be assumed to be topped by some $A \in S$ such that $\mathcal{M} \models \neg \varphi(A)$. Such a model will therefore be non-standard as c is necessarily interpreted by a non-standard integer.

From our assumption, for any countable topped model of $\text{RCA}_0^* + \mathcal{T} + \text{WF}^0(\omega_{d+\ell})$, any $i < n$ and any instance of P_i in this model, we can find a countable topped ω -extension that will be a model of $\text{RCA}_0^* + \mathcal{T} + \text{WF}^0(\omega_d)$ and will contain a solution to that instance.

So by iterating that reasoning as in the proof of Theorem 9.2.3, we can build a sequence (indexed by the standard integers) $\mathcal{M}_0 = \mathcal{M} \subseteq \mathcal{M}_1 \subseteq \dots$ of countable topped models such that $\mathcal{M}_s \models \text{RCA}_0^* + \mathcal{T} + \text{WF}^0(\omega_{c-\ell s})$ for every $s \in \omega$ and such that, for every $i < n$, every instance of P_i appearing inside one $\mathcal{M}_s$ will eventually have a solution in one $\mathcal{M}_t$. As c is non-standard, ℓ is standard, and every such s is standard, $c - \ell s$ is non-standard and every $\mathcal{M}_s$ is a model of $\bigcup_{k \in \omega} \text{WF}^0(\omega_k)$. The model $\bigcup_{s \in \omega} \mathcal{M}_s$ will therefore be a model of $\text{RCA}_0^* + \mathcal{T} + \bigcup_{k \in \omega} \text{WF}^0(\omega_k) + \exists X \neg \varphi(X) + P_0 + \dots + P_{n-1}$.

So $\text{RCA}_0^* + \mathcal{T} + \bigcup_{k \in \omega} \text{WF}^0(\omega_k) + P_0 + \dots + P_{n-1} \not\models \forall X \varphi(X)$. $\square$

Main Theorem 10.1.4. $\text{RT}_2^2 + \text{WKL}$ is Π_1^1 -conservative over $\text{RCA}_0 + \text{BS}_2^0 + \bigcup_{k \in \omega} \text{WF}^0(\omega_k)$.

Proof. For every $\alpha < \epsilon_0$, we have $\omega_2^\alpha > \sup_{\beta < \alpha} \omega^{\beta \times \omega}$. Hence, we can apply Theorem 10.4.2 with Proposition 10.2.5, Proposition 10.2.13, Proposition 10.3.1, and Proposition 10.3.18 to get the Π_1^1 -conservativity of $\text{RT}_2^2 + \text{WKL}$ over $\text{RCA}_0 + \text{BS}_2^0 + \bigcup_{k \in \omega} \text{WF}^0(\omega_k)$. $\square$

Remark 10.4.3. The combination of Proposition 10.2.5, Proposition 10.3.1 and Proposition 10.3.18 provides a so-called $\ll_2$ -basis theorem for RT_2^2 while preserving $\text{WF}^0(\epsilon_0)$ (see [46, Definition 4.9]). As shown by Ikari [67] in his dissertation, this implies that there exists a polynomial proof transformation for the Π_1^1 -conservation result of Main Theorem 10.1.3. On the other hand, it is not clear whether there exists a polynomial proof transformation for Main Theorem 10.1.4.

Chapter 11

Bounded Ramsey theorem

In this chapter, we pursue the study of BRT_2^2 , a degenerate version of Ramsey's theorem for pairs in which the homogeneous sets for color 1 are of bounded size. The study of BRT_2^2 is part of a larger project started and driven by Frittaion, who introduced the bounded Ramsey theorem for colorings of arbitrary tuples. This project was pursued by Soldà [134, Section 4.1], and then independently by Belanger, Shafer, and Yokoyama (unpublished).

BRT_2^2 was introduced in Chapter 7 as the base case for the induction on all separable permutations, and we now recall it:

Statement 7.4.1 (Bounded Ramsey theorem for pairs). BRT_2^2 is the statement “For every coloring $f : [\mathbb{N}]^2 \rightarrow 2$ and every $\ell \in \mathbb{N}$ such that there is no f -homogeneous set for color 1 of size ℓ , there is an infinite f -homogeneous set for color 0.”

BRT_2^2 is an immediate consequence of RT_2^2 in RCA_0 . In fact, Soldà [134, Theorem 4.1.10] showed that EM is sufficient to imply BRT_2^2 , and, as we will see in this chapter, it is not a consequence of ADS .

Main Theorem 11.0.1. $\text{RCA}_0 + \text{ADS} \not\vdash \text{BRT}_2^2$

BRT_2^2 is significantly weaker than RT_2^2 . From a computational viewpoint, contrary to Ramsey's theorem for pairs and to most of the statements studied so far (WKL , SEM , SADS , COH , RRT_n^k , $\dots$), BRT_2^2 is computably true, meaning that every computable instance of it has a computable solution.

The statement “ BRT_2^2 is computably true” is a first-order statement (more precisely, it is a Π_4 -statement). This follows because Σ_1 - and Π_1 -formulas (or, alternatively, Turing indexes) can be encoded as integers and quantified over within first-order arithmetic.

We are mostly interested in the relativized version of that statement: “for every set X , every X -computable instance of BRT_2^2 has an X -computable solution”. This formulation implies BRT_2^2 over RCA_0 , and, from now on, when we say “ BRT_2^2 is computably true”, we will be referring to that relativized version. This is a Π_1^1 -statement (and more precisely, a $\forall\Pi_4^0$ -statement).

Thus, the strength of the statement “ BRT_2^2 is computably true” will only depend on its first-order part, since it does not entail the existence of any non-computable set and holds in any standard model of Δ_1^0 -comprehension.

By an old argument of Hirst [64, Theorem 6.8], $\text{RCA}_0 \vdash \text{BRT}_2^2 \rightarrow \text{B}\Sigma_2^0$, so any proof that BRT_2^2 is computably true must use at least Δ_2^0 -induction. In fact, the best known proofs, due to Belanger and Yokoyama (see Lemma 7.4.2 for Belanger's proof and Soldà [134, Lemma 4.1.9] for Yokoyama's proof), require Σ_2^0 -induction, which we will show to be necessary by proving the following theorem:

Main Theorem 11.0.2 (Short version). *Every model of $\text{RCA}_0 + \text{B}\Sigma_2^0 + \neg\text{I}\Sigma_2^0$ contains an instance f of BRT_2^2 with no $\Delta_1(f)$ -definable solution.*

In particular, as $\text{RCA}_0 + \text{RT}_2^2 \not\vdash \text{I}\Sigma_2^0$ (by Chong, Slaman, and Yang [23], and also by the results of the previous chapter), there exist models of $\text{RCA}_0 + \text{B}\Sigma_2^0 + \neg\text{I}\Sigma_2^0 + \text{RT}_2^2$ and hence models of BRT_2^2 that contain an instance of BRT_2^2 with no solutions computable in that instance. Hence, the pure existence of a solution to an instance of BRT_2^2 is strictly weaker than the existence of a computable solution. It also follows that $\text{RCA}_0 + \text{B}\Sigma_2^0 \not\vdash \text{BRT}_2^2$, hence the strength of BRT_2^2 lies strictly in between $\text{I}\Sigma_2^0$ and $\text{B}\Sigma_2^0$.

This situation already appears in the literature, with the tree theorem for singletons [30, 21] and bounded-width weak König's lemma [130]. Over $\text{RCA}_0 + \text{B}\Sigma_2^0$, both statements are strictly in between $\text{I}\Sigma_2^0$ and $\text{B}\Sigma_2^0$, and their computably true version is equivalent to $\text{I}\Sigma_2^0$.

11.1 $\text{B}\Sigma_2^0$ does not imply BRT_2^2

The goal of this section is to prove Main Theorem 11.0.2 and its consequences. The constructed instance f of BRT_2^2 with no $\Delta_1(f)$ -definable solution will also serve as the BRT_2^2 -instance witnessing that $\text{RCA}_0 + \text{ADS} \not\vdash \text{BRT}_2^2$ in Section 11.2.

We will be working in models of $\text{RCA}_0 + \text{B}\Sigma_2^0 + \neg\text{I}\Sigma_2^0$. By Proposition 9.1.4, every such model $\mathcal{M} = (M, S)$ admits a proper Σ_2^0 cut. That is, there exists a Σ_2^0 -formula $\varphi(x) := \exists y \forall z \theta(x, y, z)$ such that the set $I := \{x \in M : \mathcal{M} \models \varphi(x)\}$ contains 0, is closed under successor, and is downward closed (i.e., whenever $a \in I$ and $b \leq a$, then $b \in I$), yet is not equal to M and there exists some $k \in M$ with $I < k$.

Similarly to the situation described in Section 10.2.1, if we consider, for each $x \in I$, the least witnesses $y \in M$ such that $\forall z \theta(x, y, z)$ holds, then those witnesses cannot be bounded by any $\ell \in M$. Otherwise, $\varphi(x)$ would be equivalent to $\exists y \leq \ell \forall z \theta(x, y, z)$ which would make the cut Δ_2^0 , contradicting our assumption that $\mathcal{M} \models \text{RCA}_0 + \text{B}\Sigma_2^0$.

The function g which, given $x \in I$, searches for the least $y \in M$ such that $\forall z \theta(x, y, z)$ holds is thus a Σ_2^0 function, defined on I with an M -unbounded range. Such a function $g : I \rightarrow M$ is said to be *cofinal* (in M).

g can be approximated by a Δ_1^0 sequence of functions $(g_s)_{s \in M}$ with $g_s : k \rightarrow M$ defined by letting $g_s(x)$ be the least $y \leq s$ such that $\forall z < s \theta(x, y, z)$ holds, or equal to s if no such y is found. For every $x < k$, the sequence $(g_s(x))_{s \in M}$ is non-decreasing, and is eventually constant if and only if x belongs to I , in which case the limit value is $g(x)$.

By twisting the definition of g and of the g_s a little, we can assume that they satisfy more properties:

Proposition 11.1.1. *Let $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \neg\text{I}\Sigma_2^0$, then there exists some $X \in S$ and some $\Sigma_2(X)$ -definable proper cut $I \subseteq M$ with $I < k$ for some $k \in M$ as well as some $\Sigma_2(X)$ -definable function $g : I \rightarrow M$ and some uniformly $\Delta_1(X)$ -definable functions $g_s : k \rightarrow M$ for every $s \in M$ such that:*

- g is increasing and cofinal in M .
- For every $s \in M$, g_s is non-decreasing, and for every $x < k$, $s \mapsto g_s(x)$ is non-decreasing.
- For every $x < k$, $\lim_{s \rightarrow \infty} g_s(x)$ exists if and only if $x \in I$, in which case the limit is equal to $g(x)$.
- For every $s, x \in M$, $g_s(x) \leq s$, and, if $g_{s+1}(x) \neq g_s(x)$ then $g_{s+1}(x) = s + 1$.

These functions will be of central importance for proving Main Theorem 11.0.2.

In what follows, recall that W_e^X is the e -th computably enumerable set. The core of the argument lies in the following technical lemma:

Lemma 11.1.2. $\text{I}\Sigma_1^0$ proves that, for every set X , every $b \in \mathbb{N}$ and $n \in \mathbb{N}$, there exists a $\Delta_1(X)$ coloring $f_n^b : [\mathbb{N}]^2 \rightarrow 2$ such that:

- There is no set of size 3 that is f_n^b -homogeneous for the color 1.
- For every $e < n$ such that W_e^X is infinite, there exists some $x, y \in W_e^X$ such that $b \leq x < y$ and $f_n^b(x, y) = 1$.

Furthermore, the construction is uniform in n and b .

Proof. The coloring f_n^b will be defined by stages, with $f_n^b(x, s)$ being defined at stage s for every $x < s$. At any point of the construction, any $x \in \mathbb{N}$ will be in state “0” or “1”, meaning that at any further stage s , $f_n^b(x, s)$ will be set accordingly. By default, every $x \in \mathbb{N}$ has the state “0”. We consider a trash $T_s \subseteq [0, s]$ where the elements leaving state “1” will be put, to ensure that they will never go back to that state. Our construction will ensure that $\text{card } T_s \leq n^2$ at every stage s .

For every $e < n$, we say that e is *active* at a stage s if $\text{card}(W_e^X[s] \cap [b, s]) > n^2$. Once an element is active, it stays active for the rest of the construction.

Construction: At stage s , if there exists some new $e < n$ that became active at this stage, we then put every $x < s$ that is in state “1” in state “0” and put them in the trash T_s . Then, for every $e' < n$ that is active, we pick an element in $W_{e'}^X[s] \cap [b, s] \setminus T_s$ and put it in state “1”. If no such $e < n$ was found, then all the elements stay in the same state and no new element is added to the trash.

Finally, whether or not a new e was found, we let $f_n^b(x, s) = c$ for every $c < 2$ and every $x < s$ in state “c”.

Claim 1: $\text{card } T_s \leq n^2$ for every $s \in \mathbb{N}$, hence the construction is well-defined. Indeed, there are always at most n elements in state “1” at a given time (one per active e). And such elements are put into the trash at most n times during the construction (once every time a new $e < n$ became active). Therefore, we can always pick an element in $W_{e'}^X[s] \cap [b, s] \setminus T_s$ if $W_{e'}^X$ is active at stage s .

Claim 2: There is no set of size 3 that is f_n^b -homogeneous for the color 1. Assume by contradiction that we have such a set $\{x < y < z\}$. By construction, x was in state “1” at stage y for $f_n^b(x, y)$ to be equal to 1. Then, for $f_n^b(y, z)$ to be equal to 1, we need y to be in state “1” at stage z , but y can only enter state “1” after stage y , which would put x in the trash, and make it so that $f_n^b(x, z) = 0$.

Claim 3: For every $e < n$ such that W_e^X is infinite, there exists some $x, y \in W_e^X$ such that $b \leq x < y$ and $f_n^b(x, y) = 1$. If W_e^X is infinite, then by $\text{I}\Sigma_1^0$, there is some stage s such that $\text{card}(W_e^X[s] \cap [b, s]) > n^2$, hence e is eventually active. Then, for every $y > s$, there will be some $x \in W_e^X$ that will be in state “1” at stage y , which gives $f_n^b(x, y) = 1$. Since W_e^X is infinite, there is one such y in W_e^X . $\square$

We are now ready to prove our first main theorem for this chapter.

Main Theorem 11.0.2. For every model $\mathcal{M} \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \neg \text{I}\Sigma_2^0$ and every set $Z \in \mathcal{M}$, there is a BRT_2^2 -instance $f \in \mathcal{M}$ with no $\Delta_1(f \oplus Z)$ -definable solution.

Proof. Let $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \neg \text{I}\Sigma_2^0$ and $Z \in S$. Let $X \in S$, $I \subseteq M$ the proper $\Sigma_2(X)$ cut bounded by some $k \in M$, $g : I \rightarrow M$ the cofinal $\Sigma_2(X)$ function and $(g_s)_{s \in M}$ its $\Delta_1(X)$ -approximation obtained in Proposition 11.1.1.

For every $b, n \in \mathbb{N}$, let $f_n^b : [M]^2 \rightarrow 2$ be the $\Delta_1(X \oplus Z)$ -coloring defined in Lemma 11.1.2. For every $i < k$, let $f^i : [M]^2 \rightarrow 2$ be defined by stages as follows:

Construction: At stage s : for every $i < k$, let $b_s^i \leq s$ be the last $b \leq s$ such that the value of $g_b(i)$ has changed, if it exists, otherwise, $b_s^i = 0$. Then let

$f^i(x, s) = f_{g_s(i)}^{b_s^i}(x, s) = f_{b_s^i}^{b_s^i}(x, s)$ for every $x < s$.

Finally, let $f : [M]^2 \rightarrow 2$ be defined by $f(x, y) = 1$ iff there exists some $i < k$ such that $f^i(x, y) = 1$ for every $x < y \in M$.

Claim 1: For every $i < k$, there is no f^i -homogeneous set for color 1 of size 3. Notice first that for every $n, b \in \mathbb{N}$, we have $f_n^b(x, y) = 0$ when $x < b$, therefore, if $f^i(x, y) = 1$ for some $x, y \in \mathbb{N}$, then b_y^i must be smaller than x and therefore is equal to b_x^i . This means that every f^i -homogeneous set for color 1 is also homogeneous for some coloring f_n^b , hence that no such set of size three can exist, by our assumptions on the f_n^b .

Claim 2: There is some $\ell \in M$ such that there is no f -homogeneous set for color 1 of size ℓ . Since $\mathcal{M} \models \text{IS}_1^0$, the finite Ramsey theorem for pairs and k colors is true in $\mathcal{M}$. Let $\ell \in M$ be such that every finite coloring $h : [F]^2 \rightarrow k$ admits some h -homogeneous subset of size 3 if $\text{card } F \geq \ell$. If some set $F = \{x_0, \dots, x_{\ell-1}\}$ were f -homogeneous for color 1, then, by definition of f , the coloring $h : [F]^2 \rightarrow k$ sending every pair x_a, x_b to some $i < k$ such that $f^i(x_a, x_b) = 1$ would be well-defined, and by definition of ℓ , there would exist some $i < k$ and some f^i -homogeneous set for the color 1 of size 3, contradicting Claim 1.

Claim 3: There is no $\mathcal{M}$ -infinite set which is simultaneously f -homogeneous and $\Sigma_1(X \oplus Z)$ (and a fortiori $\Delta_1(X \oplus Z)$) for color 0 in S . Assume otherwise, and let $e \in M$ be such that $W_e^{X \oplus Z}$ is M -unbounded and f -homogeneous for color 0. By definition of g , there exists some $i \in I$ such that $g(i) > e$. By assumption on the g_s , the smallest stage $s \in \mathbb{N}$ such that $g_s(i) = g(i)$ is $s = g(i)$. Pick some $x, y \in W_e^{X \oplus Z}$ such that $g(i) \leq x < y$ and such that $f_{g(i)}^{g(i)}(x, y) = 1$, then, we have $f^i(x, y) = f_{g(i)}^{g(i)}(x, y) = 1$, hence $f(x, y) = 1$ and $W_e^{X \oplus Z}$ is not f -homogeneous for color 0. $\square$

Corollary 11.1.3. $\text{RCA}_0 + \text{BS}_2^0 \not\vdash \text{BRT}_2^2$.

Proof. Let M be a countable model of $\text{IS}_1 + \text{BS}_2 + \neg \text{IS}_2$. Let $\mathcal{M} = (M, S)$ be the model whose second-order part S consists of the Δ_1 -definable sets with parameters in M . By Friedman [48], $\mathcal{M} \models \text{RCA}_0 + \text{BS}_2^0 + \neg \text{IS}_2^0$, but by Main Theorem 11.0.2, $\mathcal{M} \not\models$ “for every coloring $f : [\mathbb{N}]^2 \rightarrow 2$ and every $\ell \in \mathbb{N}$ such that there is no f -homogeneous set for color 1 of size ℓ , there is an infinite f -homogeneous set for color 0.” $\square$

Corollary 11.1.4. The statement “ BRT_2^2 is computably true” is equivalent to IS_2^0 over RCA_0 .

Proof. By Lemma 7.4.2, we have $\text{RCA}_0 + \text{IS}_2^0 \vdash$ “ BRT_2^2 is computably true”. In fact, we can directly prove the more general statement that, if IS_2^0 implies a Π_2^1 -problem P over RCA_0 , then $\text{RCA}_0 + \text{IS}_2^0 \vdash$ “ P is computably true”:

Consider such a statement P and let $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{IS}_2^0$ and let $f \in S$ be an instance of P . By Friedman [48], $\mathcal{M}_f = (M, \Delta_1\text{-Def}(f)) \models \text{IS}_2^0$, where $\Delta_1\text{-Def}(f)$ denotes the $\Delta_1(f)$ -definable subsets of M . By our assumption, $\mathcal{M}_f \models P$, so there is some $H \in \Delta_1\text{-Def}(f)$ such that $\mathcal{M}_f \models$ “ H is a P -solution to f ”. In particular, $H \in S$, so $\mathcal{M} \models$ “there is a $\Delta_1(f)$ -definable P -solution to f .” Thus, $\mathcal{M} \models$ “ P is computably true”.

Suppose now that $\mathcal{M} \models \text{RCA}_0 +$ “ BRT_2^2 is computably true”. In particular $\mathcal{M} \models \text{BRT}_2^2$. By Hirst [64], $\text{RCA}_0 \vdash \text{BRT}_2^2 \rightarrow \text{BS}_2^0$, so $\mathcal{M} \models \text{BS}_2^0$. Then, by Main Theorem 11.0.2, $\mathcal{M} \models \text{IS}_2^0$. $\square$

11.2 ADS does not imply BRT_2^2

The goal of this section is to prove Main Theorem 11.0.1. To this end, we will construct a model of $\text{RCA}_0 + \text{ADS}$ containing an instance of BRT_2^2 with no solution.

The construction of such a model will be similar to that in the proof of Theorem 9.2.3: We will build a sequence $\mathcal{M}_0 \subseteq \mathcal{M}_1 \subseteq \dots$ of ω -extensions of countable topped models of $\text{RCA}_0 + \text{BS}_2^0$ such that every instance of SADS or COH appearing in some $\mathcal{M}_i$ will eventually have a solution in some $\mathcal{M}_j$.

Our initial model $\mathcal{M}_0$ will be the model $\text{RCA}_0 + \text{BS}_2^0 + \neg \text{IS}_2^0$ constructed in Corollary 11.1.3. It contains a Δ_1 instance f of BRT_2^2 with no Δ_1 solution and we will make sure to avoid adding any solution to f in the $\mathcal{M}_i$ so that $\bigcup_{i \in \omega} \mathcal{M}_i$ does not contain any and is not a model of BRT_2^2 .

In the previous chapter, we introduced the combinatorics needed to add solutions to instances of SADS , SEM , COH and WKL to models of $\text{RCA}_0 + \text{BS}_2^0$ satisfying appropriate well-foundedness assumptions. As it was noted in Remark 10.3.29, these well-foundedness assumptions were only necessary in the case of SEM (as COH , SADS and WKL are already known to be Π_1^1 -conservative over $\text{RCA}_0 + \text{BS}_2^0$ alone). Thus, it remains to introduce the combinatorics required for adding solutions G to instances of SADS and COH to the $\mathcal{M}_i$ while ensuring that G will not compute any M -unbounded f -homogeneous set.

11.2.1 Immunity

By its definition, the coloring f satisfies properties other than simply having no $\Delta_1(f)$ -definable solutions that we will exploit in our construction. One such property is the fact that it is stable.

Proposition 11.2.1. *The coloring f constructed in the proof of Main Theorem 11.0.2 is stable (i.e., for every $x \in \mathbb{N}$, $\lim_{y \rightarrow \infty} f(x, y)$ exists).*

Proof. For every $b, n \in \mathbb{N}$, the coloring f_n^b defined in Lemma 11.1.2 is stable. Indeed, by our use of the trash T_s , every $x \in \mathbb{N}$ can change state at most twice. Therefore, every $x \in \mathbb{N}$ eventually reaches a final state, which determines its limit behavior.

Then, all the f^i constructed in Main Theorem 11.0.2 are stable. Fix $x \in M$. Either there exists some stage s such that $b_s^i > x$, in which case $f^i(x, y) = 0$ for all $y \geq s$ (because, by definition of f_n^b , we have $f_n^b(x, y) = 0$ whenever $x < b$), or no such stage exists. In the latter case, we must have $i \in I$, and the sequence $(b_s^i)_{s \in \mathbb{N}}$ eventually reaches a final value b such that $g_b(i) = g(i)$. Hence, for sufficiently large s , we have $f^i(x, s) = f_{g(x)}^b(x, s)$, and $f(x, \cdot)$ and $f_{g(x)}^b(x, \cdot)$ have the same limit behavior.

Finally, f is stable. Fix $x \in \mathbb{N}$. For each $i < k$ there exists a stage b_i beyond which $f^i(x, \cdot)$ has reached its limit behavior. By BS_2^0 , there is a uniform bound b after which all the $f^i(x, \cdot)$ have reached their limit behavior for $i < k$. Thus, f has also reached its limit behavior at stage b . $\square$

For each $i < 2$, let $A_i := \{x \in M : \forall^\infty y f(x, y) = i\}$ be the set of elements having limit behavior i . Each set A_i is Σ_2 , and since $A_0 \sqcup A_1 = M$ by stability of f , both are in fact Δ_2 . By the discussion following Statement 6.2.7 concerning the equivalence between the stable Ramsey theorem for pairs and the statement D_2^2 , every M -unbounded subset G of either A_0 or A_1 computes an M -unbounded f -homogeneous set.

The assumption that f admits no Δ_1 solution ensures that A_0 has no Δ_1 M -unbounded subset. In particular, A_1 must itself be M -unbounded, since otherwise there would exist some $k \in M$ such that $\{k, k+1, k+2, \dots\} \subseteq A_0$, yielding a Δ_1 M -unbounded subset of A_0 . Nevertheless, our construction cannot add an M -unbounded subset of A_1 . Indeed, let $\ell \in M$ be such that there is no M -finite f -homogeneous set for color 1 of size ℓ . Then every model $\mathcal{M} \models \text{RCA}_0$ with first-order part M containing an M -unbounded subset of A_1 would also contain an M -unbounded f -homogeneous set for the color 1. By IS_1^0 , this is impossible as every $\mathcal{M}$ -infinite set has an M -finite set of size ℓ , contradicting the fact that there is no M -finite f -homogeneous set for the color 1 of size ℓ .

The property for an infinite set of not having an infinite computable set is called *immunity*.

Definition 11.2.2 (immune). Fix a set Z . An infinite set $X \subseteq \mathbb{N}$ is said to be *Z-immune* if it has no infinite Z -computable subset (or equivalently no infinite Z -computably enumerable subset).

The goal is thus to ensure that A_0 remains G -immune after adding a solution G to our model. To be able to force this property, we require a more effective notion of immunity, in which we additionally restrict the finite subsets of X whose cardinalities are too large.

Definition 11.2.3. Fix a function $h : \mathbb{N} \rightarrow \mathbb{N}$ and a set Z . A set A is said to be *h-effectively Z-immune* if for every Turing index $e \in \mathbb{N}$, $\text{card } W_e^Z \geq h(e)$ implies $W_e^Z \not\subseteq A$.

A_0 can be shown to be h -effectively immune for some $\emptyset'$ -computable function h . In order to execute a Shore blocking argument, we need to consider a slight modification of the definition of an h -effectively Z -immune function in which the function h is constant on large blocks. The adequate notion of immunity which we will be able to preserve will be the following, where $2_n(a)$ is defined inductively by $2_0(a) = 1$ and $2_{n+1}(a) = 2^{2_n(a)}$ for $a \in \mathbb{N}$.

Definition 11.2.4. Fix a set Z , an unbounded $\Delta_2(Z)$ set B and some standard $n \in \omega$. A set $A \subseteq M$ is *B-block Z-immune with cost n* if for every $b \in B$, and every $x < b$, $\text{card } W_x^Z \geq 2_n(b)$ implies $W_x^Z \not\subseteq A$.

For every set Z , every $n \in \omega$ and every $\Delta_2(Z)$ set B , if a set A is B -block Z -immune with cost n , then A is h -effectively Z -immune for some Z' -computable function h . In particular, A is Z -immune.

Proposition 11.2.5. Let f be the coloring constructed in the proof of Corollary 11.1.3 and $A_0 = \{x \in M : \forall^\omega y f(x, y) = 0\}$. Then, there exists some M -unbounded Δ_2 set $B \subseteq M$ such that A_0 is B -block immune with cost 1.

Proof. Let I be the Σ_2 cut used in the construction of f , and let B be the image of the corresponding Σ_2 -definable cofinal function $g : I \rightarrow M$. Let $(g_s)_{s \in M}$ be the Δ_1 approximation of g .

Fix $b \in B$. By definition of B , there exists some $i \in I$ such that $b = g(i)$. By Proposition 11.1.1, the least stage $s \in M$ such that $g_s(i) = g(i)$ is $s = b$. Hence, the coloring f^i constructed in Main Theorem 11.0.2 satisfies $f^i(x, y) = f_b^b(x, y)$ for every $y \geq b$. Consequently, if $x < b$ is such that $\text{card } W_x \geq 2^{2^b} \geq b + b^2$, then $\text{card}(W_x \cap [b, +\infty)) \geq b^2$. Therefore, $\text{card } W_x \cap [b, +\infty)$ must contain an element having limit behavior 1 for f^i and thus also for f , by the definition of f_b^b . $\square$

11.2.2 The SADS step

We now prove the existence of ω -extensions which add solutions to instances of SADS while preserving the fact that the instance of BRT_2^2 has no solution.

Proposition 11.2.6. Let $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0$ be a countable model topped by some set Y . Let B be an M -unbounded $\Delta_2(Y)$ set and fix $n \in \omega$. Then, for every set $A \subseteq M$ that is B -block Y -immune with cost n and every instance $\mathcal{L} = (M, \leq_{\mathcal{L}})$ of SADS in S , there exists some M -unbounded set $G \subseteq M$ and some M -unbounded $\Delta_2(Y)$ set $\hat{B}$ such that:

- G is an M -infinite $\mathcal{L}$ -ascending or $\mathcal{L}$ -descending sequence;
- $Y' \geq_T (G \oplus Y)'$ and in particular $\mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0$;
- A is $\hat{B}$ -block $(G \oplus Y)$ -immune with cost $n + 2$.

Proof. As in the proof of Proposition 10.3.18, we consider $U^0 \subseteq M$ and $U^1 \subseteq M$, the sets of elements that have a finite amount of predecessors and successors, respectively. We also use the same notion of conditions, split pairs and the same forcing relations.

The difference is now that the requirements concerning the well-foundedness of an ordinal α are replaced by requirements ensuring that A stays sufficiently immune.

We want our construction to force for some $i < 2$ the following requirements for every $k \in M$:

- $\mathcal{R}_k^i$: $(G^i \oplus Y)' \upharpoonright k$ is decided.
- $\mathcal{S}_k^i$: there exists some $\ell > 2^{2^k}$ in B , such that, for every $x < \log_2(\log_2(\ell))$, if $\text{card } W_x^{G^i \oplus Y} \geq 2_n(\ell)$, then, there exists some $e < \ell$ such that $\text{card } W_e^Y \geq 2_n(\ell)$ and $W_e^Y \subseteq W_x^{G^i \oplus Y}$ (in particular, since A is B -block Y -immune with cost n , this implies that $W_e^Y \not\subseteq A$, hence that $W_x^{G^i \oplus Y} \not\subseteq A$).
- $\mathcal{T}_k^i$: There exists some $x > k$ such that $x \in G^i$.

The requirements $\mathcal{S}_k^i$ are $\Sigma_2^0(G)$ -formulas, whereas the relations $\Vdash_i$ have only been defined for $\Sigma_1^0(G)$ - and $\Pi_1^0(G)$ -formulas. To handle this case, we write $(\sigma^0, \sigma^1) \Vdash_i \exists y \Phi_e^{G^i \oplus Y}(y) \uparrow$ if there exists some $n \in M$ such that $(\sigma^0, \sigma^1) \Vdash_i \Phi_e^{G^i \oplus Y}(n) \uparrow$.

Again, we will use a pairing argument, and ensure that for every $k \in M$, the following requirements are met: $\mathcal{R}_k^0 \vee \mathcal{R}_k^1$, $\mathcal{R}_k^0 \vee \mathcal{S}_k^1$, $\mathcal{S}_k^0 \vee \mathcal{R}_k^1$, $\mathcal{S}_k^0 \vee \mathcal{S}_k^1$, and $\mathcal{T}_k^0 \wedge \mathcal{T}_k^1$, and therefore, that $(\mathcal{R}_k^0 \wedge \mathcal{S}_k^0 \wedge \mathcal{T}_k^0) \vee (\mathcal{R}_k^1 \wedge \mathcal{S}_k^1 \wedge \mathcal{T}_k^1)$ is met for every $k \in M$. Then, there will be one side $i < 2$ such that $\mathcal{R}_k^i \wedge \mathcal{S}_k^i \wedge \mathcal{T}_k^i$ is satisfied for M -unboundedly many values of k , hence that all requirements are satisfied for this side.

For simplicity of notation, the elements $b \in B$ will all be assumed to be of the form 2^{2^k} for some $k \in M$, which will then be used to construct $\hat{B}$ such that $\hat{B} \subseteq \{\log_2(\log_2(b)) : b \in B\}$. The general case requires some care in handling integer parts and inequalities, but does not change the argument.

We will define two sets G^0 and G^1 using an infinite $\Delta_1(Y')$ decreasing sequence of conditions and then pick G to be equal to one of them. We will also define two $\Delta_1(Y')$ sets B^i for $i < 2$ so that, if $G = G^i$, then A is B^i -block $(G \oplus Y)$ -immune with cost $n+2$.

The tree T obtained in Lemma 10.3.25 is given by the limit of a uniformly Y -computable increasing sequence $(T_s)_{s \in M}$ of M -finite (σ^0, σ^1) -trees. This sequence can be uniformly Y -computed in the pair $(\varphi_0(G^0, x), \varphi_1(G^1, x))$ and in the condition (σ^0, σ^1) . Thus, there exists some computable function $e : M \rightarrow M$ such that $W_e^Y(\langle \ulcorner \varphi_0 \urcorner, \ulcorner \varphi_1 \urcorner, \sigma^0, \sigma^1, \rho, i, x \urcorner \rangle)$ searches the split pair $(\sigma_\rho^0, \sigma_\rho^1)$ in position ρ of the corresponding (σ^0, σ^1) -tree T , and, if such a split pair exists, outputs the elements of $W_x^{\sigma_\rho^i \oplus Y}$.

The pair $(\hat{\sigma}^0, \hat{\sigma}^1)$ constructed in the proof of Lemma 10.3.26 has the property that, for $i < 2$ the side such that $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \neg \varphi^i(G^i, p_i + 1)$, $\hat{\sigma}^i$ is an element of some split pair appearing in T , hence there exists some $\rho \in T$ such that $\hat{\sigma}^i = \sigma_\rho^i$.

Lemma 11.2.7. *Let $k \in M$ and (σ^0, σ^1) be a condition. Then, for every requirements $\mathcal{U}_k^0 \in \{\mathcal{R}_k^0, \mathcal{S}_k^0\}$ and $\mathcal{U}_k^1 \in \{\mathcal{R}_k^1, \mathcal{S}_k^1\}$, there exists some extension $(\hat{\sigma}^0, \hat{\sigma}^1) \leq (\sigma^0, \sigma^1)$ and some $i < 2$ such that $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \mathcal{U}_k^i$. Such an extension is uniformly Y' -computable in $\mathcal{U}_k^0, \mathcal{U}_k^1$ and (σ^0, σ^1) .*

Proof. We will define two $\Sigma_1^0(Y)$ -formulas $\Phi_t(G, x)$ and $\Psi_t(G, x)$ as follows:

Let $\Phi_t(G, x)$ be the formula: $\exists e_0 < e_1 < \dots < e_x < t$ such that $\Phi_{e_i}^{G \oplus Y}(e_i) \downarrow$ for every $i \leq x$, and let $\Psi_t(G, x)$ be the formula: $\exists e_0 < e_1 < \dots < e_x < t$ such that $\text{card } W_{e_i}^{G \oplus Y} \geq 2_{n+2}(t)$ for every $i \leq x$.

Consider the function $S : M \rightarrow M$ which maps t to the largest value taken by $e(\ulcorner \varphi_0 \urcorner, \ulcorner \varphi_1 \urcorner, \sigma^0, \sigma^1, \rho, i, x)$ for every $\varphi_0, \varphi_1 \in \{\Phi_t, \Psi_t\}$, every $\rho \in 2^{\leq 2(t-1)}$, every $i < 2$ and every $x < t$.

Assuming a reasonable encoding of e and of Φ_t and Ψ_t , this function is eventually dominated by $t \mapsto 2^{2^t}$, indeed, all the possible parameters of e either directly depend on t (such as $\ulcorner \varphi_0 \urcorner$ and $\ulcorner \varphi_1 \urcorner$), are smaller than t (such as x) or are in $2^{\leq 2(t-1)}$ (like ρ), hence, in any case they can be encoded using αt bits for some $\alpha \in \omega$ and the same holds for the corresponding value taken by e . Hence, there exists some element ℓ in B such that $\ell > 2^{2^k}$ and such that $\ell' := \log_2(\log_2(\ell))$ satisfies $S(\ell') < 2^{2^{\ell'}} = \ell$.

For every $i < 2$, let $\varphi_i(G, x)$ be the formula $\Phi_{\ell'}(G, x)$ if $\mathcal{U}_k^i = \mathcal{R}_k^i$ and $\Psi_{\ell'}(G, x)$ otherwise.

Let T be the maximal $\Sigma_1(Y)$ (σ^0, σ^1)-tree T satisfying $(\varphi_0(G, x), \varphi_1(G, x))$ obtained in Lemma 10.3.25. Since $\Phi_{\ell'}(G, \ell')$ and $\Psi_{\ell'}(G, \ell')$ cannot hold, as there are only ℓ' Turing indexes smaller than ℓ' , the tree T cannot have a depth bigger than $2(\ell' - 1)$. Indeed, any $\rho \in T$ of length $> 2(\ell' - 1)$ would make either $|\rho| - \text{card } \rho$ or $\text{card } \rho$ bigger than $\ell' - 1$, contradicting the fact that T satisfies $(\varphi_0(G, x), \varphi_1(G, x))$.

Therefore, T , as a binary tree, has fewer than $2^{2^{\ell'-1}}$ elements, and the increasing sequence $(T_s)_{s \in M}$ of computable trees approaching T must have a final stage after which the approximation stabilizes and is equal to T , hence T is M -finite.

Therefore, by Lemma 10.3.26, there exists some extension $(\hat{\sigma}^0, \hat{\sigma}^1) \leq (\sigma^0, \sigma^1)$, some $i < 2$ and some $p_i \leq \ell'$ such that:

- $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \varphi_i(G^i, p_i - 1)$
- $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \neg \varphi_i(G^i, p_i)$.

Again with the convention that $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \varphi_i(G^i, -1)$ always holds.

There are two cases:

Case 1: If $\mathcal{U}_k^i = \mathcal{R}_k^i$, then $\varphi_i(G, x) = \Phi_{\ell'}(G, x)$, and, since $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \varphi_i(G^i, p_i - 1)$, there exist indices $e_0 < \dots < e_{p_i-1} < \ell'$ such that $\Phi_{e_x}^{\hat{\sigma}^i \oplus Y}(e_x) \downarrow$ for every $x < p_i$. Since $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \neg \varphi_i(G^i, p_i)$, then, for every $y < \ell'$ not equal to one of the e_x , the condition $(\hat{\sigma}^0, \hat{\sigma}^1)$ forces $\Phi_y^{G^i \oplus Y}(y) \uparrow$. Hence, the condition $(\hat{\sigma}^0, \hat{\sigma}^1)$ forces $\mathcal{R}_k^i$ as $\ell' \geq k$.

Case 2: If $\mathcal{U}_k^i = \mathcal{S}_k^i$, then $\varphi_i(G, x) = \Psi_{\ell'}(G, x)$, and, since $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \varphi_i(G^i, p_i - 1)$, there exist indices $e_0 < \dots < e_{p_i-1} < \ell'$ such that $\text{card } W_{e_x}^{\hat{\sigma}^i \oplus Y} \geq 2_{n+2}(\ell') = 2_n(\ell)$ for every $x < p_i$. Since $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \neg \varphi_i(G^i, p_i)$, then, for every $y < \ell'$ not equal to one of the e_x , the condition $(\hat{\sigma}^0, \hat{\sigma}^1)$ forces $\text{card } W_y^{G^i \oplus Y} < 2_{n+2}(\ell')$.

Let $x < p_i$ and let ρ be the index of T such that $\hat{\sigma}^i = \sigma_\rho^i$. Since the depth of T is smaller than $2\ell' - 1$, $\rho \in 2^{\leq 2(\ell'-1)}$ and, by definition of ℓ' , $e(\ulcorner \varphi_0 \urcorner, \ulcorner \varphi_1 \urcorner, \sigma^0, \sigma^1, \rho, i, e_x) < 2^{2^{\ell'}} = \ell$. By definition of the function e , $W_{e_x}^{\hat{\sigma}^i \oplus Y}$ equals $W_{e(\ulcorner \varphi_0 \urcorner, \ulcorner \varphi_1 \urcorner, \sigma^0, \sigma^1, \rho, i, e_x)}^Y$, hence,

$$\text{card } W_{e(\ulcorner \varphi_0 \urcorner, \ulcorner \varphi_1 \urcorner, \sigma^0, \sigma^1, \rho, i, e_x)}^Y \geq 2_{n+2}(\ell') = 2_n(\ell)$$

and $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i W_{e(\ulcorner \varphi_0 \urcorner, \ulcorner \varphi_1 \urcorner, \sigma^0, \sigma^1, \rho, i, e_x)}^Y \subseteq W_{e_x}^{G^i \oplus Y}$. Hence, $(\hat{\sigma}^0, \hat{\sigma}^1)$ forces $\mathcal{S}_k^i$. $\square$

Construction We Y' -computably construct a decreasing sequence (σ_s^0, σ_s^1) of conditions. In parallel, the sets $(G^0 \oplus Y)'$ and $(G^1 \oplus Y)'$ are approximated by sequences $(\sigma_s^{0'})$ and $(\sigma_s^{1'})$ of finite binary sequences and each B^i by a sequence (B_s^i) .

Let $\sigma_0^0 = \sigma_0^1 = \sigma_0^{0'} = \sigma_0^{1'} = \varepsilon$ and $B_0^0 = B_0^1 = \emptyset$. Assume that $\sigma_s^0, \sigma_s^1, \sigma_s^{0'}, \sigma_s^{1'}, B_s^0$ and B_s^1 have been defined for some $s \in M$ and let

$$k_s = 1 + \max\{|\sigma_s^0|, |\sigma_s^1|, |\sigma_s^{0'}|, |\sigma_s^{1'}|, |B_s^0|, |B_s^1|\}$$

We will satisfy one of the five requirements $\mathcal{R}_{k_s}^0 \vee \mathcal{R}_{k_s}^1, \mathcal{S}_{k_s}^0 \vee \mathcal{R}_{k_s}^1, \mathcal{R}_{k_s}^0 \vee \mathcal{S}_{k_s}^1, \mathcal{S}_{k_s}^0 \vee \mathcal{S}_{k_s}^1$ or $\mathcal{T}_{k_s}^0 \wedge \mathcal{T}_{k_s}^1$, depending on the value of s modulo 5. There are two cases:

Case 1: The requirement is of the form $\mathcal{U}_{k_s}^0 \vee \mathcal{U}_{k_s}^1$, with $\mathcal{U}_{k_s}^i \in \{\mathcal{R}_{k_s}^i, \mathcal{S}_{k_s}^i\}$. In that case, by Lemma 11.2.7, there exists some $i < 2$ and some extension $(\hat{\sigma}^0, \hat{\sigma}^1) \leq (\sigma_s^0, \sigma_s^1)$ forcing $\mathcal{U}_{k_s}^i$, we then let $\sigma_{s+1}^0 = \hat{\sigma}^0$ and $\sigma_{s+1}^1 = \hat{\sigma}^1$. Then, if $\mathcal{U}_{k_s}^i = \mathcal{R}_{k_s}^i$, let $\sigma_{s+1}^{i'}$ be such that $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash (G^i \oplus Y)' = \sigma_{s+1}^{i'}$ and let $\sigma_{s+1}^{1-i'}, B_{s+1}^0$ and B_{s+1}^1 unchanged. If $\mathcal{U}_{k_s}^i = \mathcal{S}_{k_s}^i$, let $B_{s+1}^i(x) = 0$ for every $x \in [|B_s^i|, \log_2(\log_2(\ell))]$ and $B_{s+1}^i(\log_2(\log_2(\ell))) = 1$, where ℓ is the bound obtained from the fact that $(\hat{\sigma}^0, \hat{\sigma}^1) \Vdash_i \mathcal{S}_{k_s}^i$ and let $\sigma_{s+1}^{0'}, \sigma_{s+1}^{1'}$ and B_{s+1}^{1-i} unchanged.

Case 2: The requirement is of the form $\mathcal{T}_{k_s}^0 \wedge \mathcal{T}_{k_s}^1$. Since both U^0 and U^1 are M -unbounded, there exists some condition $(\tau^0, \tau^1) \leq (\sigma_s^0, \sigma_s^1)$ with $\text{card } \tau^i > \text{card } \sigma_s^i$ for every $i < 2$. Let $\sigma_{s+1}^0 = \tau^0$ and $\sigma_{s+1}^1 = \tau^1$ and leave $\sigma_{s+1}^{0'}, \sigma_{s+1}^{1'}, B_{s+1}^0$ and B_{s+1}^1 unchanged.

Even if the $\sigma_s^0, \sigma_s^1, \sigma_s^{0'}, \sigma_s^{1'}, B_s^0$ and B_s^1 are only defined on some cut I , the values taken by the k_s are unbounded in M , hence for every $k \in M$, there exists some side $i < 2$ such that $\mathcal{R}_k^i \wedge \mathcal{S}_k^i \wedge \mathcal{T}_k^i$ is satisfied. Therefore, there exists some $i < 2$ such that, for every $k \in M$, $\mathcal{R}_k^i \wedge \mathcal{S}_k^i \wedge \mathcal{T}_k^i$ is satisfied.

Let $G = \bigcup_{s \in I} \sigma_s^i$. By construction G is Y' -computable and, as the $\mathcal{R}_k^i$ are satisfied, $(G \oplus Y)'$ is equal to $\bigcup_{s \in I} \sigma_s^{i'}$ and is also Y' -computable, hence $\mathcal{M}[G \oplus Y] = \mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0$. Since the requirements $\mathcal{T}_k^i$ are satisfied, the set G is an M -infinite $\mathcal{L}$ -ascending sequence if $i = 0$ and an M -infinite $\mathcal{L}$ -descending sequence if $i = 1$.

Let $\hat{B} = \bigcup_{s \in I} B_s^i$, then A is $\hat{B}$ -block $(G \oplus Y)$ -immune with cost $n + 2$. Indeed, every $b \in \hat{B}$ was added because some requirement $\mathcal{S}_{k_s}^i$ was forced (which happened M -unboundedly many times, hence B is M -unbounded), hence there exists some $\ell \in B$ such that $b = \log_2(\log_2(\ell))$ and for every $x < b$, if $\text{card } W_x^{G \oplus Y} \geq 2_n(\ell) = 2_{n+2}(b)$ then $W_e^Y \subseteq W_x^{G \oplus Y}$ and $\text{card } W_e^Y \geq 2_n(\ell)$ for some $e < \ell$, which implies that $W_e^Y \not\subseteq A$ and therefore $W_x^{G \oplus Y} \not\subseteq A$ by the B -block Y -immunity with cost n of A .

This completes the proof of Proposition 11.2.6. $\square$

In the proof of Proposition 11.2.6, the construction of the sets G and $\hat{B}$ does not depend on the set A . Therefore, for any other set $\tilde{A} \subseteq M$ that is B -block Y -immune with cost n , we also get that $\tilde{A}$ is $\hat{B}$ -block $(G \oplus Y)$ -immune with cost $n + 2$.

11.2.3 The COH step

We now prove the existence of ω -extensions which add solutions to instances of COH while preserving the fact that the instance of BRT₂² has no solution. Again, we reuse the same formalism as in the proof of Proposition 10.2.5.

Proposition 11.2.8. *Let $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0$ be a countable model topped by some set Y . Let B be an M -unbounded $\Delta_2(Y)$ set and fix $n \in \omega$. Then, for every set $A \subseteq M$ that is B -block Y -immune with cost n , every instance $\vec{R}$ of COH in S and every $P \gg Y'$ such that $\mathcal{M}[P] \models \text{RCA}_0^*$, there exists $G \subseteq M$ and some M -unbounded $(G \oplus Y)'$ -computable set $\hat{B}$ such that:*

- G is $\vec{R}$ -cohesive
- $\mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0$
- A is $\hat{B}$ -block $(G \oplus Y)$ -immune with cost $n + 2$.

Proof. As in the proof of Proposition 10.2.5, we let

$$\vec{R}_\sigma = \bigcap_{\sigma(i)=0} \mathbb{N} \setminus R_i \cap \bigcap_{\sigma(i)=1} R_i$$

for every $\sigma \in 2^{<M}$ and let T be the $\Sigma_1(Y)$ tree $\{\sigma \in 2^{<M} : \exists x(x > |\sigma| \wedge x \in \vec{R}_\sigma)\}$. P computes a path Q through that tree and we then consider the same notions of conditions, compatibility, and the same forcing relation using this Q .

As in Proposition 11.2.6, we will assume that all the elements of B are of the form 2^{2^k} for some $k \in M$. We will define the set G using an infinite $\Delta_1(Q \oplus Y')$ decreasing sequence of conditions. To ensure that $Q \oplus Y' \geq_T (G \oplus Y)'$ and that A is $\hat{B}$ -block $(G \oplus Y)$ -immune with cost $n + 2$, we will also construct $(G \oplus Y)'$ and $\hat{B}$ in parallel.

For this, we will need to satisfy three kinds of requirements for every $k \in M$:

- $\mathcal{R}_k$: there exists some $\sigma' \in 2^k$ such that $(G \oplus Y)' \upharpoonright k = \sigma'$.
- $\mathcal{S}_k$: there exists some $\ell > 2^{2^k}$ in B , such that, for every $x < \log_2(\log_2(\ell))$, if $\text{card } W_x^{G \oplus Y} \geq 2_n(\ell)$, then there is some $e < \ell$ such that $\text{card } W_e^Y \geq 2_n(\ell)$ and $W_e^Y \subseteq W_x^{G \oplus Y}$.
- $\mathcal{T}_k$: for every $e < k$, $G \subseteq^* R_e$ or $G \subseteq^* \mathbb{N} \setminus R_e$ and there exists some $x > k$ such that $x \in G$.

By Lemma 10.2.9, for every condition (σ, ρ) and every $k \in M$, there exists an extension forcing $\mathcal{R}_k$. And similarly for the $\mathcal{T}_k$ by Lemma 10.2.8. Remains the $\mathcal{S}_k$.

Lemma 11.2.9. *For every condition (σ, ρ) and every $k \in M$, there exists some extension $(\tau, \rho) \leq (\sigma, \rho)$ such that $(\tau, \rho) \Vdash \mathcal{S}_k$.*

Proof. For every finite set $F \subset M$, every $x \in F$ and every $t \in M$, let $W_{e(\langle F, x, t \rangle)}^Y$ search the smallest τ compatible with (σ, ρ) such that $\text{card } W_{x'}^{\tau \oplus Y} \geq 2_{n+2}(t)$ for every $x' \in F$ and output the elements of $W_x^{\tau \oplus Y}$.

Let $S : M \rightarrow M$ be the function which sends t to the largest value taken by $e(\langle F, x, t \rangle)$ for every $F \subseteq \{0, \dots, t-1\}$ and $x < t$. Assuming a reasonable encoding of e , this function is eventually dominated by $t \mapsto 2^{2^t}$, indeed, every $F \subseteq t$ and every $x \in t$ can be encoded using less than αt bits for some $\alpha \in \omega$ and the same holds for the corresponding value taken by e .

Hence, there exists some $\ell \in B$ such that $\ell > 2^{2^k}$ and such that $\ell' := \log_2(\log_2(\ell))$ satisfies $S(\ell') < 2^{2^{\ell'}} = \ell$.

Let $\mathcal{F} = \{F \subseteq \ell' : (\forall x \in F) \text{card } W_{e(\langle F, x, \ell' \rangle)}^Y \geq 2_{n+2}(\ell')\}$. The set $\mathcal{F}$ is $\Sigma_1(Y)$, non-empty (contains $\emptyset$) and bounded, hence, by $|\Sigma_1^0|$ we can pick some $F \in \mathcal{F}$ maximal for the inclusion.

Since $F \in \mathcal{F}$, there exists some τ compatible with (σ, ρ) such that $\text{card } W_x^{\tau \oplus Y} \geq 2_{n+2}(\ell') = 2_n(\ell)$ for every $x \in F$. The condition (τ, ρ) thus forces $\text{card } W_x^{G \oplus Y} \geq 2_n(\ell)$ for every $x \in F$ and, by the maximality assumption, forces $\text{card } W_x^{G \oplus Y} < 2_n(\ell)$ for every $x \in \{0, \dots, \ell' - 1\} \setminus F$, as otherwise $F \cup \{x\}$ would be in $\mathcal{F}$. Moreover, for every $x \in F$, we have $\text{card } W_{e(\langle F, x, \ell' \rangle)}^Y \geq 2_n(\ell)$ and $W_{e(\langle F, x, \ell' \rangle)}^Y = W_x^{\tau \oplus Y} \subseteq W_x^{G \oplus Y}$.

Thus, $(\tau, \rho) \Vdash \mathcal{S}_k$. $\square$

Construction: We $Q \oplus Y'$ -computably construct a decreasing sequence (σ_s, ρ_s) of conditions. In parallel, the sets $(G \oplus Y)'$ and $\hat{B}$ will be approached by sequences (σ'_s) and (B_s) of finite binary sequences.

Let $\sigma_0 = \rho_0 = \sigma'_0 = B_0 = \emptyset$. Assume that $\sigma_s, \rho_s, \sigma'_s$ and B_s have been defined for some $s \in M$ and let

$$k_s = 1 + \max\{|\sigma_s|, |\rho_s|, |\sigma'_s|, |B_s|\}$$

Using Lemma 10.2.9, Lemma 11.2.9 and Lemma 10.2.8, there exists some extension $(\tau, \rho') \leq (\sigma_s, \rho_s)$ forcing $\mathcal{R}_{k_s}, \mathcal{S}_{k_s}$ and $\mathcal{T}_{k_s}$. Furthermore, such an extension can be found $(Q \oplus Y')$ -effectively. We then let $\sigma_{s+1} = \tau$, $\rho_{s+1} = \rho'$, $\sigma'_{s+1} = \sigma'$ where σ' is the sequence obtained from the fact that $\mathcal{R}_{k_s}$ is forced and let $B_{s+1}(x) = 0$ for every $x \in [|B_s|, \log_2(\log_2(\ell))]$ and $B_{s+1}(\log_2(\log_2(\ell))) = 1$ where ℓ is the bound obtained from the fact that $\mathcal{S}_{k_s}$ is forced. We then proceed to the next stage of the construction.

The $\sigma_s, \sigma'_s, \rho_s$ and B_s may only be defined on some cut I but are nonetheless of length unbounded in M .

Thus, this construction defines two sets $G = \bigcup_{s \in I} \sigma_s$ and $\hat{B} = \bigcup_{s \in I} B_s$ that are $(Q \oplus Y')$ -computable and such that $(G \oplus Y)'$ is also $(Q \oplus Y')$ -computable. We also get that $(G \oplus Y)' \geq Q$, indeed, using $(G \oplus Y)'$, we can check if $G \subseteq^* R_i$ or if $G \subseteq^* \mathbb{N} \setminus R_i$ for every $i \in M$ and thus find the value of $Q(i)$, hence $(G \oplus Y)' \equiv_T Q \oplus Y'$ and $\hat{B}$ is $\Delta_2(G \oplus Y)$. Since the requirements $\mathcal{T}_k$ are satisfied, the set G is M -unbounded and $\vec{R}$ -cohesive. Finally, for the same reasons as in Proposition 11.2.6, we get that A is $\hat{B}$ -block $(G \oplus Y)$ -immune with cost $n + 2$. $\square$

As in Proposition 11.2.6, the previous construction makes no reference to the set A . Thus, every set $\hat{A} \subseteq M$ that is B -block Y -immune with cost n will also be $\hat{B}$ -block $(G \oplus Y)$ -immune with cost $n + 2$.

Corollary 11.2.10. Let $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0$ be a countable model topped by some set Y . Let B be an M -unbounded $\Delta_2(Y)$ set and fix $n \in \omega$. Then, for every set $A \subseteq M$ that is B -block Y -immune with cost n and every instance $\vec{R}$ of COH in S , there exists some M -unbounded set $G \subseteq M$ and some M -unbounded $(G \oplus Y)'$ -computable set $\hat{B}$ such that:

- G is $\vec{R}$ -cohesive
- $\mathcal{M}[G] \models \text{RCA}_0 + \text{B}\Sigma_2^0$
- A is $\hat{B}$ -block $(G \oplus Y)$ -immune with cost $n + 2$.

Proof. By Theorem 9.2.5, there exists a set $P \subseteq M$ such that $P \gg Y'$ and $\mathcal{M}[P] \models \text{RCA}_0^*$. The conclusion then follows from Proposition 11.2.8. $\square$

11.2.4 Main result

We are now ready to prove Main Theorem 11.0.1.

Proposition 11.2.11. Every countable topped model of $\text{RCA}_0 + \text{B}\Sigma_2^0 + \neg \text{I}\Sigma_2^0$ can be ω -extended into a model of $\text{RCA}_0 + \text{ADS} + \neg \text{BRT}_2^2$.

Proof. Let $\mathcal{M}_0 = (M, S_0)$ be a countable model of $\text{RCA}_0 + \text{B}\Sigma_2^0 + \neg \text{I}\Sigma_2^0$, topped by a set $Z_0 \in S_0$. Let $I \subseteq M$ be a $\Sigma_2(Z_0)$ -definable proper cut, witnessing that $\mathcal{M}_0 \models \neg \text{I}\Sigma_2^0$, and let $B_0 \subseteq M$ be a $\Delta_2(\mathcal{M}_0)$ -definable M -unbounded set with cardinality I . By the proof of Main Theorem 11.0.2, $\mathcal{M}_0$ contains a stable instance $f : [M]^2 \rightarrow 2$ of BRT_2^2 with no $\Delta_1(Z_0)$ -definable solution, and such that the set $A = \{x \in M : \forall^\infty y f(x, y) = 0\}$ is B_0 -block Z_0 -immune with cost 1 by the same reasoning as in Proposition 11.2.5.

Since M is countable, there is a (not necessarily computable) reordering $(\Gamma_n)_{n \in \omega}$ of the sequence $(\Phi_x)_{x \in M}$ of all Turing functionals with codes in M . By Proposition 11.2.6 and Corollary 11.2.10, define a sequence $(\mathcal{M}_n)_{n \in \omega}$ of countable models of $\text{RCA}_0 + \text{B}\Sigma_2^0$ such that for every $n \in \omega$,

- (1) $\mathcal{M}_{n+1}$ is topped by some set Z_{n+1} and ω -extends $\mathcal{M}_n$;
- (2) A is B_n -block Z_n -immune with some cost $c_n \in \omega$;
- (3) if $n = \langle 0, a, b \rangle$ and $\Gamma_a^{Z_b}$ is an instance $\mathcal{L} = (M, \leq_{\mathcal{L}})$ of SADS , then $\mathcal{M}_{n+1}$ contains an M -unbounded $\mathcal{L}$ -ascending or $\mathcal{L}$ -descending sequence;
- (4) if $n = \langle 1, a, b \rangle$ and $\Gamma_a^{Z_b}$ is an instance $\vec{R}$ of COH , then $\mathcal{M}_{n+1}$ contains an M -unbounded $\vec{R}$ -cohesive set.

Let $\mathcal{N} = \bigcup_{n \in \omega} \mathcal{M}_n$. Then $\mathcal{N} \models \text{RCA}_0 + \text{B}\Sigma_2^0$. By (1), $\mathcal{N}$ ω -extends $\mathcal{M}_0$. By (2) A is immune relative to every set in $\mathcal{N}$, so f is an instance of BRT_2^2 with no solution in $\mathcal{N}$. It follows that $\mathcal{N} \not\models \text{BRT}_2^2$. By (3) and (4), $\mathcal{N} \models \text{SADS} + \text{COH}$, so by Hirschfeldt and Shore [63], $\mathcal{N} \models \text{ADS}$. $\square$

Main Theorem 11.0.1. $\text{RCA}_0 + \text{ADS} \not\models \text{BRT}_2^2$

Proof. Let M be a countable model of $\text{I}\Sigma_1 + \text{B}\Sigma_2 + \neg\text{I}\Sigma_2$. Let $\mathcal{M} = (M, S)$ be the model whose second-order part S consists of the Δ_1 -definable sets with parameters in M . By Friedman [48], $\mathcal{M} \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \neg\text{I}\Sigma_2^0$. Moreover, $\mathcal{M}$ is countable, and topped by $\emptyset$, so by Main Theorem 11.2.11, there is an ω -extension $\mathcal{N}$ of $\mathcal{M}$ such that $\mathcal{N} \models \text{RCA}_0 + \text{ADS} + \neg\text{BRT}_2^2$. Thus, $\text{RCA}_0 + \text{ADS} \not\models \text{BRT}_2^2$. $\square$

Corollary 11.2.12. $\text{RCA}_0 + \text{ADS} + \neg\text{BRT}_2^2$ is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0 + \neg\text{I}\Sigma_2^0$.

Proof. Let $\varphi = \forall X \psi(X)$ be a Π_1^1 -sentence such that $\text{RCA}_0 + \text{B}\Sigma_2^0 + \neg\text{I}\Sigma_2^0 \not\models \varphi$. By the completeness theorem and the downward Löwenheim-Skolem theorem, there exists a countable model $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \neg\text{I}\Sigma_2^0 + \neg\varphi$. Let $X \in S$ be such that $\mathcal{M} \models \neg\psi(X)$ and $Y \in S$ be such that $\text{I}\Sigma_2^0(Y)$ fails. Let $\mathcal{M}_1 = (M, \Delta_1\text{-Def}(M, X, Y))$. By Friedman [48], $\mathcal{M}_1 \models \text{RCA}_0 + \text{B}\Sigma_2^0$. Moreover, since ψ is arithmetic, $\mathcal{M}_1 \models \neg\psi(X)$, and since $Y \in \mathcal{M}_1$, $\mathcal{M}_1 \models \neg\text{I}\Sigma_2^0$. Lastly, $\mathcal{M}_1$ is topped by $X \oplus Y$, so one can apply Main Theorem 11.2.11 to get an ω -extension $\mathcal{N}$ of $\mathcal{M}_1$ such that $\mathcal{N} \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \text{ADS} + \neg\text{BRT}_2^2$. In particular, $X \in \mathcal{N}$ and $\mathcal{N} \models \neg\psi(X)$, so $\text{RCA}_0 + \text{B}\Sigma_2^0 + \text{ADS} + \neg\text{BRT}_2^2 \not\models \varphi$. $\square$

Part IV

Partial conservation

Chapter 12

Cuts

In order to remove the use of the infinite in a proof of a finitary result, one idea could be to analyze the proof and replace each use of an infinitary statement by a finite cardinality version of it for a sufficiently large value. Unfortunately, this idea does not work in the general case. One reason is that there are (computable) properties of infinite sets that cannot be reduced simply to the fact that every finite set is of cardinality at least k for every $k \in \mathbb{N}$.

One example of such a property is ω -largeness. A finite set F is said to be ω -large if $\text{card } F > \min F$. It is clear that every infinite subset of $\mathbb{N}$ has an ω -large subset, but there exist finite sets of arbitrarily large cardinalities that are not ω -large (e.g., $\{k, k+1, \dots, 2k-1\}$ for every k). Thus, a finitary proof involving an infinite set and using the fact that it has an ω -large subset cannot directly be replaced by a proof invoking a finite set of size k for some sufficiently large k .

Paris and Harrington [108] were the first to introduce the notion of ω -large set, in order to produce a variation of the finite Ramsey theorem that is true, but not provable in Peano arithmetic:

Definition 12.0.1 (Paris-Harrington principle). Let PH be the following Π_2 -statement:

“For every $n, k, m \in \mathbb{N}$, there exists some $N \in \mathbb{N}$ such that, for every $f : [N]^n \rightarrow k$, there exists some ω -large and f -homogeneous subset $F \subseteq \{0, \dots, N-1\}$ with $\text{card } F \geq m$.”

Theorem 12.0.2 (Paris-Harrington [108]). $\text{PA} \not\vdash \text{PH}$

Nevertheless, PH is a true statement that can be deduced from the infinite Ramsey theorem via a compactness argument.

Proposition 12.0.3 (Paris-Harrington [108]). $\text{WKL}_0 + \text{RT} \vdash \text{PH}$

Proof. We can apply the same compactness argument as in the proof of the finite Ramsey theorem (Theorem 6.0.3). However, instead of considering the tree of all finite colorings for which no homogeneous subset of cardinality m exists, we consider the tree of colorings for which there is no homogeneous subset of cardinality at least m that is also ω -large.

We then conclude by using the fact that every infinite set contains a subset of cardinality greater than m which is ω -large. $\square$

In particular, $\text{RCA}_0 + \text{RT}$ is not Π_2 -conservative over PA (it is in fact not even Π_1 -conservative, as the original proof of Paris and Harrington showed that $\text{PA} + \text{PH} \vdash \text{Con}(\text{PA})$).

In this part, we will go in the opposite direction. Rather than constructing finite combinatorial statements that follow from their infinitary version, but are not provable by finitary means, we shall investigate the first-order consequences of statements from

Ramsey theory, such as Ramsey's theorem for pairs or the ordered variable word theorem. For this, we shall prove, in RCA_0 or in $\text{RCA}_0 + \text{B}\Sigma_2^0$, finite combinatorial theorems that capture some of these first-order consequences.

In this chapter, we provide a general overview of the techniques used to obtain partial conservation results over theories such as PRA , RCA_0 and $\text{RCA}_0 + \text{B}\Sigma_2^0$. We begin in Section 12.1 by introducing the notion of semi-regular cut, due to Kirby and Paris [73], which can be used to prove Π_2 -conservation results over PRA . Next, in Section 12.2, we present a generalization by Ketonen and Solovay [72] of the notion of ω -large set to other countable ordinals (called α -largeness) that was then used by Patey and Yokoyama [115] to prove $\forall\Pi_3^0$ -conservation results over RCA_0 . Finally, in Section 12.3, we introduce a parameterized version of α -largeness and see how to use it to obtain $\forall\Pi_4^0$ -conservation results above $\text{RCA}_0 + \text{B}\Sigma_2^0$. Here, a $\forall\Gamma$ -formula is any formula of the form $\forall X\varphi(X)$ where $\varphi \in \Gamma$.

12.1 Cuts and Π_2 -conservation

Contrary to the previous parts where the new models built were ω -extensions of the initial model (either in non-implication proofs over ω -models, or in the Π_1^1 -conservation proofs), here we will modify the first-order part of our models by removing some elements from it. The first-order part of the resulting model will form a cut, whose definition we now recall:

Definition 12.1.1 (Cut). A *cut* in a model M of first-order arithmetic is a nonempty subset $I \subseteq M$ which is closed under successor, and is an initial segment of M , that is, if $a \in I$ and $b \leq a$, then $b \in I$.

A cut $I \subseteq M$ is *proper* if $I \neq M$.

As in Part III, it becomes even more important to distinguish between the formal set $\mathbb{N}$, used in formulas to denote the first-order part of the model under consideration, and the set ω of standard integers. An additional layer of complexity arises from the fact that we are also dealing with ordinals inside non-standard models. To address this, we use the notations ω and ϵ_0 within formulas to denote the ordinals ω and ϵ_0 as they are interpreted in a given model.

Once a cut I is defined, the second-order part of the resulting model will be given by the sets coded by M .

Definition 12.1.2 (Coded sets). For $I \subseteq M$ a proper cut, a set $X \subseteq I$ is said to be *M -coded on I* if $X = I \cap X'$ for some M -finite set X' . Let $\text{Cod}(M/I)$ be the set of all the sets that are M -coded on I .

If some proper cut I of a first-order structure $(M, 0_M, \text{succ}_M, +_M, \times_M, <_M)$ is closed under addition and multiplication, then

$$(I, \text{Cod}(M/I), 0_M, \text{succ}_M \upharpoonright I, +_M \upharpoonright I^2, \times_M \upharpoonright I^2, <_M \upharpoonright I^2)$$

(which we will simply write $(I, \text{Cod}(M/I))$), is also a first-order structure.

Inside this structure, the infinite sets are of the form $X \cap I$ for some M -finite set X that is unbounded in I . Similarly, for $b > I$, and $n, k \in \omega$, the colorings $f : [I]^n \rightarrow k$ are those of the form $g \upharpoonright [I]^n$ for some M -finite coloring $g : [b]^n \rightarrow k$.

For conservation proofs, our goal will be to build suitable cuts I such that, when starting from a model $\mathcal{M} = (M, S)$ of $\text{RCA}_0 + \psi$ for some first-order sentence ψ , the structure $(I, \text{Cod}(M/I))$ remains a model of $\text{RCA}_0 + \psi$ plus some other axioms such as RT_2^2 . In contrast to Part III where the second-order part of both models was the same, here some care is required for $(I, \text{Cod}(M/I))$ to remain a model of ψ . In fact, our constructions only work for formulas up to Π_4^0 in the arithmetical hierarchy.

As in Part III, we can also treat the case where ψ contains second-order parameters from $\mathcal{M}$. Consequently, in the same way that we obtained Π_1^1 -conservation results in Part III instead of arithmetical conservation, here we will obtain $\forall\Pi_4^0$ -conservation

theorems. However, when ψ contains second-order parameters, we cannot directly claim that $(I, \text{Cod}(M/I)) \models \psi$, since these second-order parameters are not elements of $\text{Cod}(M/I)$. Instead, we obtain that $(I, \text{Cod}(M/I)) \models \hat{\psi}$, where $\hat{\psi}$ is the formula obtained from ψ by replacing each second-order parameter with its restriction to I , which will be sufficient for our purpose. For simplicity, we will often identify ψ with the corresponding $\hat{\psi}$ and omit the distinction in proofs.

To do so, not every cut will do. Indeed, as IS_1^0 proves the totality of every primitive recursive function, the cut I also has to be closed under all primitive recursive functions for $(I, \text{Cod}(M/I))$ to be a model of IS_1^0 . The notion of semi-regular cut encapsulates the properties needed to build a model of WKL_0 .

Definition 12.1.3 (Semi-regular cut [73]). A proper cut $I \subseteq M$ is said to be *semi-regular* if for every M -finite set $E \subseteq M$ such that $\text{card } E \in I$, $E \cap I$ is bounded in I (i.e., there exists $a \in I$ such that $E \cap I < a$).

Proposition 12.1.4 (Scott [123] and Kirby and Paris [73, Proposition 1]. See also Simpson [128, Lemma IX.3.11]). *Let $M \models \text{PRA}$ and let $I \subseteq M$ be a proper cut. Then, I is semi-regular if and only if $(I, \text{Cod}(M/I)) \models \text{WKL}_0$.*

The Π_2 -conservation of WKL_0 over PRA can then be derived by model-theoretic means, providing an arguably simpler proof, albeit a non-constructive one.

Theorem 12.1.5 (See Simpson [128, Theorem IX.3.16]). *WKL_0 is Π_2 -conservative over PRA .*

The model-theoretic proof of that result proceeds as follows: if $\text{PRA} \not\models \forall x \exists y \theta(x, y)$ for some Σ_0 -formula θ , then there exists a countable non-standard model M of $\text{PRA} + \exists x \forall y \neg \theta(x, y)$. Letting $a \in M$ be such that $M \models \forall y \neg \theta(a, y)$, we can construct some semi-regular cut I containing a . Thus, the resulting model $(I, \text{Cod}(M/I))$ will be a model of $\text{WKL}_0 + \forall y \neg \theta(a, y)$, thereby proving that $\text{WKL}_0 \not\models \forall x \exists y \theta(x, y)$.

For $(I, \text{Cod}(M/I))$ to be a model of RT_2^2 , the semi-regularity on the cut I is not sufficient, stronger assumptions are required. Indeed, any instance of RT_2^2 in $(I, \text{Cod}(M/I))$ has the form $f \upharpoonright [I]^2$ for $f : [c]^2 \rightarrow 2$ an M -finite coloring with $c > I$. In order for $(I, \text{Cod}(M/I))$ to contain a solution to this instance, we must find an f -homogeneous set $X \subseteq \{0, \dots, c-1\}$ such that $X \cap I$ is unbounded in I .

However, a direct application of the finite Ramsey theorem for pairs and two colors only guarantees the existence of a homogeneous set $X \subseteq \{0, \dots, c-1\}$ of size at least $\log_4(c)$ (using the bound $R_2(n) \leq 4^n$ of Erdős and Szekeres [42]). This does not ensure that $X \cap I$ is unbounded in I for some semi-regular cut I . Indeed, suppose X is of the form $\{c - \log_4(c), \dots, c-1\}$ with $I < c$. If $d \in X \cap I$ for some d , then by closure of I under multiplication, we would have $2d \in I$, but $2d > c$, contradiction.

In the next section, we will introduce α -large sets and study finite variants of RT -like statements that will put more weight on the smallest integers and guarantee that the resulting sets X contain sufficiently many ‘small’ elements so that a semi-regular cut I in which X is cofinal can be found.

12.2 Largeness and $\forall\Pi_3^0$ -conservation

12.2.1 Largeness notions and α -large sets

A subset $X \subseteq \mathbb{N}$ being infinite guarantees that we can find, for every $k \in \mathbb{N}$, finite subsets of X of size k . But, as shown before with the example of ω -largeness, these are not the only examples of Σ_1^0 properties that hold true for every infinite set.

Definition 12.2.1. A family $\mathcal{L}$ of finite subsets of $\mathbb{N}$ is a *largeness notion* if any infinite set has a subset in $\mathcal{L}$ and $\mathcal{L}$ is closed under supersets (i.e., if $F \in \mathcal{L}$ and $G \subseteq_{\text{fin}} \mathbb{N}$ is such that $F \subseteq G$, then $G \in \mathcal{L}$).

Any finite set in $\mathcal{L}$ is said to be *$\mathcal{L}$ -large*.

A particular family of largeness notions that will be of interest to us is the α -largeness, for some ordinal α . It is possible to define α -large sets for very large countable ordinals as long as we have a system of ordinal notations at our disposal. For our purpose, it will be sufficient to define α -large sets only for α of the form $\omega^n \cdot k$ for some $n, k \in \mathbb{N}$.

Definition 12.2.2 (α -largeness). A set $F \subseteq_{\text{fin}} \mathbb{N}$ is

- (1) ω^0 -large iff $F \neq \emptyset$
- (2) $\omega^{(n+1)}$ -large iff $F \setminus \{\min F\}$ is $(\omega^n \cdot \min F)$ -large
- (3) $\omega^n \cdot k$ -large iff there are k ω^n -large subsets of F

$$F_0 < F_1 < \dots < F_{k-1}$$

For every $n \in \omega$, RCA_0 is powerful enough to prove that ω^n -largeness is a largeness notion.

Proposition 12.2.3. *For every $n \in \omega$, RCA_0 proves that every infinite set has an ω^n -large subset.*

However, RCA_0 is not strong enough to prove that for every $n \in \mathbb{N}$, ω^n -largeness is a largeness notion. If that were the case, then RCA_0 would prove that ω^ω -largeness is a largeness notion, where a set X is ω^ω -large if $X \setminus \{\min X\}$ is $\omega^{\min X}$ -large. And, by the following result, this is not the case.

Theorem 12.2.4 (Generalized Parsons theorem for WKL_0 [115, Corollary 3.5]). *Let $\mathcal{L}$ be a Σ_1 notion of largeness, provably in WKL_0 . Then there exists some $n \in \omega$ such that IS_1^0 proves that every ω^n -large set is in $\mathcal{L}$.*

This provides further justification for using ω^n -largeness in the study of infinite sets within RCA_0 . For sufficiently large n , ω^n -large sets become increasingly difficult to distinguish from genuinely infinite sets in a model of WKL_0 (and in particular, in models of RCA_0).

We are interested in the closure of largeness notions under combinatorial statements coming from Ramsey theory (i.e., how large must a finite set X be, so that every instance of our problem on X has a solution that is $\mathcal{L}$ -large for some largeness notion $\mathcal{L}$).

Given a function $f : [\mathbb{N}]^n \rightarrow k$ and a finite set $G = \{x_0 < \dots < x_{\ell-1}\}$, let $f_G : [\ell]^n \rightarrow k$ be defined by $f_G(i_0, \dots, i_{n-1}) = f(x_{i_0}, \dots, x_{i_{n-1}})$. Most statements we are interested in will be Π_2^1 problems of the following form:

Definition 12.2.5. Given $n, k \in \mathbb{N}$, an *RT-like statement* is a Π_2^1 -formula of the form

$$(\forall f : [\mathbb{N}]^n \rightarrow k)(\exists Y)[\Phi(f, \mathbb{N}) \rightarrow (Y \text{ is infinite} \wedge \Psi(f, Y))]$$

where $\Phi(f, Y)$ and $\Psi(f, Y)$ are of the form $(\forall G \subseteq_{\text{fin}} Y)\Theta(f_G)$ with Θ a Σ_0^0 -formula.

For example, RT_2^2 , ADS and EM are all RT-like statements.

Definition 12.2.6. Let P be an RT-like statement with witnesses n, k, Φ and Ψ , and let $\mathcal{L}$ be a largeness notion. A finite set F is *P - $\mathcal{L}$ -large* if for every coloring $f : [F]^n \rightarrow k$ such that $\Phi(f, F)$ holds, there is an $\mathcal{L}$ -large subset $E \subseteq F$ such that $\Psi(f, E)$ holds.

By extension, if $(P_k)_{k \in \mathbb{N}}$ is a family of RT-like statements, then we say that a finite set F is *P - $\mathcal{L}$ -large* if it is P_k - $\mathcal{L}$ -large for every $k \leq \min F$. For instance, a set F is RT^1 - $\mathcal{L}$ -large if for every $k \leq \min F$ and every coloring $f : \mathbb{N} \rightarrow k$, there is an $\mathcal{L}$ -large f -homogeneous subset of F . Equivalently, since $\mathcal{L}$ is closed under supersets, F is RT^1 - $\mathcal{L}$ -large if for every $k \leq \min F$ and every k -partition $E_0 \sqcup \dots \sqcup E_{k-1} = F$, there is some $i < k$ such that E_i is $\mathcal{L}$ -large.

As an example, we present the following bounds for RT_k^1 and RT^1 , which will be proven in Chapter 13.

Proposition 12.2.7. *RCA_0 proves that, for every $n \geq 0$ and $a, k \geq 1$:*

- (1) *If X is $\omega^n \cdot ak$ -large, then it is RT_k^1 - $\omega^n \cdot a$ -large.*
- (2) *If X is ω^{n+1} -large, then it is RT^1 - ω^n -large.*

With an argument similar to that of Proposition 12.0.3, we can see that $\text{WKL}_0 + \text{P}$ proves that if $\mathcal{L}$ is a largeness notion, then so is $\text{P-}\mathcal{L}$ -largeness. As we will see with Theorem 12.2.9, the ability for RCA_0 itself to prove this statement gives a good criterion for the $\forall\Pi_3^0$ -conservativity of P over RCA_0 .

In a theory able to prove that $\text{P-}\mathcal{L}$ -largeness is a largeness notion for every largeness notion $\mathcal{L}$, then for every $m \in \omega$ we can prove the existence of so-called m -dense(P) sets (defined below). In contrast to $\text{P-}\alpha$ -largeness, which concerns how large a solution to a single instance of P can be, m -density(P) is a notion of largeness which involves m successive applications of P . And we will see in Theorem 12.2.9 how the existence of a c -dense(P, RT^1) set for some non-standard element $c \in M$ can then be used to build the desired cut.

Definition 12.2.8 (m -density). Let $\text{P}_0, \dots, \text{P}_{i-1}$ be RT -like statements and let $m \in \mathbb{N}$. A finite set F is m -dense($\text{P}_0, \dots, \text{P}_{i-1}$) if either:

- $m = 0$ and F is ω -large;
- $m > 0$ and the following two conditions are satisfied:
 1. for every $j < i$, every instance of P_j on F admits a solution that is $(m-1)$ -dense($\text{P}_0, \dots, \text{P}_{i-1}$);
 2. there exist subsets $F_0 < \dots < F_{\min F-1} \subseteq F \setminus \{\min F\}$ that are $(m-1)$ -dense($\text{P}_0, \dots, \text{P}_{i-1}$).

12.2.2 $\forall\Pi_3^0$ -conservation of RT -like statements

We now have sufficient tools to establish a general proof scheme for $\forall\Pi_3^0$ -conservation over RCA_0 .

Theorem 12.2.9 (Patey and Yokoyama [115, Theorem 3.1]). *Let P be an RT -like statement with witnesses n, k, Φ and Ψ . If for every $\ell \in \omega$ there exists some $\ell' \in \omega$ such that RCA_0 proves that every $\omega^{\ell'}$ -large set is $\text{P-}\omega^\ell$ -large, then $\text{WKL}_0 + \text{P}$ is $\forall\Pi_3^0$ -conservative over RCA_0 .*

Proof. Let $\varphi := \forall X \forall x \exists y \forall z \theta(X \upharpoonright z, x, y, z)$ be a $\forall\Pi_3^0$ -sentence in Kleene normal form and assume that $\text{RCA}_0 \not\vdash \varphi$.

By compactness, completeness, and the Löwenheim-Skolem theorem, there exists some countable non-standard model $\mathcal{M} = (M, S) \models \text{RCA}_0 + \neg\varphi$. Let $a \in M$ and $A \in S$ be such that $\mathcal{M} \models \text{RCA}_0 + \forall y \exists z \neg\theta(A \upharpoonright z, a, y, z)$.

Consider the following set $Y = \{y_0, y_1, \dots\}$ defined inductively by: $y_0 = a$ and, for every $i \in M$, y_{i+1} is the smallest element $> y_i$ such that $\forall y < y_{i+1} \exists z < y_{i+1} \neg\theta(A \upharpoonright z, a, y, z)$. Since $\mathcal{M} \models \text{BS}_1^0$ and $\mathcal{M} \models \forall y \exists z \neg\theta(A \upharpoonright z, a, y, z)$, Y is infinite, and since Y is Δ_1^0 , we have $Y \in S$.

By external induction on $m \in \omega$, we can prove that there exists some bound ℓ_m such that RCA_0 proves that every ω^{ℓ_m} -large set is m -dense(P, RT^1). Indeed, for $m = 0$, simply take $\ell_0 = 1$. And, assuming that there exists such an ℓ_m for some $m \in \omega$, let ℓ'_m be such that RCA_0 proves that every $\omega^{\ell'_m}$ -large set is $\text{P-}\omega^{\ell_m}$ -large and pick $\ell_{m+1} = \max\{\ell'_m + 1, \ell_m + 1\} = \ell'_m + 1$, by Proposition 12.2.7, RCA_0 proves that every $\omega^{\ell_{m+1}}$ -large set is $(m+1)$ -dense(P, RT^1).

The property of being m -dense(P, RT^1) is Σ_1^0 and $\mathcal{M} \models I\Sigma_1^0$, hence, by overspill (see Proposition 9.1.5), there must exist some non-standard element $c \in M$ and some c -dense(P, RT^1) subset F of Y . We can also assume that $\min F \geq 2$.

Fix $(E_i)_{i \in \omega}$ an enumeration containing all the M -finite sets infinitely many times, and fix $(f_i)_{i \in \omega}$ an enumeration of all the M -finite functions from $[0, \dots, \max F]^n$ to k .

Consider the following decreasing sequence $F = F_{-1} \supseteq F_0 \supseteq F_1 \supseteq \dots$ of M -finite sets, such that, for every $i \in \omega$:

- (1) F_i is $(c - 2i - 2)$ -dense(P, RT^1).
- (2) $\Phi(f_i, \{0, \dots, \min F_{2i}\})$ doesn't hold, or $\Psi(f_i, F_{2i})$ holds.
- (3) If $\text{card } E_i < \min F_{2i+1}$ then $[\min F_{2i+2}, \max F_{2i+2}) \cap E_i = \emptyset$.

Assume F_i has already been defined for some $i \in \omega \cup \{-1\}$.

If $i + 1 = 2j$ for some $j \in \omega$, then, since F_i is $(c - 2i - 2)$ -dense(P, RT^1) and $\min F_i \geq 2$, there exist $F_i^0 < F_i^1 \subseteq F_i$ that are $(c - 2i - 3)$ -dense(P, RT^1). If $\Phi(f_j, \{0, \dots, \max F_i^0\})$ holds, let $F_{i+1} \subseteq F_i^0$ be some $(c - 2i - 4)$ -dense(P, RT^1) set such that $\Psi(f_j, F_{i+1})$ holds. Otherwise, let $F_{i+1} = F_i^1$. Since $\min F_i^1 > \max F_i^0$, this ensures that $\Phi(f_j, \{0, \dots, \min F_{i+1}\})$ does not hold.

If $i + 1 = 2j + 1$ for some $j \in \omega$: if $\min F_i \leq \text{card } E_j$ then let $F_{i+1} = F_i$ otherwise, consider the coloring $g : F_i \rightarrow \min F_i$ defined by $g(x) = \text{card}\{y \in E_j : y < x\}$ and let F_{i+1} be some g -homogeneous and $(c - 2i - 3)$ -dense(P, RT^1) subset of F_i , in that case we have $[\min F_{i+1}, \max F_{i+1}) \cap E_j = \emptyset$ as $g(\min F_{i+1}) = g(\max F_{i+1})$ implies that $\text{card}\{y \in E_j : \min F_{i+1} \leq y < \max F_{i+1}\} = 0$.

By (1), this construction can be carried out for every $i \in \omega$. Let $I = \sup\{\min(F_i) : i \in \omega\} := \{x \in M : \exists i \in \omega, x < \min F_i\}$. Notice that, as the F_i form a decreasing sequence of sets, every F_i is cofinal in I .

I forms a semi-regular cut of M . Indeed, consider some M -finite set $E \subseteq M$ with $\text{card } E \in I$. By definition of I , there exists some $j \in \omega$ such that $\text{card } E < \min F_j$. Then, since $(E_i)_{i \in \omega}$ contains all the M -finite sets infinitely many times, there exists some $i \geq j$ such that $E = E_i$, and, as $F_{2i+1} \subseteq F_j$ (since $2i + 1 \geq j$), we also have $\text{card } E < \min F_{2i+1}$. By (3), this implies that $[\min F_{2i+2}, \max F_{2i+2}) \cap E_i = \emptyset$, and hence that $E \cap I$ is bounded in I (as $\min F_{2i+2} \in I$ and $\max F_{2i+2} > I$ by construction). By Proposition 12.1.4, the semi-regularity of I implies $(I, \text{Cod}(M/I)) \models \text{WKL}_0$.

For every $i \in M$ such that $y_i \in Y \cap I$, we have that $y_{i+1} \in Y \cap I$. Hence, $\forall y \in I \exists z \in I \neg \theta(A \upharpoonright z, a, y, z)$ holds. Since $\min Y = a$, we have that $a \in I$, and letting $A' = A \cap I$ denote the M -coded set by A on I , it follows that $(I, \text{Cod}(M/I)) \models \forall y \exists z \neg \theta(A' \upharpoonright z, a, y, z)$, and therefore $(I, \text{Cod}(M/I)) \models \neg \varphi$.

Finally, for every instance $f : [I]^n \rightarrow k$ of P in $(I, \text{Cod}(M/I))$, there exists some index $i \in \omega$ such that $f = f_i \upharpoonright [I]^n$. If $\Phi(f_i, \{0, \dots, \min F_{2i}\})$ doesn't hold, then neither does $\Phi(f_i, I)$ as $\min F_{2i} \in I$, contradicting our assumption that f is an instance of P in $(I, \text{Cod}(M/I))$. Thus, by (2) we have that $\Psi(f_i, F_{2i})$ holds, and hence $\Psi(f, F_{2i} \cap I)$ holds, and, as F_{2i} is cofinal in I , the set $F_{2i} \cap I$ is infinite in $(I, \text{Cod}(M/I))$, thus $(I, \text{Cod}(M/I)) \models P$.

Thus, $(I, \text{Cod}(M/I)) \models \text{WKL}_0 + P + \neg \varphi$, and $\text{WKL}_0 + P \not\models \varphi$. $\square$

The original proof by Patey and Yokoyama of the $\forall \Pi_3^0$ -conservation of RT_2^2 over RCA_0 [115] did not provide explicit bounds for the amount of $\omega^{\ell'}$ -largeness required for a set to be RT_2^2 - ω^{ℓ} -large. Such bounds were later obtained by Kołodziejczyk and Yokoyama [79], who showed that every $\omega^{144(n+1)}$ -large set is RT_2^2 - ω^n -large.

Deriving explicit bounds can be of theoretical interest. Indeed, the fact that the function $n \mapsto 144(n + 1)$ is polynomial in n was exploited by Kołodziejczyk et al. [78]. Adapting the general scheme of forcing interpretation as introduced by Avigad [4], they proved that every proof of a $\forall \Pi_3^0$ -statement in $\text{RCA}_0 + RT_2^2$ could be translated to a proof of that same statement in RCA_0 with at most a polynomial increase in size.

12.3 Largeness(θ) and $\forall\Pi_4^0$ -conservation

To move from a proof of $\forall\Pi_3^0$ -conservation to one of $\forall\Pi_4^0$ -conservation, several adjustments are required. When, starting with a model M that satisfies an $\exists\Sigma_4^0$ -formula, or equivalently, after fixing witnesses for the existential quantifiers, a Π_3^0 -formula, we aim to construct a cut that preserves this property. The main difficulty is that for $\forall x\exists y\forall z\theta(x, y, z)$ a Π_3^0 -formula, there is no guarantee that there exists an infinite set $Y = \{y_0, y_1, \dots\}$ such that $M \models \forall x < y_i \exists y < y_{i+1} \forall z \theta(x, y, z)$ for every $i \in \mathbb{N}$ as such a set would require Π_1^0 -comprehension to exist. Instead, we will modify our notion of α -largeness so that it now depends on the formula θ .

Another point to consider is that $\text{RCA}_0 + \text{RT}_2^2$ proves BS_2^0 (and the same holds for many other Ramsey-like statements). Since BS_2^0 is a $\forall\Pi_4^0$ -statement not provable in RCA_0 , it follows that we cannot prove the $\forall\Pi_4^0$ -conservativity of RT_2^2 over RCA_0 . Instead, we will establish $\forall\Pi_4^0$ -conservativity results over the stronger base theory $\text{RCA}_0 + \text{BS}_2^0$.

12.3.1 Definitions and combinatorial properties

Definition 12.3.1. Let $\theta(x, y, z)$ be a Σ_0^0 -formula and let $b \in \mathbb{N}$ be the largest first-order constant appearing in θ . Two finite sets $b \leq X < Y$ are θ -*apart* if

$$\forall x < \max X \exists y < \min Y \forall z < \max Y \theta(x, y, z)$$

Note that θ -apartness is a transitive relation. Moreover, if $X < Y$ are θ -apart and $X_0 \subseteq X$ and $Y_0 \subseteq Y$, then X_0, Y_0 are θ -apart.

The use of the largest first-order constant appearing in θ in the definition of θ -apartness and α -largeness(θ) is to ensure that the cut constructed will contain all first-order constants appearing in θ .

Definition 12.3.2 (α -largeness(θ)). Let $\theta(x, y, z)$ be a Σ_0^0 -formula and let $b \in \mathbb{N}$ be the largest first-order constant appearing in θ . A set $X \subseteq_{\text{fin}} \mathbb{N}$ with $b \leq \min X$ is

- ω^0 -*large*(θ) if $X \neq \emptyset$.
- $\omega^{(n+1)}$ -*large*(θ) if $X \setminus \{\min X\}$ is $(\omega^n \cdot \min X)$ -*large*(θ)
- $\omega^n \cdot k$ -*large*(θ) if there are k pairwise θ -apart ω^n -*large*(θ) subsets of X

$$X_0 < X_1 < \dots < X_{k-1}$$

Likewise, we can define what it means for a set to be P - α -*large*(θ) and m -*dense*($\theta, \text{P}_0, \dots, \text{P}_i$):

Definition 12.3.3 (m -density(θ)). Let $\theta(x, y, z)$ be a Σ_0^0 -formula and let $\text{P}_0, \dots, \text{P}_{i-1}$ be RT -like statements. A finite set F is m -*dense*($\theta, \text{P}_0, \dots, \text{P}_{i-1}$) if either:

- $m = 0$ and F is ω -*large*(θ);
- $m > 0$ and the following two conditions are satisfied:
 1. for every $j < i$, every instance of P_j on F admits a solution that is $(m-1)$ -*dense*($\theta, \text{P}_0, \dots, \text{P}_{i-1}$);
 2. there exists pairwise θ -apart subsets $F_0 < \dots < F_{\min F-1} \subseteq F \setminus \{\min F\}$ that are $(m-1)$ -*dense*($\theta, \text{P}_0, \dots, \text{P}_{i-1}$).

Clearly, every α -*large*(θ) set is also α -*large*. The main combinatorial difference between α -largeness(θ) and ordinary α -largeness is that α -*large*(θ) sets cannot easily be grouped together (if a set A is θ -apart from two sets B and C , then A is not necessarily θ -apart from $B \cup C$), giving us worse bounds for RT^1 -largeness(θ). In

what follows, we will often use the term bound to refer specifically to the exponent n when discussing how much ω^n -largeness(θ) is required for a given result. For instance, we shall say that P admits a polynomial bound if there is a polynomial Q such that every $\omega^{Q(n)}$ -large(θ) set is P - ω^n -large(θ).

The following statement, whose proof is deferred to Chapter 14, gives a $4n + 4$ bound for a set to be RT^1 - ω^n -large(θ). In contrast to the case of RT^1 - ω^n -largeness (see Proposition 12.2.7), it is not possible to obtain a bound of the form $n + O(1)$. Indeed, for every $n \in \mathbb{N}$, there is a set F and a formula θ_F depending on F such that F is ω^{2n-1} -large(θ_F) but not RT^1 - ω^n -large(θ_F) (see [88, Section 7]).

Lemma 12.3.4. $|\Sigma_1^0$ proves that, for every Σ_0^0 -formula $\theta(x, y, z)$ and every $b \in \mathbb{N}$, if $F \subseteq_{\text{fin}} \mathbb{N}$ is ω^{4b+4} -large(θ), then it is RT^1 - ω^b -large(θ).

Just as RCA_0 proved that ω^n -largeness is a largeness notion for every $n \in \omega$, we have that, for every $n \in \omega$, $RCA_0 + B\Sigma_2^0$ proves that ω^n -largeness(θ) is a largeness notion when $\forall x \exists y \forall z \theta(x, y, z)$ holds.

Lemma 12.3.5. $RCA_0 + B\Sigma_2^0$ proves that, if $\forall x \exists y \forall z \theta(x, y, z)$ holds, then, for every $b \in \mathbb{N}$, if ω^b -largeness(θ) is a largeness notion, then

1. For every $c \in \mathbb{N}$, $\omega^b \cdot c$ -largeness(θ) is a largeness notion ;
2. ω^{b+1} -largeness(θ) is a largeness notion.

Proof. Suppose that $\forall x \exists y \forall z \theta(x, y, z)$ holds and that ω^b -largeness(θ) is a largeness notion.

Fix an infinite set X . We prove by internal $\Sigma_1^0(X)$ -induction over $c \geq 1$, that X contains an $\omega^b \cdot c$ -large(θ) subset. The case $c = 1$ follows from the fact that ω^b -largeness(θ) is a largeness notion. Assume the case c holds. Let $F \subseteq X$ be an $\omega^b \cdot c$ -large(θ) set. Since $\forall x \exists y \forall z \theta(x, y, z)$ holds, then, by $B\Sigma_2^0$, there is some $d \in X$ such that

$$\forall x < \max F \exists y < d \forall z \theta(x, y, z)$$

Since $X \setminus [0, d]$ is infinite and ω^b -largeness(θ) is a largeness notion, there is an ω^b -large(θ) subset $G \subseteq X \setminus [0, d]$. Then $F \cup G$ is an $\omega^b \cdot (c + 1)$ -large(θ) subset of X . Since every infinite set X contains an $\omega^b \cdot c$ -large(θ) subset, then $\omega^b \cdot c$ -largeness(θ) is a largeness notion.

Suppose now that for every $c \geq 1$, $\omega^b \cdot c$ -largeness(θ) is a largeness notion. In particular, for every infinite set X , there is an $\omega^b \cdot (\min X)$ -large(θ) set $F \subseteq X \setminus \{\min X\}$, hence $\{\min X\} \cup F$ is an ω^{b+1} -large(θ) subset of X . Thus, ω^{b+1} -largeness(θ) is a largeness notion. $\square$

Proposition 12.3.6. For every $n \in \omega$, $RCA_0 + B\Sigma_2^0$ proves that, for every Σ_0^0 -formula $\theta(x, y, z)$ such that $\forall x \exists y \forall z \theta(x, y, z)$ holds, then for every $c \geq 1$, $\omega^n \cdot c$ -largeness(θ) is a largeness notion.

Proof. We prove by external induction over n that ω^n -largeness(θ) is a largeness notion. The statement of the proposition then follows from Lemma 12.3.5. For the base case, since every non-empty finite set $X > b$ is ω^0 -large(θ) for b the largest first-order constant appearing in θ , then ω^0 -largeness(θ) is a largeness notion. The induction case is the second item of Lemma 12.3.5. $\square$

12.3.2 $\forall \Pi_4^0$ -conservation of RT-like statements

Similarly, we have the following proof scheme for $\forall \Pi_4^0$ -conservation over $RCA_0 + B\Sigma_2^0$.

Theorem 12.3.7. Let P be an RT-like statement with witnesses n, k, Φ and Ψ . If for every $\ell \in \omega$ there exists some $\ell' \in \omega$ such that $RCA_0 + B\Sigma_2^0$ proves, for every Σ_0^0 -formula $\theta(x, y, z)$, that every $\omega^{\ell'}$ -large(θ) set is P - ω^ℓ -large(θ), then $WKL_0 + B\Sigma_2^0 + P$

is $\forall\Pi_4^0$ -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$.

Proof. Let $\varphi := \forall X \forall x \exists y \forall z \exists t \theta'(X \upharpoonright t, x, y, z, t)$ be a $\forall\Pi_4^0$ -formula in Kleene normal form and assume that $\text{RCA}_0 + \text{B}\Sigma_2^0 \not\models \varphi$.

By compactness, completeness, and the Löwenheim-Skolem theorem, there exists some countable non-standard model $\mathcal{M} = (M, S) \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \neg\varphi$. Let $a \in M$ and $A \in S$ be such that $\mathcal{M} \models \forall y \exists z \forall t \neg\theta'(A \upharpoonright t, a, y, z, t)$. Let $\theta(y, z, t)$ be the Σ_0^0 -formula $\neg\theta'(A \upharpoonright t, a, y, z, t)$.

Since $\mathcal{M} \models \text{RCA}_0 + \text{B}\Sigma_2^0 + \forall y \exists z \forall t \theta(y, z, t)$, then, by Proposition 12.3.6, there exist ω^ℓ -large(θ) subsets of M for every $\ell \in \omega$. Thus, by assumption, there exist $\text{P-}\omega^\ell$ -large(θ) subsets of M for every $\ell \in \omega$, and by Lemma 12.3.4, there exist RT^1 - ω^ℓ -large(θ) subsets of M for every $\ell \in \omega$. Thus, there exist m -dense($\theta, \text{P}, \text{RT}^1$) sets for every $m \in \omega$.

By overspill, (as the property of being m -dense($\theta, \text{P}, \text{RT}^1$) is Σ_1^0 and $\mathcal{M} \models \text{I}\Sigma_1^0$) there must exist some non-standard element $c \in M$ and some c -dense($\theta, \text{P}, \text{RT}^1$) set $F \subseteq_{\text{fin}} M$.

Fix $(E_i)_{i \in \omega}$ an enumeration containing all the M -finite sets infinitely many times, and fix $(f_i)_{i \in \omega}$ an enumeration of all the M -finite functions from $[0, \dots, \max F]^n$ to k .

Consider the following decreasing sequence $F = F_{-1} \supseteq F_0 \supseteq F_1 \supseteq \dots$ such that, for every $i \in \omega$:

- (1) F_i is $(c - 2i - 2)$ -dense($\theta, \text{P}, \text{RT}^1$).
- (2) $\Phi(f_i, \{0, \dots, \min F_{3i}\})$ doesn't hold, or $\Psi(f_i, F_{3i})$ holds.
- (3) If $\text{card } E_i < \min F_{3i}$ then $[\min F_{3i+1}, \max F_{3i+1}) \cap E_i = \emptyset$.
- (4) $\forall y < \min F_{3i+1} \exists z < \min F_{3i+2} \forall t < \max F_{3i+2} \theta(y, z, t)$ holds.

Assume F_i has already been defined for some $i \in \omega \cup \{-1\}$.

If $i + 1 = 3j$ for some $j \in \omega$, then, since F_i is $(c - 2i - 2)$ -dense($\theta, \text{P}, \text{RT}^1$) and $\min F_i \geq 2$, there exist $F_i^0 < F_i^1 \subseteq F_i$ that are $(c - 2i - 3)$ -dense($\theta, \text{P}, \text{RT}^1$). If $\Phi(f_j, \{0, \dots, \max F_i^0\})$ holds, let $F_{i+1} \subseteq F_i^0$ be some $(c - 2i - 4)$ -dense($\theta, \text{P}, \text{RT}^1$) set such that $\Psi(f_j, F_{i+1})$ holds. Otherwise, let $F_{i+1} = F_i^1$. Since $\min F_i^1 > \max F_i^0$, this ensures that $\Phi(f_j, \{0, \dots, \min F_{i+1}\})$ does not hold.

If $i + 1 = 3j + 1$ for some $j \in \omega$: if $\min F_i \leq \text{card } E_j$ then let $F_{i+1} = F_i$ otherwise, consider the coloring $g : F_i \rightarrow \min F_i$ defined by $g(x) = \text{card}\{y \in E_j : y < x\}$ and let F_{i+1} be some g -homogeneous and $(c - 2i - 3)$ -dense($\theta, \text{P}, \text{RT}^1$) subset of F_i , in that case we have $[\min F_{i+1}, \max F_{i+1}) \cap E_j = \emptyset$.

If $i + 1 = 3j + 2$ for some $j \in \omega$: as F_i is $(c - 2i - 2)$ -dense($\theta, \text{P}, \text{RT}^1$) and $\min F_i \geq 2$, let $F_i^0 < F_i^1 \subseteq F_i$ be $(c - 2i - 3)$ -dense($\theta, \text{P}, \text{RT}^1$) and θ -apart and let $F_{i+1} = F_i^1$. By the θ -apartness of F_i^0 and F_i^1 , we have $\forall y < \max F_i^0 \exists z < \min F_i^1 \forall t < \max F_i^1 \theta(y, z, t)$, and, as $\max F_i^0 > \min F_i$, we have that $\forall y < \min F_i^1 \exists z < \min F_{i+1} \forall t < \max F_{i+1} \theta(y, z, t)$ holds.

By assumption (1), this construction can be carried out for every $i \in \omega$. Let $I = \sup\{\min(F_i) : i \in \omega\}$. Then, as in the proof of Theorem 12.3.7, by (3), I forms a semi-regular cut of M , hence $(I, \text{Cod}(M/I)) \models \text{WKL}_0$ and as the F_i form a decreasing sequence, every F_i is cofinal in I . By definition of α -large(θ) sets, we have that $a < \min F$, hence $a \in I$. By (2), we have $(I, \text{Cod}(M/I)) \models \text{P}$.

Finally, let $y \in I$, by definition of I there exists some $i \in \omega$ such that $y < \min F_{3i+1}$. Then, as (4) holds, there exists some $z < \min F_{3i+2}$ such that $\theta(y, z, t)$ holds for every $t < \max F_{3i+2}$. As $z < \min F_{3i+2}$, we have $z \in I$, and as $\max F_{3i+2} > I$, we have that $\forall y \in I \exists z \in I \forall t \in I \neg\theta'(A \upharpoonright t, a, y, z, t)$ holds, hence, as $a \in I$, we have $(I, \text{Cod}(M/I)) \models \neg\varphi$.

Thus, $(I, \text{Cod}(M/I)) \models \text{WKL}_0 + \text{B}\Sigma_2^0 + \text{P} + \neg\varphi$, and $\text{WKL}_0 + \text{B}\Sigma_2^0 + \text{P} \not\models \varphi$. $\square$

Chapter 13

α -largeness

Before proving bounds for RT_2^2 -largeness(θ) in the next chapter, we shall first familiarize ourselves with the standard notion of α -largeness, as developed by Ketonen and Solovay [72]. To do so, we will first study the general definition of α -largeness for ordinals smaller than ϵ_0 and prove the following partition theorem. Here, $\oplus$ denotes the natural (or Hessenberg) sum over ordinals, as introduced in Chapter 10. Moreover, a set F is said to be *at most α -large* if $F = \emptyset$ or if $F \setminus \{\max F\}$ is not α -large.

Main Theorem 13.0.1. *Let $\beta, \gamma < \epsilon_0$ be two ordinals, and $B, C \subseteq \mathbb{N} \setminus \{0\}$ be at most β -large and γ -large respectively. Then $B \cup C$ is at most $(\beta \oplus \gamma)$ -large.*

This theorem has several consequences, including a generalization of Ketonen and Solovay [72, Lemma 4.6] and pigeonhole generalizations of Bigorajska and Kotlarski [8, Theorem 12] for α -largeness. For this, we introduce a Hardy-like hierarchy of fast-growing functions, and adapt and generalize the techniques of Bigorajska and Kotlarski.

Then, in the rest of this chapter, we shall focus on ordinals $\alpha < \omega^\omega$ and give a tight upper bound (up to additive constant) for RT_k^2 - ω^n -largeness.

Main Theorem 13.0.2. *Let $n, k \geq 1$. If $X \subseteq_{\text{fin}} \mathbb{N}$ is ω^{kn+4} -large and $\min X \geq 2$, then it is RT_k^2 - ω^n -large.*

In the case $k = 2$, we obtain ω^{2n+4} -largeness, improving over the $\omega^{144(n+1)}$ -largeness upper bound of Kołodziejczyk and Yokoyama [79]. This bound is tight, in the sense that Kotlarski et al. [82, Theorem 5.4] proved (in a slightly different formalism) that every RT_k^2 - ω^n -large set is ω^{kn-1} -large.

The proof of Main Theorem 13.0.2 goes through the computation of upper bounds of closure for the usual decomposition of RT_2^2 into $\text{ADS} + \text{EM}$, or rather to one of their generalization to k -colorings in the form of the transitive Ramsey theorem trRT_k^2 and the fallow Erdős-Moser theorem (fEM) due to Towsner and Yokoyama [139].

We prove in particular that trRT_k^2 is solely responsible for the lower bound of RT_k^2 - ω^n -largeness, in that every ω^{n+3} -large set X with $\min X \geq 7$ is fEM - ω^n -large (Corollary 13.2.14). This latter result improves the ω^{18n} -largeness upper bound of Towsner and Yokoyama for fEM - ω^n -largeness.

This chapter is of purely combinatorial nature, with no meta-mathematical considerations. Although improving the bounds for RT_2^2 - ω^n -largeness can have theoretical consequences (with most notably, the proof of Kołodziejczyk et al. [78] of the absence of speedup between $\text{RCA}_0 + \text{RT}_2^2$ and RCA_0 for $\forall \Pi_3^0$ -formulas), there are, to our knowledge, no further theoretical implications arising from refining the bound $144n + O(1)$ to $2n + O(1)$.

In general, the statement “ α -largeness is a largeness notion” is equivalent, over RCA_0 , to the statement $\text{WF}^0(\alpha)$, asserting the well-foundedness of α . Many

proofs in Section 13.1 rely on induction up to ϵ_0 and therefore cannot be formalized directly in RCA_0 . Nevertheless, the results in this chapter are still formalizable in RCA_0 , indeed, most of them are of the form “For every $\alpha < \epsilon_0$, every α -large set X ...” or similar. Such statements are usually provable in RCA_0 , as the existence of the α -large set is assumed rather than derived.

13.1 Largeness below ϵ_0

In this section, we study the general definition of α -largeness as introduced by Ketonen and Solovay in [72], which applies for every ordinal $\alpha < \epsilon_0$ and establish some of its combinatorial properties. By Ketonen and Solovay [72, Lemma 4.6], the two definitions of α -largeness coincide for α of the form $\omega^n \cdot k$.

Recall that every ordinal $\alpha < \epsilon_0$ admits a unique *Cantor normal form*

$$\alpha = \omega^{\alpha_s} \cdot a_s + \cdots + \omega^{\alpha_0} \cdot a_0$$

for some $\alpha_0 < \alpha_1 < \cdots < \alpha_s < \alpha$ and $a_0, \dots, a_s \in \mathbb{N} \setminus \{0\}$. Given two ordinals $\alpha, \beta < \epsilon_0$, we write $\alpha \ll \beta$ ¹ if either α or β equals 0, or if the smallest exponent of β is at least the largest exponent of α in their Cantor normal form. If $\alpha \neq 0$, we shall also use its *short Cantor normal form* $\alpha = \beta + \omega^\gamma$ with $\beta \gg \omega^\gamma$.

Definition 13.1.1. Given a non-zero ordinal $\alpha < \epsilon_0$ with short Cantor normal form $\alpha = \beta + \omega^\gamma$, and $x \in \mathbb{N}$, let

$$\{\alpha\}(x) = \begin{cases} \beta & \text{if } \gamma = 0, \\ \beta + \omega^{\gamma-1} \cdot x & \text{if } \gamma \text{ is successor,} \\ \beta + \omega^{\{\gamma\}(x)} & \text{if } \gamma \text{ is limit.} \end{cases}$$

Finally, let $\{0\}(x) = 0$.

The sequence $(\{\alpha\}(x))_{x \in \mathbb{N}}$ is called the (canonical) *fundamental sequence* of α . When α is a limit ordinal, then it is a distinguished increasing sequence of ordinals that is cofinal in α .

The $\gg$ relation admits two desirable properties. First, if $\beta \gg \alpha$, then the sum $\beta + \alpha$ coincides with the natural sum (or Hessenberg sum) $\beta \oplus \alpha$ corresponding to the component-wise sum of the multiplicative factors. Second, if $\beta \gg \alpha$ and $\alpha \neq 0$, then the operation $\{\beta + \alpha\}(x)$ can be reduced to a computation of $\{\alpha\}(x)$:

Lemma 13.1.2 (Ketonen and Solovay [72]). *Let $\alpha, \beta < \epsilon_0$ be such that $\beta \gg \alpha$ and $\alpha \neq 0$. Then $\{\beta + \alpha\}(x) = \beta + \{\alpha\}(x)$.*

In general, fundamental sequences are not compatible with ordinal inequality, in the sense that there exist some ordinals $\beta > \alpha$ and some $x \in \mathbb{N}$ such that $\{\beta\}(x) < \{\alpha\}(x)$. For instance, if $\beta = \omega$ and $\alpha = 10$, then $\{\omega\}(1) = 1 < \{10\}(1) = 9$. However, this is the case if x is larger than every coefficient appearing in the Cantor normal form of α . This motivates the definition of the pseudo-norm:

Definition 13.1.3 (Bigorajska and Kotlarski [9]). The *pseudo-norm* $\text{psn}(\alpha)$ of α of Cantor normal form $\omega^{\alpha_s} \cdot a_s + \cdots + \omega^{\alpha_0} \cdot a_0$ is defined as

$$\text{psn}(\alpha) = \max(\text{psn}(\alpha_0), \dots, \text{psn}(\alpha_s), a_0, \dots, a_s)$$

The following lemma is an immediate consequence of Bigorajska and Kotlarski [9, Lemma 2.1].

¹The notation $B \ll A$ was introduced in Chapter 4 and used in Chapter 10 to denote that A is PA over B . As there is no risk of notational conflict, we will overload the symbol $\gg$.

Lemma 13.1.4 (Bigorajska and Kotlarski [9]). *Let $\alpha < \beta < \mathfrak{e}_0$ and $x \in \mathbb{N}$ such that $\text{psn}(\alpha) < x$. Then $\{\beta\}(x) \geq \alpha$ (with equality if and only if $\beta = \alpha + 1$).*

The following easy lemma relates $\text{psn}(\{\alpha\}(n))$ to $\text{psn}(\alpha)$ for any $n \in \mathbb{N}$.

Lemma 13.1.5. *Let $\alpha < \mathfrak{e}_0$ and let $n \in \mathbb{N} \setminus \{0\}$. Then $\text{psn}(\alpha) - 1 \leq \text{psn}(\{\alpha\}(n)) \leq \max\{\text{psn}(\alpha), n\}$*

Proof. The result is clear for $\alpha = 0$, then, if $\alpha \neq 0$, the operation $\{\alpha\}(n)$ can only add n 's and decrease coefficients by 1 in the (recursive) Cantor normal form of α . $\square$

The notion of fundamental sequence generalizes to finite sets as follows:

Definition 13.1.6 (Ketonen and Solovay [72]). Given an ordinal $\alpha < \mathfrak{e}_0$ and a finite set $X = \{x_0 < \dots < x_s\}$, let

$$\{\alpha\}(X) = \{\dots \{\{\alpha\}(x_0)\}(x_1) \dots\}(x_s)$$

A set X is α -large if $\{\alpha\}(X) = 0$, otherwise, it is α -small.

Accordingly, a finite set X is *exactly* α -large if it is α -large and $X \setminus \{\max X\}$ is α -small. A finite set X is *at most* α -large if $X \setminus \{\max X\}$ is α -small. By convention, we consider that $\emptyset$ is at most 0-large.

Note that a set X might be α -large, $\alpha + 1$ -small, but not at most α -large. For instance, $\{3, 5, 6, 7, 8\}$ is ω -large, not $\omega + 1$ -large, and not at most ω -large.

Ketonen and Solovay [72] proved many of the basic properties of α -large sets. In particular:

Proposition 13.1.7 (Ketonen and Solovay [72, Section 4.5]). *Let $\alpha < \mathfrak{e}_0$ and $X = \{x_0, \dots, x_{n-1}\} \subseteq_{\text{fin}} \mathbb{N}$ be α -large.*

- (Closure by superset) *If $Y \subseteq_{\text{fin}} \mathbb{N}$ is such that $X \subseteq Y$, then Y is also α -large.*
- (Regularity) *If $Y = \{y_0, \dots, y_{n-1}\}$ with $y_i \leq x_i$ for every $i < n$, then Y is also α -large.*

13.1.1 Hardy-like hierarchy

Ratajczyk [120, Section 5] defined a notion of largeness based on the Hardy hierarchy of functions [59]. As noted by Ratajczyk, this notion is slightly different from Ketonen and Solovay's α -largeness, and was extensively studied by Bigorajska and Kotlarski [8, 9, 81]. We now define a Hardy-like hierarchy whose induced notion of largeness corresponds exactly to Ketonen and Solovay's definition.

Definition 13.1.8. Fix an increasing function $h : A \rightarrow \mathbb{N}$ for some $A \subseteq \mathbb{N}$ with $h(a) > a$ for every $a \in A$. We define inductively a sequence of partial functions h_α for every $\alpha < \mathfrak{e}_0$ as follows: for $x \in \mathbb{N}$ and $\alpha > 0$, let $h_0(x) = x$ and let $h_\alpha(x) = h_{\{\alpha\}(x)}(h(x))$ if $h(x) \in \text{dom}(h_{\{\alpha\}(x)})$ and $x \in \text{dom}(h)$, otherwise leave $h_\alpha(x)$ undefined. We write $h_\alpha(x) \downarrow$ if $x \in \text{dom } h_\alpha$ and $h_\alpha(x) \uparrow$ otherwise.

Note that for every $\alpha < \mathfrak{e}_0$ and $x \in \text{dom } h_\alpha$, there is some $n \in \mathbb{N}$ such that $h_\alpha(x) = h^{(n)}(x)$. However, this n depends on x : if $\alpha \geq \omega$, there is no n such that $h_\alpha = h^{(n)}$.

Remark 13.1.9. In the definition of the original Hardy hierarchy [59] based on h , if α is limit, then $h_\alpha(x) = h_{\{\alpha\}(x)}(x)$, while in our Hardy-like hierarchy, we define $h_\alpha(x) = h_{\{\alpha\}(x)}(h(x))$. This variation reflects the main difference between the notion of largeness from Ratajczyk [120, Section 5] and the one of Ketonen and Solovay [72]. In the former case, a set A is α -large for a limit ordinal α if it is $\{\alpha\}(\min A)$ -large, while in the latter case, a set A is α -large if $A \setminus \{\min A\}$ is $\{\alpha\}(\min A)$ -large. Some

further differences exist: in the finite case, a set A is n -large in the sense of Ratajczyk if $\text{card } A \geq n+1$ while in Ketonen and Solovay's sense, a set A is n -large iff $\text{card } A \geq n$. As a consequence, under both definitions, a set A is ω -large iff $\text{card } A > \min A$.

Lemma 13.1.10. *If $h_\alpha(x) \downarrow$ for some $\alpha < \epsilon_0$ and $x \in \mathbb{N} \setminus \{0\}$, then $h_\alpha(x) \geq \text{psn}(\alpha) + x$. Furthermore, the inequality is strict when $\alpha \geq \omega$.*

Proof. By induction on α . For $\alpha = 0$, let $x \in \mathbb{N} \setminus \{0\}$ be such that $h_0(x) \downarrow$, then $h_0(x) = x = \text{psn}(0) + x$. Let $\alpha < \epsilon_0$ and assume the property to hold for every $\alpha' < \alpha$. Let $x \in \mathbb{N}$ be such that $h_\alpha(x) \downarrow$, then $h_\alpha(x) = h_{\{\alpha\}(x)}(h(x))$. By the inductive hypothesis, we have $h_{\{\alpha\}(x)}(h(x)) \geq \text{psn}(\{\alpha\}(x)) + h(x)$, but, as $h(x) \geq x + 1$ and as $\text{psn}(\{\alpha\}(x)) \geq \text{psn}(\alpha) - 1$ by Lemma 13.1.5, we have $h_{\{\alpha\}(x)}(h(x)) \geq \text{psn}(\alpha) + x$. This concludes our induction.

For strictness: when $\alpha = \omega$, we have $\{\alpha\}(x) = x$, so $\text{psn}(\{\alpha\}(x)) = \text{psn}(x) = x > \text{psn}(\alpha) - 1 = 0$. Thus, the inequality $\text{psn}(\{\alpha\}(x)) \geq \text{psn}(\alpha) - 1$ used above is strict, and hence so is $h_\alpha(x) \geq \text{psn}(\alpha) + x$. The remaining cases $\alpha > \omega$ follow by the same induction as above, propagating strictness upward. $\square$

Fix an increasing function $h : A \rightarrow \mathbb{N}$ for some $A \subseteq \mathbb{N}$, with $h(x) > x$ for every $x \in A$. The Hardy-like iterations of h yield faster-growing functions on a restricted domain, that is, if $\alpha > 0$, then $\text{dom } h_\alpha \subseteq \text{dom } h$ and h_α dominates h on $\text{dom } h_\alpha$ (see Lemma 13.1.14(4)).

On the other hand, given two ordinals $\alpha \leq \beta < \epsilon_0$ and some $x \in \text{dom } h_\beta$, it is not true in general that $x \in \text{dom } h_\alpha$, nor that if $x \in \text{dom } h_\beta \cap \text{dom } h_\alpha$, then $h_\beta(x) \geq h_\alpha(x)$. However, if x is sufficiently large — $x > \text{psn}(\alpha)$ suffices —, then this is the case (see Lemmas 13.1.12 and 13.1.14). However, the $x > \text{psn}(\alpha)$ hypothesis is too strong for our purposes, so we will use a more general relation $\beta \Rightarrow_x \alpha$ which depends on both β and α , and such that if $\beta \Rightarrow_x \alpha$, then $\beta \geq \alpha$ and $h_\beta(x) \geq h_\alpha(x)$.

Definition 13.1.11 (Ketonen and Solovay [72]). Given two ordinals $\alpha, \beta < \epsilon_0$ and $x \geq 0$, let $\beta \Rightarrow_x \alpha$ if either $\beta = \alpha$, or $\{\beta\}(x) \Rightarrow_x \alpha$.

Bigorajska and Kotlarski [8, Lemma 2(viii)] proved that the definition is unchanged if one replaces $\{\beta\}(x)$ by $\{\beta\}(y)$ for some $y \leq x$. As one expects from the previous discussion, the relation $\beta \Rightarrow_x \alpha$ coincides with the inequality $\beta \geq \alpha$ whenever $x > \text{psn}(\alpha)$. The whole interest of this relation is therefore in the case $x \leq \text{psn}(\alpha)$.

The following lemma strengthens Ketonen and Solovay [72, Proposition 2.8.1] and is an immediate consequence of Bigorajska and Kotlarski [9, Lemma 2.1]. It appears in this form in the posthumous book of Kotlarski [81, Lemma 1.4.3].

Lemma 13.1.12 (Kotlarski [81]). *Let $\alpha \leq \beta < \epsilon_0$ and $x \in \mathbb{N}$ such that $\text{psn}(\alpha) < x$, then $\beta \Rightarrow_x \alpha$.*

Ketonen and Solovay [72] proved many properties about the relation $\Rightarrow_x$. The following lemma compiles a list of basic facts that shall be useful to prove our main theorem.

Lemma 13.1.13 (Ketonen and Solovay [72]). *Let $\alpha, \beta, \lambda < \epsilon_0$ and $x \geq 0$.*

- (1) *If $\lambda \gg \alpha$ and $\alpha \Rightarrow_x \beta$, then $(\lambda + \alpha) \Rightarrow_x (\lambda + \beta)$;*
- (2) *If $k < \ell \in \mathbb{N}$, then $\omega^\alpha \cdot \ell \Rightarrow_x \omega^\alpha \cdot k$;*
- (3) *If $\alpha \Rightarrow_x \beta$ and $\alpha > \beta + 1$, then $\alpha \Rightarrow_{x+1} \beta + 1$.*
- (4) *If $\alpha \Rightarrow_x \beta$ and $x \geq 1$, then $\omega^\alpha \Rightarrow_x \omega^\beta$;*
- (5) *If $x \geq 1$ and $k < \ell \in \mathbb{N}$, then $\{\alpha\}(\ell) \Rightarrow_x \{\alpha\}(k)$.*

Proof. (1) is [72, Lemma 1, part 2.4], (2) is [72, Lemma 3, part 2.4], (3) is [72, Lemma 2, part 2.6], (4) is [72, Lemma 5, part 2.4], (5) is [72, Theorem 2.4] for α limit. If α is 0 or successor, then $\{\alpha\}(\ell) = \{\alpha\}(k)$. $\square$

The following lemma states that the $\Rightarrow_x$ relation is a sufficient hypothesis for the Hardy-like hierarchy of functions to behave nicely. Lemma 13.1.14(1,3) is a Ketonen-Solovay counterpart of Bigorajska and Kotlarski [8, Lemma 4]:

Lemma 13.1.14. *Let h be as above. Let $0 < \alpha, \beta < \mathfrak{e}_0$ and $x \in \mathbb{N}$. Then:*

- (1) h_β is increasing
- (2) if $x \in \text{dom } h$ and $h(x) \in \text{dom } h_\beta$ then $x \in \text{dom } h_\beta$
- (3) if $\beta \Rightarrow_x \alpha$ and $x \in \text{dom } h_\beta$ then $x \in \text{dom } h_\alpha$ and $h_\alpha(x) \leq h_\beta(x)$
- (4) if $x \in \text{dom } h_\beta$, then $h_\beta(x) \geq h(x)$

Proof. We prove (1),(2) and (3) by mutual induction on β . If $\beta = 1$ then the property holds as $h_1 = h$ is increasing and the only $0 < \alpha < \mathfrak{e}_0$ such that $\beta \Rightarrow_x \alpha$ is $\alpha = 1$.

Let $\beta > 1$ and assume that (1),(2) and (3) hold for every $0 < \beta' < \beta$. Let $x, y \in \text{dom } h_\beta$ such that $x < y$, then $h_\beta(y) = h_{\{\beta\}(y)}(h(y))$. As $\{\beta\}(y) < \beta$ and as, by Lemma 13.1.13(5), $\{\beta\}(y) \Rightarrow_{h(y)} \{\beta\}(x)$, then by (3) for $\{\beta\}(y)$ we have $h(y) \in \text{dom } h_{\{\beta\}(x)}$ and $h_{\{\beta\}(y)}(h(y)) \geq h_{\{\beta\}(x)}(h(y))$. From $x \in \text{dom } h_\beta$ we get $h(x) \in \text{dom } h_{\{\beta\}(x)}$, hence, by (1) for $\{\beta\}(x)$, we get $h_{\{\beta\}(x)}(h(y)) > h_{\{\beta\}(x)}(h(x))$. Thus, $h_\beta(y) > h_\beta(x)$ and h_β is increasing. Thus (1) holds for β .

Let $x \in \text{dom } h$ be such that $h(x) \in \text{dom } h_\beta$. Then we have $h_\beta(h(x)) = h_{\{\beta\}(h(x))}(h(h(x)))$ hence $h(h(x)) \in \text{dom } h_{\{\beta\}(h(x))}$. By (2) for $\{\beta\}(h(x))$ we have $h(x) \in \text{dom } h_{\{\beta\}(h(x))}$. Then, as Lemma 13.1.13(5) $\{\beta\}(h(x)) \Rightarrow_{h(x)} \{\beta\}(x)$, we also have $h(x) \in \text{dom } h_{\{\beta\}(x)}$ by (3) for $\{\beta\}(h(x))$, hence $x \in \text{dom } h_\beta$. Thus (2) holds for β .

Let $x \in \mathbb{N}$ and $0 < \alpha$ be such that $\beta \Rightarrow_x \alpha$ and $h_\beta(x) \downarrow$. Since $\beta \Rightarrow_x \alpha$, either $\beta = \alpha$ or $\{\beta\}(x) \Rightarrow_x \alpha$. If $\beta = \alpha$ then $h_\alpha(x) \downarrow$ and $h_\alpha(x) = h_\beta(x)$, in which case (3) holds. If $\{\beta\}(x) \Rightarrow_x \alpha$, then as $h_\beta(x) = h_{\{\beta\}(x)}(h(x))$, we have $x \in \text{dom } h_{\{\beta\}(x)}$ by (2) for $\{\beta\}(x)$ and $h_{\{\beta\}(x)}(h(x)) > h_{\{\beta\}(x)}(x)$ by (1) for $\{\beta\}(x)$. From $\{\beta\}(x) \Rightarrow_x \alpha$ and (3) for $\{\beta\}(x)$, we get $x \in \text{dom } h_\alpha$ and $h_{\{\beta\}(x)}(x) \geq h_\alpha(x)$. So $h_\beta(x) = h_{\{\beta\}(x)}(h(x)) > h_{\{\beta\}(x)}(x) \geq h_\alpha(x)$, Thus (3) holds for β .

Last, since $\beta \Rightarrow_x 1$ for every $\beta > 0$ and $x \in \mathbb{N}$, then (4) follows from (3). $\square$

The following lemma is the Ketonen-Solovay counterpart of Bigorajska and Kotlarski [8, Lemma 7]. Formulated in terms of largeness, it corresponds to Ketonen and Solovay [72, Lemma 4.6].

Lemma 13.1.15. *Let $\alpha \ll \beta$ and $x \in \mathbb{N}$, if $h_{\beta+\alpha}(x) \downarrow$ then $h_\beta(h_\alpha(x)) \downarrow$ and $h_{\beta+\alpha}(x) = h_\beta(h_\alpha(x))$.*

Proof. By induction on α . If $\alpha = 0$ then, for $x \in \mathbb{N}$ and $\beta \gg \alpha$ such that $h_{\beta+\alpha}(x) \downarrow$, we have $h_\beta(x) \downarrow$, then, as $h_0(x) = x$ we have $h_\beta(x) = h_\beta(h_\alpha(x)) \downarrow$ and $h_\beta(h_\alpha(x)) = h_{\beta+\alpha}(x)$.

Let $\alpha > 0$ and assume the property to hold for every $\alpha' < \alpha$. Let $x \in \mathbb{N}$ and $\beta \gg \alpha$, if $h_{\beta+\alpha}(x) \downarrow$ then $h_{\{\beta+\alpha\}(x)}(h(x)) \downarrow$ and $h_{\{\beta+\alpha\}(x)}(h(x)) = h_{\beta+\alpha}(x)$. By Lemma 13.1.2, $\{\beta+\alpha\}(x) = \beta + \{\alpha\}(x)$, so $h_{\{\beta+\alpha\}(x)}(h(x)) = h_{\beta+\alpha}(x)$. As $\{\alpha\}(x) < \alpha$ and $\{\alpha\}(x) \ll \beta$, by the induction hypothesis we have $h_\beta(h_{\{\alpha\}(x)}(h(x))) \downarrow$ and $h_\beta(h_{\{\alpha\}(x)}(h(x))) = h_{\beta+\{\alpha\}(x)}(h(x)) = h_{\beta+\alpha}(x)$. And, as $h_{\{\alpha\}(x)}(h(x)) = h_\alpha(x)$, we have $h_{\beta+\alpha}(x) = h_\beta(h_\alpha(x))$. $\square$

We now prove our core proposition. The intuition goes as follows: Suppose $0 < \lambda < \mathfrak{e}_0$ and $x > 0$. The evaluation of $h_\lambda(x)$ unfolds into $h_{\{\lambda\}(x)}(h(x))$, then $h_{\{\{\lambda\}(x)\}(h(x))}(h(h(x)))$, and so on. Thus, the i th iteration of the evaluation of $h_\lambda(x)$, if defined, corresponds to $h_{\lambda_i}(x_i)$, where x_i and λ_i are defined inductively by $x_0 = x$,

$\lambda_0 = \lambda$ and $x_{i+1} = h(x_i)$, $\lambda_{i+1} = \{\lambda_i\}(x_i)$. Note that since the function h is increasing, so is the sequence $x_0 < x_1 < \dots$.

The definition of $\{\lambda\}(x)$ is done by considering the short Cantor normal form, by reducing first the smallest component of the Cantor normal form. The evaluation either decreases a successor ordinal by one, or replaces a limit ordinal by a smaller one, with coefficient x . Since the sequence $\lambda_0 > \lambda_1 > \dots$ is defined by successive evaluations with an increasing sequence $x_0 < x_1 < \dots$, the higher components will be evaluated with larger values than the smaller components, and therefore this evaluation order ensures that the sequence $\lambda_0 > \lambda_1 > \dots$ decreases as slowly as possible.

Based on this evaluation order, if $\lambda = \beta + \alpha$, with $\beta \gg \alpha$, the sequence $\lambda_0 > \lambda_1 > \dots$ corresponds to evaluating α successively with $x_0, x_1, \dots$ until we reach a stage i such that $\alpha_i = 0$. Then we evaluate β with $x_{i+1}, x_{i+2}, \dots$.

One could consider other evaluation orders, in which one would alternate some evaluation steps of β , and some of α . In terms of the Hardy-like hierarchy, one evaluation step of β corresponds to $h_{\{\beta\}(x) \oplus \alpha}(h(x))$ while one evaluation step of α corresponds to $h_{\beta + \{\alpha\}(x)}(h(x))$. Since the standard evaluation order is the one making the sequence $\lambda_0 > \lambda_1 > \dots$ decrease as slowly as possible, this translates into the following inequality:

$$h_{\beta + \alpha}(x) = h_{\beta + \{\alpha\}(x)}(h(x)) \geq h_{\{\beta\}(x) \oplus \alpha}(h(x))$$

The following proposition formalizes this intuition in the case $\beta \gg \alpha$. It will then be proven in its most general form in Theorem 13.1.18.

Proposition 13.1.16. *Let $0 < \alpha, \beta < \epsilon_0$ with $\beta \gg \alpha$ and $x > 0$ an integer. If $h_{\beta + \alpha}(x) \downarrow$, then $h_{\{\beta\}(x) \oplus \alpha}(h(x)) \downarrow$ and $h_{\beta + \alpha}(x) = h_{\beta + \{\alpha\}(x)}(h(x)) \geq h_{\{\beta\}(x) \oplus \alpha}(h(x))$.*

Proof. The proof is by case analysis on the Cantor normal forms, and not by induction. Write $\alpha = \omega^{\alpha_n} \cdot a_n + \dots + \omega^{\alpha_0} \cdot a_0$ and $\beta = \omega^{\beta_k} \cdot b_k + \dots + \omega^{\beta_0} \cdot b_0$ their Cantor normal form. From the assumption $\beta \gg \alpha$, we have $\beta_0 \geq \alpha_n$.

Case 1: $n = 0$ and $\alpha_n = \beta_0$. In that case, $\beta + \{\alpha\}(x) = \{\beta\}(x) \oplus \alpha$, hence $h_{\{\beta\}(x) \oplus \alpha}(h(x)) \downarrow$ and $h_{\beta + \{\alpha\}(x)}(h(x)) = h_{\{\beta\}(x) \oplus \alpha}(h(x))$.

If we are not in this case, we can freely assume that $\alpha_n < \beta_0$ in the Cantor normal form of α and β . Indeed, if $\alpha_n = \beta_0$, we can consider $\alpha' = \omega^{\alpha_{n-1}} a_{n-1} + \dots + \omega^{\alpha_0} a_0$ and $\beta' = \omega^{\beta_k} b_k + \dots + \omega^{\beta_0} (b_0 + a_n)$, we then have $0 < \alpha', \beta' < \epsilon_0$, $\alpha' \ll \beta'$, $\beta + \alpha = \beta' + \alpha'$, $\beta + \{\alpha\}(x) = \beta' + \{\alpha'\}(x)$ and $\{\beta\}(x) \oplus \alpha = \{\beta'\}(x) \oplus \alpha'$, making it sufficient to prove the result for α' and β' .

Case 2: $\alpha_n < \beta_0$ and β_0 is successor. Let β'_0 be such that $\beta_0 = \beta'_0 + 1$. In that case, write $\beta = \delta + \omega^{\beta'_0 + 1}$. By our assumption that $\alpha_n < \beta_0$, we have $\omega^{\beta'_0 + 1} > \alpha$ and $\omega^{\beta'_0} \gg \alpha$. By Lemma 13.1.15, letting $c = h_{\{\alpha\}(x)}(h(x)) = h_\alpha(x)$, we have

$$h_{\beta + \{\alpha\}(x)}(h(x)) = h_\beta(c) = h_{\delta + \omega^{\beta'_0 + 1} \cdot c}(h(c))$$

Consider the following two subcases:

Subcase 2.1: $\alpha < \omega$. In that case, $\alpha = k$ for some $k \in \mathbb{N}$. This gives us $h_{\beta + \alpha}(x) = h_\beta(h^k(x)) = h_{\{\beta\}(h^k(x))}(h^{k+1}(x))$ (where h^ℓ denotes the composition of h with itself ℓ times). Then, as h is increasing, we have $\{\beta\}(h^k(x)) \Rightarrow_{h^{k+1}(x)} \{\beta\}(x)$ by Lemma 13.1.13(5). Hence, by Lemma 13.1.14(3), we have $h_{\{\beta\}(x)}(h^{k+1}(x)) \downarrow$ and

$$h_{\{\beta\}(h^k(x))}(h^{k+1}(x)) \geq h_{\{\beta\}(x)}(h^{k+1}(x)) = h_{\{\beta\}(x) + \alpha}(h(x)) = h_{\{\beta\}(x) \oplus \alpha}(h(x))$$

Hence $h_{\beta + \alpha}(x) \geq h_{\{\beta\}(x) \oplus \alpha}(h(x))$, which is what we wanted.

Subcase 2.2: $\alpha \geq \omega$. In that case, we can use the strict inequality obtained in Lemma 13.1.10 to get $c > \text{psn}(\alpha) + x$.

As $\beta_0 > \alpha_n$, we get $\beta'_0 \geq \alpha_n$ and therefore have $\omega^{\beta'_0} \cdot (\text{psn}(\alpha) + 1) > \alpha$ by definition of $\text{psn}(\alpha)$. Thus, by Lemma 13.1.12, we have $\omega^{\beta'_0} \cdot (\text{psn}(\alpha) + 1) \Rightarrow_{h(c)} \alpha$ (using the fact that $h(c) > c > \text{psn}(\alpha)$).

We have $\omega^{\beta'_0} \cdot c \Rightarrow_{h(c)} \omega^{\beta'_0} \cdot (x + \text{psn}(\alpha) + 1) = \omega^{\beta'_0} \cdot x + \omega^{\beta'_0} \cdot (\text{psn}(\alpha) + 1)$ by Lemma 13.1.13(2). Then, as $\omega^{\beta'_0} \cdot x \gg \omega^{\beta'_0} \cdot (\text{psn}(\alpha) + 1)$, we can apply Lemma 13.1.13(1) and obtain $\omega^{\beta'_0} \cdot x + \omega^{\beta'_0} \cdot (\text{psn}(\alpha) + 1) \Rightarrow_{h(c)} \omega^{\beta'_0} \cdot x + \alpha$. Hence that $\omega^{\beta'_0} \cdot c \Rightarrow_{h(c)} \omega^{\beta'_0} \cdot x + \alpha$.

By another application of Lemma 13.1.13(1) (using the fact that $\delta \gg \omega^{\beta'_0} \cdot c$), we obtain $\delta + \omega^{\beta'_0} \cdot c \Rightarrow_{h(c)} \delta + \omega^{\beta'_0} \cdot x + \alpha$.

By Lemma 13.1.14(3), we thus have $h_{\delta + \omega^{\beta'_0} \cdot x + \alpha}(h(c)) \downarrow$ and

$$h_{\delta + \omega^{\beta'_0} \cdot c}(h(c)) \geq h_{\delta + \omega^{\beta'_0} \cdot x + \alpha}(h(c))$$

As $h(c) = h(h_\alpha(x)) = h^\ell(h(x))$ for some $\ell \in \mathbb{N}$, we have, by several applications of Lemma 13.1.14(2) that $h_{\delta + \omega^{\beta'_0} \cdot x + \alpha}(h(x)) \downarrow$ and that

$$h_{\delta + \omega^{\beta'_0} \cdot x + \alpha}(h(c)) \geq h_{\delta + \omega^{\beta'_0} \cdot x + \alpha}(h(x)) = h_{\{\beta\}(x) \oplus \alpha}(h(x))$$

Which is what we wanted.

Case 3: $\alpha_n < \beta_0$ and β_0 **limit**. In that case, write $\beta = \delta + \omega^{\beta_0}$. Similarly, by Lemma 13.1.2 and Lemma 13.1.15, letting $c = h_{\{\alpha\}(x)}(h(x)) = h_\alpha(x)$, we have

$$h_{\beta + \alpha}(x) = h_{\beta + \{\alpha\}(x)}(h(x)) = h_\beta(c) = h_{\delta + \omega^{\{\beta_0\}(c)}}(h(c))$$

Since $c > x$, we have $\{\beta_0\}(c) \Rightarrow_y \{\beta_0\}(x) + 1$ for every $y \geq 2$ by Lemma 13.1.13(3,5).

By Lemma 13.1.10, $c \geq \text{psn}(\alpha) + x > \text{psn}(\alpha)$. Thus, by Lemma 13.1.4, we have $\{\beta_0\}(c) > \alpha_n$, and, by Lemma 13.1.12, $\{\beta_0\}(c) \Rightarrow_c \alpha_n$. By Lemma 13.1.13(3), we have $\{\beta_0\}(c) \Rightarrow_{h(c)} \alpha_n + 1$, as $h(c) > c$.

Thus, we have both $\{\beta_0\}(c) \Rightarrow_{h(c)} \{\beta_0\}(x) + 1$ and $\{\beta_0\}(c) \Rightarrow_{h(c)} \alpha_n + 1$. Depending on whether α_n or $\{\beta_0\}(x)$ appears first in the sequence $\{\beta_0\}(c) \Rightarrow_{h(c)} \{\{\beta_0\}(c)\}(h(c)) \Rightarrow_{h(c)} \{\{\{\beta_0\}(c)\}(h(c))\}(h(c)) \Rightarrow_{h(c)} \dots$, we either have $\{\beta_0\}(x) \Rightarrow_{h(c)} \alpha_n$ or $\alpha_n \Rightarrow_{h(c)} \{\beta_0\}(x)$.

We can then consider the following two subcases:

Subcase 3.1: $\alpha_n > \{\beta_0\}(x)$, in which case $\alpha_n \Rightarrow_{h(c)} \{\beta_0\}(x)$. By Lemma 13.1.13(2,4), we have

$$\omega^{\{\beta_0\}(c)} \Rightarrow_{h(c)} \omega^{\alpha_n + 1} \Rightarrow_{h(c)} \omega^{\alpha_n} \cdot h(c) \Rightarrow_{h(c)} \omega^{\alpha_n} \cdot (c + 1)$$

Thus, by Lemma 13.1.13(1), since $\delta \gg \omega^{\{\beta_0\}(c)}$,

$$\delta + \omega^{\{\beta_0\}(c)} \Rightarrow_{h(c)} \delta + \omega^{\alpha_n} \cdot (c + 1)$$

Therefore, by Lemma 13.1.14(3), $h_{\delta + \omega^{\alpha_n} \cdot (c+1)}(h(c)) \downarrow$, and, letting $d = h_{\omega^{\alpha_n}}(h(c))$,

$$h_{\delta + \omega^{\{\beta_0\}(c)}}(h(c)) \geq h_{\delta + \omega^{\alpha_n} \cdot (c+1)}(h(c)) = h_{\delta + \omega^{\alpha_n} \cdot c}(d)$$

As $\omega^{\alpha_n} \Rightarrow_{h(c)} \omega^{\{\beta_0\}(x)}$ by Lemma 13.1.13(4), we have $h_{\omega^{\{\beta_0\}(x)}}(h(c)) \downarrow$ and $d \geq h_{\omega^{\{\beta_0\}(x)}}(h(c))$ by Lemma 13.1.14(3).

By Lemma 13.1.10, $h_{\omega^{\{\beta_0\}(x)}}(h(c)) \geq \text{psn}(\{\beta_0\}(x)) + h(c)$. Hence, we have $d \geq \text{psn}(\{\beta_0\}(x)) + h(c) > \text{psn}(\{\beta_0\}(x)) + \text{psn}(\alpha)$ as $h(c) > c > \text{psn}(\alpha)$.

Then, from $c > \text{psn}(\alpha)$ and $\alpha_n > \{\beta_0\}(x)$, we get $\omega^{\alpha_n} \cdot c > \omega^{\{\beta_0\}(x)} \oplus \alpha$, and, as $d > \text{psn}(\omega^{\{\beta_0\}(x)}) + \text{psn}(\alpha) \geq \text{psn}(\omega^{\{\beta_0\}(x)} \oplus \alpha)$, we have $\omega^{\alpha_n} \cdot c \Rightarrow_d \omega^{\{\beta_0\}(x)} \oplus \alpha$ by Lemma 13.1.12.

Again, as $\delta \gg \omega^{\{\beta_0\}(c)}$, we can apply Lemma 13.1.13(1) and obtain $\delta + \omega^{\alpha_n} \cdot c \Rightarrow_d \delta + (\omega^{\{\beta_0\}(x)} \oplus \alpha) = \{\beta\}(x) \oplus \alpha$.

Thus, by Lemma 13.1.14(3), $h_{\{\beta\}(x) \oplus \alpha}(d) \downarrow$ and

$$h_{\delta + \omega^{\alpha_n} \cdot c}(d) \geq h_{\{\beta\}(x) \oplus \alpha}(d)$$

As $d = h^\ell(h(x))$ for some $\ell \in \mathbb{N}$, we have, by several applications of Lemma 13.1.14(2) that $h_{\{\beta\}(x) \oplus \alpha}(h(x)) \downarrow$ and that $h_{\{\beta\}(x) \oplus \alpha}(d) \geq h_{\{\beta\}(x) \oplus \alpha}(h(x))$.

Combined with the other inequalities, we get $h_{\beta+\alpha}(x) \geq h_{\{\beta\}(x) \oplus \alpha}(h(x))$.

Subcase 3.2: $\alpha_n \leq \{\beta_0\}(x)$, in which case $\{\beta_0\}(x) \Rightarrow_{h(c)} \alpha_n$. By Lemma 13.1.13(2,4), we have

$$\omega^{\{\beta_0\}(c)} \Rightarrow_{h(c)} \omega^{\{\beta_0\}(x)+1} \Rightarrow_{h(c)} \omega^{\{\beta_0\}(x)} \cdot h(c) \Rightarrow_{h(c)} \omega^{\{\beta_0\}(x)} \cdot (c+1)$$

Then, from $c > \text{psn}(\alpha)$ and $\alpha_n \leq \{\beta_0\}(x)$, we get $\omega^{\{\beta_0\}(x)} \cdot c > \alpha$, hence, by Lemma 13.1.12 we get $\omega^{\{\beta_0\}(x)} \cdot c \Rightarrow_{h(c)} \alpha$. As $\omega^{\{\beta_0\}(x)} \gg \alpha$, Lemma 13.1.13(1) gives us

$$\omega^{\{\beta_0\}(x)} \cdot (c+1) = \omega^{\{\beta_0\}(x)} + \omega^{\{\beta_0\}(x)} \cdot c \Rightarrow_{h(c)} \omega^{\{\beta_0\}(x)} + \alpha$$

Then, as $\delta \gg \omega^{\{\beta_0\}(c)}$, another application of Lemma 13.1.13(1) gives us

$$\delta + \omega^{\{\beta_0\}(c)} \Rightarrow_{h(c)} \delta + \omega^{\{\beta_0\}(x)} + \alpha = (\delta + \omega^{\{\beta_0\}(x)}) \oplus \alpha = \{\beta\}(x) \oplus \alpha$$

Thus, by Lemma 13.1.14(3), $h_{\{\beta\}(x) \oplus \alpha}(x) \downarrow$ and

$$h_{\delta + \omega^{\{\beta_0\}(c)}}(h(c)) \geq h_{\{\beta\}(x) \oplus \alpha}(h(c))$$

Again, by several applications of Lemma 13.1.14(2), we have $h_{\{\beta\}(x) \oplus \alpha}(h(x)) \downarrow$ and $h_{\{\beta\}(x) \oplus \alpha}(h(c)) \geq h_{\{\beta\}(x) \oplus \alpha}(h(x))$.

Hence, we have $h_{\beta+\alpha}(x) \geq h_{\{\beta\}(x) \oplus \alpha}(h(x))$.

This completes our proof of Proposition 13.1.16. $\square$

Lemma 13.1.17. *Let $\alpha, \beta < \epsilon_0$, then there exists $\alpha', \beta' < \epsilon_0$ with $\alpha' \ll \beta'$ and such that $\beta \oplus \alpha = \beta' + \alpha'$ and $\{\beta\}(x) \oplus \alpha = \{\beta'\}(x) \oplus \alpha'$ for every $x \in \mathbb{N}$*

Proof. If $\beta = 0$ then take $\alpha' = \alpha$ and $\beta' = \beta$. If $\beta \neq 0$, then write $\alpha = \omega^{\gamma_n} \cdot a_n + \dots + \omega^{\gamma_0} \cdot a_0$ and $\beta = \omega^{\gamma_n} \cdot b_n + \dots + \omega^{\gamma_0} \cdot b_0$ with $a_0, \dots, a_n, b_0, \dots, b_n \in \mathbb{N}$ and $\gamma_0 < \dots < \gamma_n < \epsilon_0$. Let $i \leq n$ be the smallest index such that $b_i \neq 0$, and let $\beta' = \omega^{\gamma_n} \cdot (b_n + a_n) + \dots + \omega^{\gamma_i} \cdot (b_i + a_i)$ and $\alpha' = \omega^{\gamma_{i-1}} \cdot (b_{i-1} + a_{i-1}) + \dots + \omega^{\gamma_0} \cdot (b_0 + a_0)$.

We have $\beta \oplus \alpha = \beta' + \alpha' = \omega^{\gamma_n} (b_n + a_n) + \dots + \omega^{\gamma_0} (b_0 + a_0)$ and $\{\beta'\}(x) \oplus \alpha' = (\omega^{\gamma_n} (b_n + a_n) + \dots + \omega^{\gamma_i} (b_i + a_i - 1) + \dots + \omega^{\gamma_0} (b_0 + a_0)) \oplus \{\omega^{\gamma_i}\}(x) = \{\beta\}(x) \oplus \alpha$ $\square$

The following theorem informally states that the evaluation order of the fundamental sequences is the one which makes the Hardy-like hierarchy grow the fastest. In the case $\alpha \ll \beta$, $\beta + \alpha = \beta \oplus \alpha$, hence it generalizes Proposition 13.1.16.

Theorem 13.1.18 (Optimal growth). *Let $\alpha, \beta < \epsilon_0$ with $\beta > 0$ and $x > 0$ an integer. If $h_{\beta \oplus \alpha}(x) \downarrow$ then $h_{\{\beta\}(x) \oplus \alpha}(h(x)) \downarrow$ and $h_{\beta \oplus \alpha}(x) = h_{\{\beta \oplus \alpha\}(x)}(h(x)) \geq h_{\{\beta\}(x) \oplus \alpha}(h(x))$.*

Proof. Let $\alpha', \beta' < \epsilon_0$ be obtained from Lemma 13.1.17 applied on α, β . As $\beta \oplus \alpha = \beta' + \alpha'$ and $\{\beta\}(x) \oplus \alpha = \{\beta'\}(x) \oplus \alpha'$, it is sufficient to show that $h_{\{\beta'\}(x) \oplus \alpha'}(h(x)) \downarrow$ and that $h_{\beta' \oplus \alpha'}(x) = h_{\{\beta' \oplus \alpha'\}(x)}(h(x)) \geq h_{\{\beta'\}(x) \oplus \alpha'}(h(x))$.

If $\alpha' = 0$, then the result holds by definition of $h_{\beta'}$.

If $\alpha' \neq 0$, then the result is that of Proposition 13.1.16. $\square$

13.1.2 Partition theorem

We now translate this abstract study of the Hardy-like hierarchy into the framework of α -largeness, to prove Main Theorem 13.0.1.

Definition 13.1.19. Let $A = \{x_0, \dots, x_{n-1}\} \subseteq \mathbb{N}$ a finite set. We define $h^A : A \setminus \{\max A\} \rightarrow A$ by $h^A(x_i) = x_{i+1}$ for every $i < n-1$. In other words, h^A is the successor operation on the ordered set $(A, <_{\mathbb{N}})$.

The following lemma is the heart of the correspondence between Hardy-like hierarchies and of Ketonen and Solovay's α -largeness.

Lemma 13.1.20. *Let $0 < \alpha < \mathfrak{e}_0$, then A is at most α -large iff $h_\alpha^A(\min A) \uparrow$.*

Proof. By induction on α . The result holds for $\alpha = 1$, as $h_1^A(\min A) \uparrow$ if and only if $\text{card } A \leq 1$. Let $1 < \alpha < \mathfrak{e}_0$ and assume the property holds for every $\beta < \alpha$. Let $A = \{x_0, x_1, \dots, x_n\} \subseteq \mathbb{N}$ be such that $h_\alpha^A(x_0) \uparrow$, then, $h_{\{\alpha\}(x_0)}^A(x_1) \uparrow$ and therefore $h_{\{\alpha\}(x_0)}^{A \setminus \{x_0\}}(x_1) \uparrow$. By the inductive hypothesis, $A \setminus \{x_0\}$ is therefore at most $\{\alpha\}(x_0)$ -large, hence A is therefore not at most α -large.

Reciprocally, if $h_\alpha^A(x_0) \downarrow$, then we also have $h_{\{\alpha\}(x_0)}^A(x_1) \downarrow$, hence $A \setminus \{x_0\}$ is not at most $\{\alpha\}(x_0)$ -large by the inductive hypothesis, and A is not at most α -large. $\square$

The following theorem is an iterated counterpart of Theorem 13.1.18 in the language of α -largeness. It has several consequences, such as the splitting property (Theorem 13.1.21) and multiple versions of the pigeonhole principle (Theorems 13.1.22 and 13.1.23)

Main Theorem 13.0.1. Let $\beta, \gamma < \mathfrak{e}_0$ be two ordinals, and $B, C \subseteq \mathbb{N} \setminus \{0\}$ be at most β -large and at most γ -large respectively. Then $B \cup C$ is at most $(\beta \oplus \gamma)$ -large.

Proof. Without loss of generality, we can assume that $B \cap C = \emptyset$. By assumption we have $h_\beta^B(\min B) \uparrow$ and $h_\gamma^C(\min C) \uparrow$. We want to show that $h_{\beta \oplus \gamma}^{B \cup C}(\min\{\min B, \min C\}) \uparrow$.

Write $B \cup C = \{x_0 < x_1 < \dots < x_{k-1}\}$ and consider the sequences of ordinals $(\beta_i)_{i < k}$ and $(\gamma_i)_{i < k}$ defined inductively as follows: $\beta_0 = \beta$ and $\gamma_0 = \gamma$ and if β_i and γ_i have been defined, let $\beta_{i+1} = \{\beta_i\}(x_i)$ and $\gamma_{i+1} = \gamma_i$ if $x_i \in B$ and $\beta_{i+1} = \beta_i$ and $\gamma_{i+1} = \{\gamma_i\}(x_i)$ if $x_i \in C$.

Since B is at most β -large, then $\{\beta\}(B \setminus \{\max B\}) > 0$, thus, if $x_i \in B$ then $\beta_i \neq 0$, similarly, if $x_i \in C$ then $\gamma_i \neq 0$. Hence, if $x_{k-1} \in B$ then $\beta_{k-1} \neq 0$ and if $x_{k-1} \in C$ then $\gamma_{k-1} \neq 0$, in every case $\beta_{k-1} \oplus \gamma_{k-1} \neq 0$.

If $h_{\beta_0 \oplus \gamma_0}^{B \cup C}(x_0) \downarrow$, then, by induction on $i < k$, using Theorem 13.1.18, we claim that for every $i < k$, we have $h_{\beta_i \oplus \gamma_i}^{B \cup C}(x_i) \downarrow$ and $h_{\beta_i \oplus \gamma_i}^{B \cup C}(x_i) \geq h_{\beta_{i+1} \oplus \gamma_{i+1}}^{B \cup C}(x_{i+1})$. Thus, if $h_{\beta_0 \oplus \gamma_0}^{B \cup C}(x_0) \downarrow$ then $h_{\beta_{k-1} \oplus \gamma_{k-1}}^{B \cup C}(x_{k-1}) \downarrow$, which is only possible if $\beta_{k-1} \oplus \gamma_{k-1} = 0$, contradiction. Thus $h_{\beta_0 \oplus \gamma_0}^{B \cup C}(x_0) \uparrow$ and $B \cup C$ is at most $(\beta \oplus \gamma)$ -large. $\square$

The following splitting property generalizes [72, Lemma 4.6]. Note that the reversal holds if $\beta \gg \alpha$.

Theorem 13.1.21 (Splitting property). *Let $X \subseteq \mathbb{N} \setminus \{0\}$ be $(\alpha \oplus \beta)$ -large for $\alpha, \beta < \mathfrak{e}_0$, then there exists $X_0 < X_1 \subseteq X$ such that X_0 is α -large and X_1 is β -large.*

Proof. Let X_0 be the prefix of X that is exactly α -large, i.e., if $X = \{x_0, \dots, x_{k-1}\}$, we let $X_0 = \{x_0, \dots, x_i\}$ for $i < k$ the smallest index such that $\{\alpha\}(x_0, \dots, x_i) = 0$. Such a set X_0 exists, as every coefficient in the Cantor normal form of $\alpha \oplus \beta$ is at least as big as the corresponding one in the Cantor normal form of α .

Let $X_1 = X \setminus X_0$, $\star = \max X + 1$ and let $X' = X \cup \{\star\}$. By the contrapositive of Main Theorem 13.0.1, since $X' = X_0 \cup (X_1 \cup \{\star\})$ is not at most $(\alpha \oplus \beta)$ -large, either X_0 is not at most α -large or $X_1 \cup \{\star\}$ is not at most β -large. Thus, $X_1 \cup \{\star\}$ is not at most β -large, and X_1 is therefore β -large. $\square$

We now deduce two versions of the pigeonhole principle.

Theorem 13.1.22 (First pigeonhole principle). *Let $X = X_0 \cup X_1 \subseteq \mathbb{N} \setminus \{0\}$ be $(\alpha \oplus \beta)$ -large for some $\alpha, \beta < \mathfrak{e}_0$, then either $X_0 \setminus \{\max X_0\}$ is α -large or $X_1 \setminus \{\max X_1\}$ is β -large or X_0 and X_1 are exactly α -large and exactly β -large, respectively.*

Proof. Suppose X_0 is at most α -large, otherwise $X_0 \setminus \{\max X_0\}$ is α -large and we are done. Similarly, suppose X_1 is at most β -large, otherwise $X_1 \setminus \{\max X_1\}$ is β -large and we are done. Let $\star = \max X + 1$. The set $X_0 \cup (X_1 \cup \{\star\})$ is not at most $(\alpha \oplus \beta)$ -large, hence, by the contrapositive of Main Theorem 13.0.1, $X_1 \cup \{\star\}$ is not at most β -large.

Since X_1 is at most β -large, it is exactly β -large. By applying the contrapositive of Main Theorem 13.0.1 on the partition $(X_0 \cup \{\star\}) \cup X_1$, we get that $X_0 \cup \{\star\}$ is not at most α -large, hence since X_0 is at most α -large, X_0 is exactly α -large and we are also done since both X_0 and X_1 are exactly β -large. $\square$

Theorem 13.1.23 (Second pigeonhole principle). *Let $X = X_0 \cup X_1 \cup \{\star\} \subseteq \mathbb{N} \setminus \{0\}$ be $(\alpha \oplus \beta)$ -large for some $\alpha, \beta < \mathfrak{e}_0$ and some $\star > X_0 \cup X_1$. Then either X_0 is α -large or X_1 is β -large.*

Proof. Let $\star' = \max X + 1$. The set $X \cup \{\star'\} = (X_0 \cup \{\star\}) \cup (X_1 \cup \{\star'\})$ is not at most $(\alpha \oplus \beta)$ -large, hence, by the contrapositive of Main Theorem 13.0.1, either $X_0 \cup \{\star\}$ is not at most α -large, in which case X_0 is α -large, or $X_1 \cup \{\star'\}$ is not at most β -large, in which case X_1 is β -large. $\square$

Since α -largeness is closed under supersets, if $X_0 \cup X_1$ is $(\alpha \oplus \beta)$ -large, then so is $X_0 \cup X_1 \cup \{\star\}$ for any $\star > \max X_1$, in which case either X_0 is α -large or X_1 is β -large. This formulation of the pigeonhole principle, although more natural, is less optimal than Theorem 13.1.23.

13.2 Largeness below ω^ω

We now return to the study of α -largeness for ordinals below ω^ω . Although, for our investigation of the first-order consequences of various theorems, we are mainly interested in ordinals of the form $\omega^n \cdot k$ (or even ω^n) for some $n, k \in \mathbb{N}$. The general definition of α -largeness will allow us to obtain tighter upper bounds that are of independent combinatorial interest.

13.2.1 Construction and deconstruction

Any ω^{n+1} -large set X can be seen as an increasing sequence of $\min X$ many ω^n -large blocks. In this section, we prove two important propositions: a *construction* property, which quantifies how many ω^n -large blocks are sufficient to produce an ω^{n+m} -large set, and a *deconstruction* property, which counts the number of ω^n -large blocks which can be extracted from an ω^{n+m} -large set.

The quantification is expressed in terms of α -largeness. There are multiple ways to give a meaning to the sentence “the sequence of blocks $X_0 < \dots < X_{d-1}$ is α -large”. One could ask that $\{\max X_i : i < d\}$ is α -large, or that any sequence in $X_0 \times \dots \times X_{d-1}$ is α -large. Thanks to the regularity of α -largeness (see Proposition 13.1.7), the two definitions are equivalent. However, this is no longer the case with largeness(θ). In the next chapter, we prove in Lemma 14.0.3 the decomposition property for largeness(θ) in the strongest setting, where any sequence in $X_0 \times \dots \times X_{d-1}$ is required to be α -large(θ).

Proposition 13.2.1 (Construction). *For every $n, m, d \in \mathbb{N}$ and for every sequence of ω^n -large sets $X_0 < \dots < X_{d-1}$ such that $\{\max X_i : i < d\}$ is ω^m -large, then $X_0 \cup \dots \cup X_{d-1}$ is ω^{n+m} -large.*

Proof. By induction on m . The case $m = 0$ is immediate, since any ω^0 -large set is non-empty, so $d \geq 1$ and X_0 is ω^n -large. Suppose it holds for m . If $Y = \{\max X_i : i < d\}$ is ω^{m+1} -large, then $Y \setminus \{\min Y\}$ is $\omega^m \cdot \min Y$ -large. Let $Y_0 < \dots < Y_{\min Y - 1}$ be ω^m -large subsets of $Y \setminus \{\min Y\}$. For every $i < \min Y$, let $Z_i = \bigcup_{\max X_j \in Y_i} X_j$. By induction hypothesis, each Z_i is ω^{n+m} -large, so $\{\min Y\} \cup Z_0 \cup \dots \cup Z_{\min Y - 1}$ is ω^{n+m+1} -large. By upward-closure of α -largeness, $X_0 \cup \dots \cup X_{d-1}$ is ω^{n+m+1} -large. $\square$

The following deconstruction proposition is an adaptation of Kołodziejczyk and Yokoyama [79, Lemma 2.1]. Note that the hypothesis requires an extra ω^n -large set,

while the conclusion also contains some extra ω^n -large set. This might seem sub-optimal but the ω^n -large set of the hypothesis is on the left part of X , while the one from the conclusion is on the right, which is weaker.

Proposition 13.2.2 (Deconstruction). *For every $n, m \in \mathbb{N}$ and for every $\omega^{n+m} + \omega^n$ -large set X , there are some $d \in \mathbb{N}$ and some ω^n -large subsets $X_0 < \dots < X_d$ of X such that $\{\max X_i : i < d\}$ is ω^m -large.*

Proof. By induction on m . The case $m = 0$ is immediate by Ketonen and Solovay [72, Lemma 4.6], as X is $\omega^n \cdot 2$ -large, so there are two ω^n -large sets $X_0 < X_1 \subseteq X$. Note that $\{\max X_0\}$ is non-empty, hence ω^0 -large.

Suppose it holds for m . Let X be $\omega^{n+m+1} + \omega^n$ -large. By Ketonen and Solovay [72, Lemma 4.6] (or by Theorem 13.1.21), there are some $A < B \subseteq X$ such that A is ω^n -large and B is ω^{n+m+1} -large. Let $Z_0 < \dots < Z_{\min B - 1}$ be ω^{n+m} -large subsets of $B \setminus \{\min B\}$.

Let $d_{-1} = 0$ and $X_{d_{-1}}^{-1} = Z_0$. Note that $X_{d_{-1}}^{-1}$ is ω^{n+m} -large, hence ω^n -large. We are going to build inductively for every $i < \min B - 1$ a family of ω^n -large sets $X_0^i < \dots < X_{d_i}^i$ such that

- (1) $X_0^i \cup \dots \cup X_{d_i}^i \subseteq X_{d_{i-1}}^{i-1} \cup Z_i$;
- (2) $\{\max X_j^i : j < d_i\}$ is ω^m -large.

Assuming $X_{d_{i-1}}^{i-1} < Z_i$ is defined, the set $X_{d_{i-1}}^{i-1} \cup Z_i$ is $\omega^{n+m} + \omega^n$ -large, so by induction hypothesis, there is a sequence $X_0^i, \dots, X_{d_i}^i$ of ω^n -large sets satisfying (1) and (2). It follows that $\{\max X_j^i : i < \min B - 1, j < d_i\}$ is $\omega^m \cdot (\min B - 1)$ -large. Since $\max A < \min B$, then $\{\max A\} \cup \{\max X_j^i : i < \min B - 1, j < d_i\}$ is ω^{m+1} -large, so altogether, these ω^n -large sets form the desired sequence. $\square$

Corollary 13.2.3 (Generalized deconstruction). *For every $n, k, \ell \geq 1$ and $m \geq 0$, for every $\omega^{n+m} \cdot (k\ell + 1)$ -large set X such that $k(\ell + 1) \leq \min X$, there are some $d \in \mathbb{N}$ and some $\omega^n \cdot k$ -large subsets $X_0 < \dots < X_{d-1}$ of X such that $\{\max X_i : i < d\}$ is $\omega^m \cdot \ell$ -large.*

Proof. Suppose $m = 0$. Then X is $\omega^n \cdot (k\ell + 1)$ -large, and a fortiori $\omega^n \cdot k\ell$ -large hence there is a sequence $X_0 < \dots < X_{k\ell-1}$ of ω^n -large subsets of X . For every $i < \ell$, let $Z_i = X_{ki} \cup \dots \cup X_{k(i+1)-1}$. Each Z_i is $\omega^n \cdot k$ -large, and $\{\max Z_i : i < \ell\}$ is ℓ -large.

Suppose $m > 0$. In particular, X is $\omega^{n+m} \cdot k\ell + \omega^{n+1}$ -large. Since $k(\ell + 1) \leq \min X$, then X is $\omega^{n+m} \cdot k\ell + \omega^n \cdot k(\ell + 1)$ -large. By Theorem 13.1.21, there is a sequence $X_{-1} < X_0 < \dots < X_{k\ell-1}$ of subsets of X such that X_{-1} is $\omega^n \cdot k$ -large and $X_0, \dots, X_{k\ell-1}$ are $\omega^{n+m} + \omega^n$ -large. By Proposition 13.2.2, for each i such that $0 \leq i < k\ell$, there is a sequence $X_i^0 < \dots < X_i^{d_i-1}$ of ω^n -large subsets of X_i such that $\{\max X_i^j : j < d_i\}$ is ω^m -large. Let $W_0, \dots, W_{p-1}$ be the sequence $(X_i^j)_{i < k\ell, j < d_i}$. In particular, the set $Z = \{\max W_i : i < p\}$ is $\omega^m \cdot k\ell$ -large. Let $g : Z \rightarrow k$ be such that $g(x)$ is the remainder of the Euclidean division of i by k , where $x = \max W_i$. By Theorem 13.1.23, there is an $\omega^m \cdot \ell$ -large g -homogeneous subset $H = \{x_0 < \dots < x_{d-1}\} \subseteq Z$. For each $s < d$, let i_s be such that $\max W_{i_s} = x_s$. Let $H_0 = X_{-1} \cup W_{i_0}$ and $H_{s+1} = \bigcup_{i_s < j \leq i_{s+1}} W_j$. Note that for each $s < d$, $\max H_s = x_s$, so $\{\max H_s : s < d\}$ is $\omega^m \cdot \ell$ -large. Since X_{-1} is $\omega^n \cdot k$ -large, so is H_0 . By g -homogeneity of H , $i_{s+1} \geq i_s + k$, so H_{s+1} is $\omega^n \cdot k$ -large. $\square$

13.2.2 Sparsity

Every notion of largeness $\mathcal{L}$ induces a dual notion of $\mathcal{L}$ -sparsity as follows:

Definition 13.2.4. Let $\mathcal{L}$ be a largeness notion, a set $X \subseteq \mathbb{N}$ is $\mathcal{L}$ -sparse if for every $x, y \in X$ such that $x < y$, $(x, y]_{\mathbb{N}} \in \mathcal{L}$.

In the special case of α -largeness, we will talk about α -sparse sets.

By closure of largeness notions under supersets, any subset of an $\mathcal{L}$ -sparse set is again $\mathcal{L}$ -sparse. Any increasing function $g : \mathbb{N} \rightarrow \mathbb{N}$ induces a notion of largeness $\mathcal{L}_g = \{F : g(\min F - 1) < \max F\}$. Then, an $\mathcal{L}_g$ -sparse set X is such that for $x, y \in X$ with $x < y$, $g(x) < y$. We simply say that X is g -sparse. We will be particularly interested in g -sparsity for primitive recursive functions g . Indeed, by the following theorem, g -sparse sets for primitive recursive functions g are closely linked to α -sparse sets for $\alpha < \omega^\omega$.

Theorem 13.2.5 (Ketonen and Solovay [72]). *For every primitive recursive function $g : \mathbb{N} \rightarrow \mathbb{N}$, there is some $n \in \mathbb{N}$ such that for every ω^n -large set F , $\max F > g(\min F)$. Thus, every ω^n -sparse set is also g -sparse.*

In particular, $y > 2x$ whenever $(x, y]$ is ω -large and $y > x2^x$ whenever $(x, y]$ is ω^2 -large. We shall be particularly interested in $\omega^2 \cdot 3$ -sparse sets as they provide sufficient sparsity to bound Ramsey numbers.

Recall that we have the following upper bounds on $R_k(m)$, the least number R such that for every coloring $f : [R]^2 \rightarrow k$, there is an f -homogeneous set of size m . By Erdős and Szekeres [42], $R_2(m) \leq \binom{2m-1}{m-1} \leq 4^m$, and this bound generalizes to $R_k(m) \leq k^{km}$ (see Graham, Rothschild and Spencer [53, Section 1.1]).

Lemma 13.2.6. *If X is $\omega^2 \cdot 3$ -sparse and $\min X \geq 7$, then it is $(x \mapsto x^{R_x(2x+2)})$ -sparse.*

Proof. Any ω^2 -sparse set is $(x \mapsto x2^x)$ -sparse, so any $\omega^2 \cdot 3$ -sparse set is $(x \mapsto x2^{x+x2^x+x2^{x2^x}})$ -sparse, and in particular $(x \mapsto 2^{x2^{x2^x}})$ -sparse. Whenever $x \geq 7$, then $x(2x+2) \leq 2^x$, so $2^{x2^{x2^x}} \geq x^{x^2(2x+2)} \geq x^{x^{x(2x+2)}}$. Thus X is $(x \mapsto x^{x^{x(2x+2)}})$ -sparse.

Since $R_x(2x+2) \leq x^{x(2x+2)}$ (see Graham, Rothschild and Spencer [53, Section 1.1]), $x^{R_x(2x+2)} \leq x^{x^{x(2x+2)}}$, so X is $(x \mapsto x^{R_x(2x+2)})$ -sparse. $\square$

Thanks to the deconstruction proposition Proposition 13.2.2, one can relate largeness to sparsity.

Proposition 13.2.7. *For every primitive recursive and increasing function $g : \mathbb{N} \rightarrow \mathbb{N}$, there exists some bound $n \in \mathbb{N}$ such that, for every $m \in \mathbb{N}$, every ω^{n+m} -large set X with $\min X \geq 2$ contains an ω^m -large and g -sparse subset.*

Proof. By Theorem 13.2.5, there exists some $n' \in \mathbb{N}$ such that every $\omega^{n'}$ -sparse set is g -sparse. Let $n = n' + 1$ and let $X \subseteq_{\text{fin}} \mathbb{N}$ be ω^{n+m} -large. As $\min X \geq 2$, X is also $\omega^{n'+m} \cdot 2$ -large and therefore $\omega^{n'+m} + \omega^{n'}$ -large, hence, by Proposition 13.2.2 there are some $d \in \mathbb{N}$ and some $\omega^{n'}$ -large subsets $X_0 < \dots < X_d$ of X such that $Y := \{\max X_i : i < d\}$ is ω^m -large.

We claim that Y is $\omega^{n'}$ -sparse and therefore g -sparse. Indeed, let $x < y \in Y$, there exists some $i < d$ such that $y = \max X_i$, hence $X_i \subseteq (x, y]$ and $(x, y]$ is thus $\omega^{n'}$ -large. $\square$

The following lemma specifies this bound in the case of $\omega^2 \cdot 3$ -sparsity.

Lemma 13.2.8. *For every $\ell \geq 1$ and $m \geq 0$, for every ω^{m+3} -large set X such that $3\ell + 2 \leq \min X$, there is some $\omega^2 \cdot 3$ -sparse $\omega^m \cdot \ell + 1$ -large subset of X .*

Proof. Since $3\ell + 2 \leq \min X$, then $X \setminus \{\min X\}$ is $\omega^{m+2} \cdot (3\ell + 2)$ -large, hence $\omega^{m+2} \cdot (3\ell + 1) + \omega^{m+2}$ -large. Then there are two subsets $A < B \subseteq X$ such that A is ω^{m+2} -large and B is $\omega^{m+2} \cdot (3\ell + 1)$ -large. Since $3\ell + 2 \leq \min A$, then $3(\ell + 1) \leq \max B$, so by Corollary 13.2.3 applied to B , there is a sequence $X_0 < \dots < X_{d-1}$ of $\omega^2 \cdot 3$ -large subsets of B such that $Z = \{\max X_i : i < d\}$ is $\omega^m \cdot \ell$ -large. The set $\{\min X\} \cup Z$ is $\omega^m \cdot \ell + 1$ -large, and $\omega^2 \cdot 3$ -sparse. $\square$

Although sparsity can be translated into some largeness assumption, it is convenient to separate the largeness and sparsity hypothesis in the ordinal computation of largeness, as sparsity has a better behavior with respect to composition of RT-like statements and yields better bounds.

13.2.3 Largeness for combinatorial theorems

Bounds for Ramsey's theorem for pairs were extensively studied. Ketonen and Solovay [72] proved that every ω^{k+4} -large set is RT_k^2 - ω -large when $k \geq 2$. Patey and Yokoyama [115] proved that for every $k \in \mathbb{N}$, there is some $n \in \mathbb{N}$ such that every ω^n -large set is RT_k^2 - ω^n -large. Kołodziejczyk and Yokoyama [79] computed some explicit bounds, and proved that every $\omega^{144(n+1)}$ -large set is RT_k^2 - ω^n -large. Our goal is to give a tight bound to RT_k^2 - ω^n -largeness, up to an additive constant.

To do so, we will again use the decomposition of Bovykin and Weiermann [11] of Ramsey's theorem for pairs and two colors into $\text{EM} + \text{ADS}$. More precisely, the bounds obtained for EM - ω^n -largeness actually apply to a stronger statement called fEM_k , formally introduced by Towsner and Yokoyama [139], but already used in Patey [111, Theorem 8.2.1]:

Definition 13.2.9. Given a coloring $f : [\mathbb{N}]^2 \rightarrow k$, a set $H \subseteq \mathbb{N}$ is f -fallow if for every $x < y < z \in H$, $f(x, z) \in \{f(x, y), f(y, z)\}$.

Clearly, any f -fallow set is f -transitive, but the converse is not true. Both EM and ADS therefore admit two generalized statements, based on transitivity and fallowness. We shall only consider the stronger version of each statement, that is, the fallow version of the Erdős-Moser theorem and transitive Ramsey theorem for pairs.

Statement 13.2.10 (Fallow Erdős-Moser theorem). Given $k \in \mathbb{N}$, fEM_k is the statement “For every coloring $f : [\mathbb{N}]^2 \rightarrow k$, there is an infinite f -fallow set”. And fEM is the statement $\forall k \text{fEM}_k$.

Statement 13.2.11 (Transitive Ramsey theorem for pairs). Given $k \in \mathbb{N}$, trRT_k^2 is the statement “For every transitive coloring $f : [\mathbb{N}]^2 \rightarrow k$, there is an infinite f -homogeneous set”.

The statements fEM_k and trRT_k^2 admit largeness counterparts, as does Ramsey's theorem. In the remainder of this chapter, we shall compute bounds for fEM - ω^n -largeness and trRT_k^2 - ω^n -largeness, and deduce bounds for RT_k^2 - ω^n -largeness.

13.2.4 Pigeonhole principle

We already studied α -large versions of the pigeonhole principle for $\alpha < \epsilon_0$ in Section 13.1. The following proposition is an immediate consequence of Theorem 13.1.23, restricted to the setting of ω^n -largeness. It is however useful to state it under this form as it will be used all over the remainder of the chapter. This strengthens Kołodziejczyk and Yokoyama [79, Lemma 2.2] by improving the bounds and removing the sparsity assumptions.

Proposition 12.2.7. Fix $n \geq 0$ and $a, k \geq 1$.

- (1) If X is $\omega^n \cdot ak$ -large, then it is RT_k^1 - $\omega^n \cdot a$ -large.
- (2) If X is ω^{n+1} -large, then it is RT^1 - ω^n -large.

Proof. (1) By induction on k . The case $k = 1$ is trivial. Suppose it holds for k . Let X be $\omega^n \cdot a(k+1)$ -large, and let $f : X \rightarrow k+1$. Let $Y_0 = \{x \in X : f(x) < k\}$ and $Y_1 = \{x \in X : f(x) = k\}$. By Theorem 13.1.23, either Y_0 is $\omega^n \cdot ak$ -large, in which case by induction hypothesis, there is an f -homogeneous set which is $\omega^n \cdot a$ -large, or Y_1 is $\omega^n \cdot a$ -large, and is f -homogeneous for color k by definition.

(2) Let $f : X \rightarrow \min X$. Since X is ω^{n+1} -large, $X \setminus \{\min X\}$ is $\omega^n \cdot \min X$ -large. By (1), $X \setminus \{\min X\}$ is $\text{RT}_{\min X}^1$ - ω^n -large, so there is an ω^n -large f -homogeneous subset. $\square$

13.2.5 Grouping principle

The inductive proofs of Kołodziejczyk and Yokoyama for largeness for EM heavily depended on the notion of grouping. We shall only use the following lemma, which, in the terminology of grouping, means that every $\omega^n \cdot R_k(2d)$ -large $(x \mapsto x2^x)$ -sparse set admits an $(\omega^n + 1, d)$ -grouping. This decomposed their ω^{n+3} bound [79, Lemma 2.6] into a largeness and a sparsity assumption.

Lemma 13.2.12. *Let $k > 0$ and let $X \subseteq_{\text{fin}} \mathbb{N}$ be $\omega^n \cdot R_k(2d)$ -large and $(x \mapsto xk^x)$ -sparse, for some $d \in \mathbb{N}$ such that $k^{R_k(2d)} \leq \min X$ and let $f : [X]^2 \rightarrow k$ be a coloring. Then, there exists a sequence $Y_0 < \dots < Y_{d-1}$ of $\omega^n + 1$ -large subsets of X and some color $c < k$ such that $f(x, y) = c$ for every $x \in Y_s$ and $y \in Y_t$ for $s < t < d$.*

Proof. If $n = 0$, the result holds by definition of R and ω^0 -largeness. Indeed, let $f : [X]^2 \rightarrow k$ be a coloring. Since $\text{card } X \geq R_k(2d)$, by finite Ramsey theorem for pairs, there is an f -homogeneous set $\{y_0 < y_1 < \dots < y_{2d-1}\} \subseteq X$. Since any singleton element is ω^0 -large, the set $Y_i = \{y_{2i}, y_{2i+1}\}$ is $\omega^0 + 1$ -large. Assume from now on that $n > 0$.

Let $X_0 < \dots < X_{R_k(2d)-1}$ be a decomposition of X into ω^n -large sets. Consider the sequence $(f_i)_{i < R_k(2d)}$ of colorings and the sequence $Z_0 < \dots < Z_{R_k(2d)-1}$ of sets defined inductively as follows:

Assume that $Z_{i+1}, \dots, Z_{R_k(2d)-1}$ have been defined for some $i < R_k(2d)$, and consider the coloring $f_i : X_i \rightarrow k^{R_k(2d)-1-i+\sum_{j<i} \text{card } X_j}$ defined by

$$f_i(x) = (f(y, x))_{y \in \bigcup_{j<i} X_j} \sqcup (f(x, y))_{y \in \{\min Z_j : i < j < R_k(2d)\}}$$

As $\min X_0 \geq R_k(2d)$ and as the X_i are disjoint, we get that for $i > 0$, $R_k(2d) + \sum_{j<i} \text{card } X_j \leq \max X_{i-1}$. By $(x \mapsto xk^x)$ -sparsity of X , $k^{\max X_{i-1}} \cdot \max X_{i-1} \leq \min X_i$, so $k^{R_k(2d)+\sum_{j<i} \text{card } X_j} \cdot \max X_{i-1} \leq \min X_i$. By Proposition 12.2.7(1), there exists some $\omega^{n-1} \cdot (\max X_{i-1})$ -large and f_i -homogeneous subset $Z_i \subseteq X_i$. For $i = 0$, since $k^{R_k(2d)} \leq \min X_0$, by Proposition 12.2.7(1), there exists some ω^{n-1} -large and f_0 -homogeneous subset $Z_0 \subseteq X_0$.

The $(Z_i)_{i < R_k(2d)}$ are defined such that, for every $i < j < R_k(2d)$ and $x \in Z_i$, $y \in Z_j$, we have $f(x, y) = f(\min Z_i, \min Z_j)$. By definition of $R_k(2d)$, there exists some subset $\{i_0 < \dots < i_{2d-1}\} \subseteq \{0, \dots, R_k(2d) - 1\}$ and some color $c < k$ such that, for every $j < \ell < 2d$, $x \in Z_{i_j}$ and $y \in Z_{i_\ell}$, $f(x, y) = c$.

Finally, let $Y_j = \{z_{2j}^0, z_{2j}^1\} \cup Z_{i_{2j+1}}$ for every $j < d$, where z_{2j}^0 and z_{2j}^1 are the two last elements of $Z_{i_{2j}}$. Every Y_j is $\omega^n + 1$ -large. $\square$

13.2.6 Erdős-Moser theorem

The goal of this section is to prove that every ω^{n+3} -large set X with $\min X \geq 5$ is fEM- ω^n -large (and thus EM- ω^n -large). The previous known bound of Kołodziejczyk and Yokoyama [79] was ω^{36n+4} -largeness.

Proposition 13.2.13. *Let $n \geq 0$ and $k, d \geq 2$, let $X \subseteq_{\text{fin}} \mathbb{N}$ be $(x \mapsto x^{R_x(2x)})$ -sparse,*

- (1) *If X is $\omega^n + 1$ -large with $2 \leq \min X$, then it is fEM- ω^n -large.*
- (2) *If X is $\omega^n \cdot R_k(2d)$ -large with $k^{R_k(2d)} \leq \min X$, then it is fEM $_k$ - $\omega^n \cdot d$ -large.*
- (3) *If X is $\omega^n \cdot 2k$ -large, then it is fEM $_k$ - $\omega^n + 1$ -large.*

Proof. We prove (1) and (2) by mutual induction on n . (1) holds trivially for $n = 0$, as any ω^0 -large set is non-empty, and any singleton is $\mathbf{fEM}$ - ω^0 -large.

Suppose (1) holds for n . Let us show that (2) holds for n . Let $k, d \in \mathbb{N}$, let $X \subseteq_{\mathbf{fin}} \mathbb{N}$ be $(x \mapsto x^{R_x(2x)})$ -sparse, $(\omega^n \cdot R_k(2d))$ -large and such that $k^{R_k(2d)} \leq \min X$ and fix some coloring $f : [X]^2 \rightarrow k$. As $k^{R_k(2d)} \leq \min X$, we have $k \leq \min X$, hence X is $(x \mapsto x^{k^x})$ -sparse, and by Lemma 13.2.12, there exists some sequence $Z_0 < \dots < Z_{d-1}$ of $\omega^n + 1$ -large subsets of X and some color $c < k$ such that $f(x, y) = c$ for every $x \in Z_s$ and $y \in Z_t$ for $s < t < d$. By (1) for n , for each $i < d$, there is an f -fallow ω^n -large subset $W_i \subseteq Z_i$. Then, the set $W = \bigcup_{i < d} W_i$ is $\omega^n \cdot d$ -large and f -fallow.

Indeed, pick by contradiction some $x < y < z \in W$ such that $f(x, z) \notin \{f(x, y), f(y, z)\}$. We cannot have $Z_i \cap \{x, y, z\} = \{x\}$ or $Z_i \cap \{x, y, z\} = \{z\}$ for some $i < d$ as it would imply $f(x, y) = f(x, z) = c$ or $f(x, z) = f(y, z) = c$ by our application of Lemma 13.2.12. Similarly, we cannot have $Z_i \cap \{x, y, z\} = \{y\}$ for some $i < d$ as it implies the existence of some $j < d$ such that $Z_j \cap \{x, y, z\} = \{z\}$, since $x < y < z$. And finally, we cannot have $Z_i \cap \{x, y, z\} = \{x, y, z\}$ for some $i < d$ by our application of the inductive hypothesis.

Suppose (2) holds for n . Let us show (1) holds for $n + 1$. Let $X \subseteq_{\mathbf{fin}} \mathbb{N}$ be $(x \mapsto x^{R_x(2x)})$ -sparse and $(\omega^{n+1} + 1)$ -large and let $x_0 = \min X$. Fix some coloring $f : [X]^2 \rightarrow x_0$, then, $X \setminus \{x_0\}$ is $\omega^n \cdot x_0 R_{x_0}(2x_0)$ -large as $x_0 R_{x_0}(2x_0) \leq x_0^{R_{x_0}(2x_0)}$ when $x_0 \geq 2$ and $x_0^{R_{x_0}(2x_0)} \leq \min X \setminus \{x_0\}$. By Proposition 12.2.7(1), there exists some $\omega^n \cdot R_{x_0}(2x_0)$ -large subset $Y \subseteq X \setminus \{x_0\}$ such that $f(x_0, y) = f(x_0, z)$ for every $y, z \in Y$. By (2) for n , there is some f -fallow $\omega^n \cdot x_0$ -large subset $Z \subseteq Y$. Then, the set $W = \{x_0\} \cup Z$ is f -fallow and ω^{n+1} -large.

Indeed, pick by contradiction some $x < y < z \in W$ such that $f(x, z) \notin \{f(x, y), f(y, z)\}$. We cannot have $x = x_0$ as $f(x_0, y) = f(x_0, z)$ for every $y, z \in Y$. And we cannot have $x \in Z$, otherwise $\{x, y, z\} \subseteq Z$, contradicting the fact that Z is f -fallow.

(3) For $n = 0$, if X is $\omega^0 \cdot 2k$ -large, then $\text{card } X \geq 2k$. In particular, any two-element subset of X is $\mathbf{fEM}$ - $\omega^0 + 1$ -large. Suppose $n > 0$. Fix a coloring $f : [\mathbb{N}]^2 \rightarrow k$. Let $x_0 = \min X$ and $g : X \rightarrow k$ be defined by $g(x_0) = 0$ and $g(y) = f(x_0, y)$ for $y \neq x_0$. By Proposition 12.2.7(1), there is an $\omega^n \cdot 2$ -large g -homogeneous subset $Y \subseteq X$. Let $Y_0 < Y_1$ be its decomposition into ω^n -large sets. Let $x_1 = \max Y_0$. Since $n > 0$, $\text{card } Y_0 \geq 2$, so $x_1 > x_0$ and $\{x_1\} \cup Y_1$ is $\omega^n + 1$ -large. By (1), there is a ω^n -large f -fallow subset $Z \subseteq \{x_1\} \cup Y_1$. Since Z is g -homogeneous, then $\{x_0\} \cup Z$ is f -fallow, and $\omega^n + 1$ -large. $\square$

Corollary 13.2.14. Let $n \geq 0$. If $X \subseteq_{\mathbf{fin}} \mathbb{N}$ is ω^{n+3} -large and $\min X \geq 7$, then it is $\mathbf{fEM}$ - ω^n -large.

Proof. Since $\min X \geq 5$, then by Lemma 13.2.8, there is an $\omega^2 \cdot 3$ -sparse $\omega^n + 1$ -large subset $Y \subseteq X$. By Lemma 13.2.6, since $\min X \geq 7$, Y is $(x \mapsto x^{R_x(2x+2)})$ -sparse, so by Proposition 13.2.13, Y is $\mathbf{fEM}$ - ω^n -large. $\square$

13.2.7 Transitive colorings

The strategy to compute bounds for transitive Ramsey theorem for pairs and k -colors is different from the one for fallow Erdős-Moser theorem. Patey and Yokoyama [115, Lemma 4.4] proved that every ω^{2n+6} -large set is $\mathbf{ADS}$ - ω^n -large, using a slightly modified notion of largeness. The proof consisted in defining a coloring $f : [\mathbb{N}]^2 \rightarrow 2n + 2$ and applying Ketonen and Solovay's bound [72]. Kołodziejczyk and Yokoyama [79, Theorem 2.11] gave a more direct but less optimal proof by getting rid of the intermediate notion of largeness, and showed that any ω^{4n+4} -large set X with $\min X \geq 3$ is trRT_2^2 - ω^n -large.

In this section, we give a direct and optimal proof, by generalizing the bound of Ketonen and Solovay to ω - g -largeness. Given an increasing function $g : \mathbb{N} \rightarrow \mathbb{N}$, a set X is ω - g -large if $\text{card } X > g(\min X)$. Whenever g is the identity function, this yields the standard notion of ω -largeness. The following proposition is a generalization of Ketonen and Solovay [72, Lemma 6.3] and of Kołodziejczyk and Yokoyama [79, Theorem 1.5]:

Proposition 13.2.15. *Fix a primitive recursive function $g : \mathbb{N} \rightarrow \mathbb{N}$. If X is $(\omega^k + 1)$ -large and g -sparse for $k = 1$ and $(x \mapsto k^{R_k(2g(x))})$ -sparse for $k \geq 2$, then it is RT_k^2 - ω - g -large.*

Proof. By induction on k . Suppose $k = 1$. Let $f : [X]^2 \rightarrow 1$ be the trivial constant coloring. Then X is f -homogeneous. Since X is $\omega^k + 1$ -large, then, letting $X = \{x_0 < x_1\} \sqcup Y$, then $\text{card } Y \geq x_1$ and by g -sparsity of X , $x_1 > g(x_0)$, so $\text{card } X > g(\min X)$.

Assume the property to be true for some $k > 0$. Let X be $(\omega^{k+1} + 1)$ -large and $(x \mapsto (k+1)^{R_{k+1}(2g(x))})$ -sparse and consider a coloring $f : [X]^2 \rightarrow k+1$. Let $x_0 = \min X$, by sparsity, $\min(X \setminus \{x_0\}) \geq (k+1)^{R_{k+1}(2g(x_0))}$ and $X \setminus \{x_0\}$ is an $(\omega^k \cdot (k+1) \times R_{k+1}(2g(x_0)))$ -large set.

Consider the coloring $h : X \setminus \{x_0\} \rightarrow k+1$ defined by $h(y) = f(x_0, y)$. By Proposition 12.2.7(1), there exists some $(\omega^k \cdot R_{k+1}(2g(x_0)))$ -large and h -homogeneous subset Y of $X \setminus \{x_0\}$, and let $c_0 \leq k$ be the corresponding color.

By Lemma 13.2.12, as $\min Y \geq (k+1)^{R_{k+1}(2g(x_0))}$, there exists a sequence $Y_0 < \dots < Y_{g(x_0)-1}$ of $\omega^k + 1$ -large subsets of Y and some color c_1 such that $f(x, y) = c_1$ for every $x \in Y_s$ and $y \in Y_t$ for $s < t < g(x_0)$.

If $c_0 = c_1$, then the set $\{x_0\} \cup \{\min Y_0 < \dots < \min Y_{g(x_0)-1}\}$ is f -homogeneous and ω - g -large and we are done. So assume $c_0 \neq c_1$ and consider the coloring $f' : [X]^2 \rightarrow k$ to be the same coloring as f , except that the colors c_0 and c_1 are fused into one. By the inductive hypothesis, there exists some ω - g -large f' -homogeneous subset $Z_1 \subseteq Y_1$.

If Z_1 is f homogeneous, then we are done. Otherwise, $f([Z_1]^2) = \{c_0, c_1\}$, and, since $\min Z_1 > R_2(g(\min Y_0))$ by sparsity of X , there exists some f -homogeneous subset $W \subseteq Z_1$ of size $\geq g(\min Y_0)$. If $f([W]^2) = \{c_0\}$, then $\{x_0\} \cup W$ is f -homogeneous, and if $f([W]^2) = \{c_1\}$, then $\{\min Y_0\} \cup W$ is f -homogeneous, and both of these sets are ω - g -large. $\square$

By generalizing the proof of Patey and Yokoyama [115, Lemma 4.4] to trRT_k^2 and replacing the bound of Ketonen and Solovay by Proposition 13.2.15, we obtain the following proposition:

Proposition 13.2.16. *Let $n \in \mathbb{N}$, $k \geq 1$ and let $X \subseteq \mathbb{N}$ be $(\omega^{kn} + 1)$ -large and $(x \mapsto (kn)^{R_{kn}(2x+2)})$ -sparse, then it is trRT_k^2 - ω^n -large.*

Proof. Consider a transitive coloring $f : [X]^2 \rightarrow k$ and let $\bar{f} : [X]^2 \rightarrow kn$ be defined by $\bar{f}(x, y) = kj + i$ if $f(x, y) = i$ and if j is the smallest index such that there is no ω^{j+1} -large subset $H \subseteq [x, y] \cap X$ with $x \in H$ and such that $H \cup \{y\}$ is f -homogeneous for color i , or, if no such index exists, take $j = n - 1$.

By Proposition 13.2.15, take $Y \subseteq X$ such that Y is ω -($x \mapsto x + 1$)-large and $\bar{f}$ -homogeneous. Write $Y = \{y_0 < y_1 < \dots < y_\ell\}$ for some $\ell > y_0$ and let $\bar{f}([Y]^2) = kj + i$. For every $s < \ell$, let $H_s \subseteq [y_s, y_{s+1})$ be ω^j -large, f -homogeneous for the color i and such that $y_s \in H_s$. Let $H = \{y_0\} \cup \bigcup_{s=1}^{\ell-1} H_s$, then, as $\ell > y_0$, H is ω^{j+1} -large and $H \cup \{y_\ell\}$ is f -homogeneous for the color i . Thus, we have $j = n$ as $\bar{f}(y_0, y_\ell) = nj + i$, hence H is f -homogeneous and ω^n -large. $\square$

Corollary 13.2.17. *Let $n \in \mathbb{N}$, $k \geq 1$. If $X \subseteq_{\text{fin}} \mathbb{N}$ is ω^{n^k+3} -large and $\min X \geq 8$, then it is trRT_k^2 - ω^n -large.*

Proof. Since $\min X \geq 8$, then by Lemma 13.2.8, there is an $\omega^2 \cdot 3$ -sparse $\omega^{nk} \cdot 2 + 1$ -large subset $Y \subseteq X$. By regularity, there are two subsets $Y_0 < Y_1$ of Y such that Y_0 is ω^{nk} -large and Y_1 is $\omega^{nk} + 1$ -large.

By Lemma 13.2.6, since Y_1 is $\omega^2 \cdot 3$ -sparse and $\min X \geq 7$, then it is $(x \mapsto x^{R_x(2x+2)})$ -sparse. Since Y_0 is ω^{nk} -large, then $\text{card } Y_0 \geq nk$, so $\min Y_1 \geq nk$. It follows that Y_1 is $(x \mapsto (kn)^{R_{kn}(2x+2)})$ -sparse, so by Proposition 13.2.16, Y_1 is trRT_k^2 - ω^n -large. $\square$

Corollary 13.2.18. Let $n \in \mathbb{N}$. If $X \subseteq_{\text{fin}} \mathbb{N}$ is ω^{2n+3} -large and $\min X \geq 8$, then it is $\text{ADS-}\omega^n$ -large.

Proof. Immediate by Corollary 13.2.17. $\square$

Note that this bound is optimal, up to an additive constant, in that the coloring witnessing the lower bound of RT_2^2 by Kotlarski et al. [82, Theorem 5.4] is transitive.

13.2.8 Ramsey's theorem for pairs

We now have all the necessary ingredients to prove our optimal bounds for RT_k^2 - ω^n -largeness, using the bounds for the fallow Erdős-Moser theorem and for transitive Ramsey theorem for pairs. We first state it with largeness and sparsity assumptions, and then use Section 13.2.2 to compute purely largeness bounds.

Theorem 13.2.19. Let $n, k \geq 1$. If $X \subseteq_{\text{fin}} \mathbb{N}$ is $\omega^{kn} \cdot 2k$ -large and $(x \mapsto x^{R_x(2x+2)})$ -sparse with $kn \leq \min X$, then it is RT_k^2 - ω^n -large.

Proof. Let $f : [X]^2 \rightarrow k$ be a coloring. X is $\omega^{kn} \cdot 2k$ -large and $(x \mapsto x^{R_x(2x+2)})$ -sparse, hence, by Proposition 13.2.13(3), there exists some $\omega^{kn} + 1$ -large and f -fallow subset $Y \subseteq X$. In particular, Y is f -transitive.

As $kn \leq \min X$, Y is $(x \mapsto (kn)^{R_{kn}(2x+2)})$ -sparse, and by Proposition 13.2.16 there exists some ω^n -large f -homogeneous subset $Z \subseteq Y$. $\square$

Note that if we applied fEM and trRT_k^2 under their explicit largeness assumptions (Corollaries 13.2.14 and 13.2.17) instead of their hybrid sparsity/largeness versions (Propositions 13.2.13 and 13.2.16), this would have yielded significantly worse bounds. It is therefore useful to keep a sparsity version of the statements in case of their use as an intermediate object in the proof of a stronger combinatorial statement. When translating the sparsity assumptions into largeness ones, we obtain our main theorem.

Main Theorem 13.0.2. Let $n, k \geq 1$. If $X \subseteq_{\text{fin}} \mathbb{N}$ is ω^{kn+4} -large and $\min X \geq 2$, then it is RT_k^2 - ω^n -large.

Proof. Since $\min X \geq 2$, then there exist $X_0 < X_1 \subseteq X \setminus \{\min X\}$ that are ω^{kn+3} -large. As X_0 is ω^{kn+3} -large, we have $\min X_1 > \max X_0 \geq 6k + 2$, hence there is an $\omega^2 \cdot 3$ -sparse and $\omega^{nk} \cdot (2k) + 1$ -large subset $Y \subseteq X_1$ by Lemma 13.2.8. By Lemma 13.2.6, Y is thus $(x \mapsto x^{R_x(2x+2)})$ -sparse.

We also have $\min Y > \max X_0 \geq kn$, hence by Theorem 13.2.19, Y (and therefore X) is RT_k^2 - ω^n -large. $\square$

Chapter 14

$\forall\Pi_4^0$ -conservation of Ramsey's theorem for pairs

In this chapter, we pursue the search for the first-order part of RT_2^2 by extending the characterization due to Patey and Yokoyama [115] of the $\forall\Pi_3^0$ -consequences of $\text{RCA}_0 + \text{RT}_2^2$ to a characterization of its $\forall\Pi_4^0$ -consequences, via the following result:

Main Theorem 14.0.1. $\text{WKL}_0 + \text{RT}_2^2$ is $\forall\Pi_4^0$ -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$.

This constitutes a step toward a characterization of the full first-order part of RT_2^2 , as well as toward resolving whether RT_2^2 is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$. By Fiori-Carones et al. [46], it would suffice to go one step further and prove the $\forall\Pi_5^0$ -conservation of RT_2^2 over $\text{RCA}_0 + \text{B}\Sigma_2^0$, in order to obtain the full Π_1^1 -conservation.

The original article [88], which established the $\forall\Pi_4^0$ -conservation of $\text{RT}_2^2 + \text{WKL}_0$ over $\text{RCA}_0 + \text{B}\Sigma_2^0$, derived bounds for RT_2^2 - ω^n -largeness(θ) (although the bound for ADS - ω^n -largeness(θ) as initially given was flawed and later corrected in [86]). The combinatorial arguments in [88] adapt those of Kołodziejczyk and Yokoyama [79] to the setting of largeness(θ) and are quite involved.

Fortunately, we can again exploit the decomposition of RT_2^2 into $\text{EM} + \text{ADS}$, together with the fact that ADS is already known to be Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$ (Chong, Slaman and Yang [22]). This allows us to focus solely on bounds for EM - ω^n -largeness(θ) (or rather, for fEM - ω^n -largeness(θ)) and, by Theorem 12.3.7, to obtain the $\forall\Pi_4^0$ -conservation of $\text{WKL}_0 + \text{EM}$ over $\text{RCA}_0 + \text{B}\Sigma_2^0$. By an amalgamation theorem of Patey and Yokoyama (Theorem 14.0.9), we can then combine those two conservation results, yielding a proof of Main Theorem 14.0.1.

By restricting ourselves to EM , we can adapt the combinatorics used in Section 13.2 to largeness(θ) instead of those of Kołodziejczyk and Yokoyama [79], leading to a shorter proof.

The following lemma is the analogue of Proposition 12.2.7 for largeness(θ). The bound obtained here, $2n$, is worse than the bound $n+1$ given in Proposition 12.2.7. It was proved in [88, Section 7] that this bound is tight, in the sense that for every $n \in \mathbb{N}$, there is a set F and a formula θ_F depending on F such that F is ω^{2n-1} -large(θ_F) but not RT_2^1 - ω^n -large(θ_F). This difference in the bounds propagates throughout the rest of the proof, yielding an exponential bound for EM - ω^n -largeness(θ) in contrast to the linear bound of Corollary 13.2.14.

This exponential bound prevents us from directly adapting the framework of Kołodziejczyk et al. [78] to show that every proof of a $\forall\Pi_4^0$ -statement in $\text{WKL}_0 + \text{RT}_2^2$ can be transformed into a proof of the same statement in $\text{RCA}_0 + \text{B}\Sigma_2^0$ with only a polynomial increase in size. Nevertheless, polynomial bounds for RT_2^2 - ω^n -largeness(θ) can in fact be obtained, yielding such a result, as shown in [91]. The derivation of these bounds relies on more involved combinatorics than those presented here,

including a tree partition theorem based on Milliken's tree theorem. Given the elaboration of the argument and the complexity of the framework of Kołodziejczyk et al. [78], we shall not present those results.

In this chapter, for every proposition, the formula θ is taken to be universally quantified over all Σ_0^0 -formulas within the considered theory.

Lemma 14.0.2. $|\Sigma_1^0$ proves that for all $b \in \mathbb{N}$, every ω^{2b} -large(θ) and $(x \mapsto 4^x)$ -sparse set X is RT^1 - ω^b -large(θ).

Proof. By induction over b .

Case $b = 0$. For every ω^0 -large(θ) set X and every $f : X \rightarrow \min X$, every 1-element subset of X is f -homogeneous and ω^0 -large(θ).

Case $b > 0$. Fix an ω^{2b} -large(θ) $(x \mapsto 4^x)$ -sparse set X and $f : X \rightarrow \min X$. Since X is ω^{2b} -large(θ), there are $\min X$ pairwise θ -apart ω^{2b-1} -large(θ) subsets $X_0 < \dots < X_{\min X-1}$ of $X \setminus \{\min X\}$. Suppose X_0 is not f -homogeneous, otherwise we are done. Let $t < \min X$ be maximal such that $f[\bigcup_{i < t} X_i] \supseteq f[X_t]$. Note that $t \geq 1$ exists since otherwise,

$$2 \leq \text{card } f[X_0] < \text{card } f[X_0 \cup X_1] < \dots < \text{card } f[X_0 \cup \dots \cup X_{\min X-1}]$$

and therefore $\text{card } f[X] \geq \min X + 1$, contradicting the assumption that $f : X \rightarrow \min X$.

Since X_t is ω^{2b-1} -large(θ), there are $\min X_t$ pairwise θ -apart ω^{2b-2} -large(θ) subsets $Y_0 < \dots < Y_{\min X_t-1}$ of $X_t \setminus \{\min X_t\}$.

By induction hypothesis, for every $j < \min X_t$, there is a ω^{b-1} -large(θ) subset $Z_j \subseteq Y_j$ which is f -homogeneous for some color $c_j < \min X$. Since X is $(x \mapsto 4^x)$ -sparse, $4^{\max X_{t-1}} < \min X_t$ so $(\max X_{t-1})^2 < \min X_t$, thus $\min X \times \max X_{t-1} < \min X_t$. Thus, by the finite pigeonhole principle, there is a subset $J \subseteq \{0, \dots, \min X_t - 1\}$ of size at least $\max X_{t-1}$ and some color $c < \min X$ such that for every $j \in J$, $c_j = c$. By choice of t , $f[\bigcup_{i < t} X_i] \supseteq f[X_t]$, so there is some $x \in \bigcup_{i < t} X_i$ with $f(x) = c$. Thus, the set $\{x\} \cup \bigcup_{j \in J} Z_j$ is an ω^b -large(θ) f -homogeneous subset of X . $\square$

As explained in Section 13.2.1, there are two natural ways to define when a sequence $X_0 < \dots < X_{d-1}$ of blocks is $\mathcal{L}$ -large, where $\mathcal{L}$ is a largeness notion. One may either require that the set $\{\max X_i : i < d\}$ is $\mathcal{L}$ -large, or instead require that every sequence in $X_0 \times \dots \times X_{d-1}$ is $\mathcal{L}$ -large. For α -largeness(θ), these two definitions no longer coincide. The following lemma provides an analogue of Proposition 13.2.2 for largeness(θ) when the second definition is used. As with Lemma 14.0.2, the $2b$ in the bound is tight because of the apartness requirement. This allows us to prove Proposition 14.0.4, a counterpart of Proposition 13.2.7 for largeness(θ) which will be used to obtain sufficiently sparse ω^m -large(θ) sets.

Lemma 14.0.3. $|\Sigma_1^0$ proves that for every a, b and for every ω^{a+2b+1} -large(θ) set X , there are some $c \in \mathbb{N}$ and some ω^a -large(θ) pairwise θ -apart subsets $X_0 < \dots < X_{c-1}$ of X such that every $H \in \prod_{i < c} X_i$ is ω^b -large(θ) (here, we abuse the product notation and see H as a subset of $\mathbb{N}$ intersecting every X_i exactly once rather than a tuple in the product $X_0 \times \dots \times X_{c-1}$).

Proof. By Π_1^0 -induction over b , we prove the following statement that directly implies the lemma (since $\min X \geq 2$): for all a , for all θ -apart pairs $Y_0 < Y_1$ of ω^{a+2b} -large(θ) sets, there exists some $c < \max Y_1$ and some ω^a -large(θ) subsets $X_1 < \dots < X_{c-1}$ of Y_1 such that, letting $X_0 = Y_0$, every $H \in \prod_{i < c} X_i$ is ω^b -large(θ) and $X_0, \dots, X_{c-1}$ are pairwise θ -apart. This is indeed a Π_1^0 -formula, as being θ -apart, or being ω^a -large(θ) are Σ_0^0 -statements.

Case $b = 0$. The result is clear, as every non-empty set is ω^0 -large(θ).

Case $b > 0$. Let $Y_0 < Y_1$ be ω^{a+2b} -large(θ) and θ -apart, let $Z_0 < \dots < Z_{\min Y_1-1}$ be ω^{a+2b-1} -large(θ) pairwise θ -apart subsets of Y_1 . For every $i < \min Y_1$, let Z_i^0, Z_i^1 be ω^{a+2b-2} -large(θ) and θ -apart subsets of Z_i .

We can then apply the inductive hypothesis to the pairs

$$(Z_0^0, Z_0^1), (Z_1^0, Z_1^1) \dots, (Z_{\max Y_0 - 1}^0, Z_{\max Y_0 - 1}^1)$$

to get for every $j < \max Y_0$, families of pairwise θ -apart ω^a -large(θ) subsets $Z_j^0 = X_{j,0} < \dots < X_{j,c_j}$ of $Z_j^0 \cup Z_j^1$ such that every $H \in \prod_{i \leq c_j} X_{j,i}$ is ω^{b-1} -large(θ).

Consider the family $Y_0 < X_{0,0} < X_{0,1} < \dots < X_{0,c_0} < X_{1,0} < X_{1,1} < \dots < X_{\max Y_0 - 1, c_{\max Y_0 - 1}}$. Every block of this family is ω^a -large(θ). Since $X_{0,0} \subseteq Y_1$, then Y_0 and $X_{0,0}$ are θ -apart. Moreover, for all $j < \max Y_0 - 1$, since $X_{j,c_j} \subseteq Z_j$ and $X_{j+1,0} = Z_{j+1}^0$, X_{j,c_j} and $X_{j+1,0}$ are θ -apart.

Let $H \in Y_0 \times X_{0,0} \times \dots \times X_{0,c_0} \times X_{1,0} \times \dots \times X_{\max Y_0 - 1, c_{\max Y_0 - 1}}$. Every element of $\prod_{i < c_j} X_{j,i}$ is ω^{b-1} -large(θ), therefore H is ω^b -large(θ).

This completes the proof. $\square$

Lemma 12.3.4. $|\Sigma_1^0$ proves that, for every $a \in \mathbb{N}$, if $X \subseteq_{\text{fin}} \mathbb{N}$ is ω^{4a+4} -large(θ), then it is RT^1 - ω^a -large(θ).

Proof. By Lemma 14.0.3, there exists $c \in \mathbb{N}$ and pairwise θ -apart subsets $X_0 < \dots < X_{c-1}$ of X , each ω^3 -large, such that the set $Y := \{\max X_i : i < c\}$ is ω^{2a} -large(θ). The set Y is ω^3 -sparse and, in particular, $(x \mapsto 4^x)$ -sparse. By Lemma 14.0.2, it follows that Y is RT^1 - ω^a -large(θ), and hence X is also RT^1 - ω^a -large(θ). $\square$

Proposition 14.0.4. *For every primitive recursive and increasing function $g : \mathbb{N} \rightarrow \mathbb{N}$, there exists some bound $n \in \omega$ such that $|\Sigma_1^0$ proves that, for every $b \in \mathbb{N}$, every ω^{n+2b} -large(θ) set F contains an ω^b -large(θ) and g -sparse subset.*

Proof. By Theorem 13.2.5, there exists some $n' \in \mathbb{N}$ such that every $\omega^{n'}$ -sparse set is g -sparse. Let $n = n' + 1$ and let $X \subseteq_{\text{fin}} \mathbb{N}$ be ω^{n+2b} -large(θ). By Lemma 14.0.3, there exists some $c \in \mathbb{N}$ and some $\omega^{n'}$ -large(θ) subsets $X_0 < \dots < X_{c-1}$ such that $Y := \{\max X_i : i < c\}$ is ω^b -large(θ).

We claim that Y is $\omega^{n'}$ -sparse and therefore g -sparse. Indeed, let $x < y \in Y$, there exists some $i < c$ such that $y = \max X_i$, hence $X_i \subseteq (x, y]$ and $(x, y]$ is thus $\omega^{n'}$ -large. $\square$

The following proof follows Lemma 13.2.12 but with the bound obtained in Lemma 14.0.2 instead of the one of Proposition 12.2.7.

Lemma 14.0.5. $|\Sigma_1^0$ proves that, for $\ell > 1$, $a, d \in \mathbb{N}$, and $X \subseteq_{\text{fin}} \mathbb{N}$ an $\omega^{2a} \cdot R_\ell(d)$ -large(θ) and $(x \mapsto x\ell^x)$ -sparse set such that $\ell^{R_\ell(d)} \leq \min X$, and for every coloring $f : [X]^2 \rightarrow \ell$, there exists a sequence $Y_0 < \dots < Y_{d-1}$ of pairwise θ -apart ω^a -large(θ) subsets of X and some color $c < \ell$ such that $f(x, y) = c$ for every $x \in Y_s$ and $y \in Y_t$ for $s < t < d$.

Proof. For $a = 0$: let $\{x_0\} < \dots < \{x_{R_\ell(d)-1}\} \subseteq X$ be ω^0 -large(θ) and pairwise θ -apart. By definition of R there is an f -homogeneous subset $\{y_0 < y_1 < \dots < y_{d-1}\} \subseteq \{x_0, \dots, x_{R_\ell(d)-1}\}$ and pick $Y_i = \{y_i\}$ for every $i < d$. The Y_i are ω^0 -large(θ) and θ -apart (as the $\{x_i\}$ are θ -apart).

Assume from now on that $a > 0$. Let $X_0 < \dots < X_{R_\ell(d)-1}$ be a decomposition of X into θ -apart ω^{2a} -large(θ) sets. Consider the sequence $(f_i)_{i < R_\ell(d)}$ of colorings and the sequence $Z_0 < \dots < Z_{R_\ell(d)-1}$ of sets defined inductively as follows:

Assume that $Z_{i+1}, \dots, Z_{R_\ell(d)-1}$ have been defined for some $i < R_\ell(d)$, and consider the coloring $f_i : X_i \rightarrow \ell^{R_\ell(d)-1-i+\sum_{j < i} \text{card } X_j}$ defined by

$$f_i(x) = (f(y, x))_{y \in \bigcup_{j < i} X_j} \sqcup (f(x, y))_{y \in \{\min Z_j : i < j < R_\ell(d)\}}$$

As $\min X_0 \geq R_\ell(d)$ and as the X_i are disjoint, we get that for $i > 0$, $R_\ell(d) + \sum_{j < i} \text{card } X_j \leq \max X_{i-1}$. By $(x \mapsto x\ell^x)$ -sparsity of X , $\ell^{\max X_{i-1}} \cdot \max X_{i-1} \leq \min X_i$, so $\ell^{R_\ell(d)+\sum_{j < i} \text{card } X_j} \cdot \max X_{i-1} \leq \min X_i$. By Lemma 14.0.2, there exists

some ω^a -large(θ) and f_i -homogeneous subset $Z_i \subseteq X_i$. For $i = 0$, since $\ell^{R_\ell(d)} \leq \min X_0$, by Lemma 14.0.2, there exists some ω^a -large(θ) and f_0 -homogeneous subset $Z_0 \subseteq X_0$.

The $(Z_i)_{i < R_\ell(d)}$ are defined such that, for every $i < j < R_\ell(d)$ and $x \in Z_i$, $y \in Z_j$, we have $f(x, y) = f(\min Z_i, \min Z_j)$. By definition of $R_\ell(d)$, there exists some subset $\{i_0 < \dots < i_{d-1}\} \subseteq \{0, \dots, R_\ell(d) - 1\}$ and some color $c < \ell$ such that, for every $j < k < d$, $x \in Z_{i_j}$ and $y \in Z_{i_k}$, $f(x, y) = c$.

Finally, let $Y_j = Z_{i_j}$ for every $j < d$. Every Y_j is ω^a -large(θ), and, as the X_i are pairwise θ -apart, they are also pairwise θ -apart. $\square$

And finally, we can adapt Proposition 13.2.13 to largeness(θ), which gives us an exponential bound.

Proposition 14.0.6. $|\Sigma_1^0$ proves that, for $a \in \mathbb{N}$, and $X \subseteq_{\text{fin}} \mathbb{N}$ an ω^{4^a-1} -large(θ) and $(x \mapsto x^{R_x(x)})$ -sparse set, X is $\text{fEM-}\omega^a$ -large(θ).

Proof. We proceed by induction on a .

The result holds trivially for $a = 0$, as any ω^0 -large(θ) set is non-empty, and any singleton is $\text{fEM-}\omega^0$ -large(θ).

Suppose the property holds for some a . Let $f : [X]^2 \rightarrow \min X$ be an instance of fEM on X . Let $X_0 < X_1$ be $\omega^{4^{a+1}-2}$ -large(θ) subsets of X and let $x_0 = \min X_0$. Consider the coloring $g : X_1 \rightarrow \min X$ defined by $g(y) = f(x_0, y)$, by Lemma 14.0.2, there exists some $\omega^{4^a \cdot 2-1}$ -large(θ) subset $Y \subseteq X_1$ such that $f(x_0, y) = f(x_0, y')$ for every $y, y' \in Y$.

As $x^{R_x(x)} > R_x(x)$ for every $x > 1$, then, by $(x \mapsto x^{R_x(x)})$ -sparsity of X , $Y \setminus \{\min Y\}$ is $\omega^{4^a \cdot 2-2} \cdot R_{x_0}(x_0)$ -large(θ) and $x_0^{R_{x_0}(x_0)} \leq \min Y \setminus \{\min Y\}$. Then, as $x^{R_x(x)} > x^{x+1}$ for every $x > 1$, $Y \setminus \{\min Y\}$ is $(x \mapsto xx^x)$ -sparse, so, by Lemma 14.0.5, there exists a color $c < x_0$ and a sequence $Y_0 < \dots < Y_{x_0-1}$ of pairwise θ -apart ω^{4^a-1} -large(θ) subsets of $Y \setminus \{\min Y\}$ such that $f(x, y) = c$ for every $x \in Y_s$ and $y \in Y_t$ for $s < t < x_0$.

We can then apply the inductive hypothesis on each of the Y_i and obtain a sequence $Z_0 < \dots < Z_{x_0-1}$ of f -fallow and pairwise θ -apart ω^a -large(θ) subsets of $Y \setminus \{\min Y\}$.

Thus, the set $Z := \{x_0\} \cup \bigcup_{i < x_0} Z_i$ is an ω^{a+1} -large(θ) subset of X . We claim that Z is f -fallow. Suppose otherwise and let $x < y < z \in Z$ be such that $f(x, z) \notin \{f(x, y), f(y, z)\}$, then x cannot be equal to x_0 as $f(x_0, y) = f(x_0, z)$ by definition of Y , hence there exists some $i \leq j \leq k < x_0$ such that $x \in Z_i$, $y \in Z_j$ and $z \in Z_k$. If $j < k$, then, by our use of Lemma 14.0.5, $f(x, z) = f(y, z) = c$ which contradicts our assumption, similarly, if $i < j$, then $f(x, y) = f(x, z) = c$, again contradicting our assumption. Hence, we necessarily have $i = j = k$, which is impossible since Z_i is f -fallow by our use of the inductive hypothesis. Hence, Z is indeed f -fallow. $\square$

Corollary 14.0.7. There exists some $n \in \mathbb{N}$ such that $|\Sigma_1^0$ proves that, for $a \in \mathbb{N}$, every $\omega^{4^a \cdot 2+n}$ -large(θ) set is $\text{fEM-}\omega^a$ -large(θ).

Proof. The function $(x \mapsto x^{R_x(x)})$ being primitive recursive, by Proposition 14.0.4 there exists some $n \in \mathbb{N}$ such that, for every $a \in \mathbb{N}$, every $\omega^{4^a \cdot 2+n}$ -large(θ) set contains an ω^{4^a-1} -large(θ) and $(x \mapsto x^{R_x(x)})$ -sparse subset, and, by Proposition 14.0.6, is therefore $\text{fEM-}\omega^a$ -large(θ). $\square$

Proposition 14.0.8. $\text{WKL}_0 + \text{fEM}$ is $\forall\Pi_4^0$ -conservative over $\text{RCA}_0 + \text{BS}_2^0$.

Proof. By Theorem 12.3.7, using the bounds obtained in Corollary 14.0.7. $\square$

To obtain the $\forall\Pi_4^0$ -conservation of RT_2^2 over $\text{RCA}_0 + \text{BS}_2^0$, we will use the following theorem of Patey and Yokoyama [115]. It is a generalization of an amalgamation theorem for Π_1^1 -conservation by Yokoyama [142], that was used in Chapter 10 (Theorem 10.4.1).

Theorem 14.0.9 (Amalgamation (Patey and Yokoyama [115, Theorem 3.6])). *Fix $n \geq 1$. Let $\mathcal{T}$ be a theory extending $\text{I}\Sigma_1^0$ which consists of sentences of the form $\forall X \exists Y \theta(X, Y)$ where θ is Π_{n+2}^0 , and let Γ_1 and Γ_2 be sentences of the same form. If $\mathcal{T} + \Gamma_i$ is $\forall \Pi_{n+2}^0$ -conservative over $\mathcal{T}$ for $i = 1, 2$, then $\mathcal{T} + \Gamma_1 + \Gamma_2$ is $\forall \Pi_{n+2}^0$ -conservative over $\mathcal{T}$.*

Main Theorem 14.0.1. $\text{WKL}_0 + \text{RT}_2^2$ is $\forall \Pi_4^0$ -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$.

Proof. By a result of Chong, Slaman, and Yang [22], ADS is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$ (and therefore $\forall \Pi_4^0$ -conservative). By Theorem 14.0.8, so is $\text{EM} + \text{WKL}$. Hence, by Theorem 14.0.9, $\text{WKL}_0 + \text{RT}_2^2$ is $\forall \Pi_4^0$ -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$. $\square$

As discussed at the beginning of the chapter, it is possible to obtain polynomial bounds for RT_2^2 - ω^n -largeness(θ), as shown in [91], using more involved combinatorics. Then, using the framework of Kołodziejczyk et al. [78], we obtain the following result, which shows that the above conservation can be witnessed by a polynomial-time procedure.

Theorem 14.0.10 ([91, Main Theorem 1.1]). *We can effectively transform every proof of a $\forall \Pi_4^0$ -statement in $\text{WKL}_0 + \text{RT}_2^2$ into a proof of that same statement in $\text{RCA}_0 + \text{B}\Sigma_2^0$ with only up to a polynomial increase in size.*

Chapter 15

$\forall\Pi_4^0$ -conservation of the ordered variable word theorem

15.1 Introduction

In this chapter, we will use the formalism introduced in Section 12.3 and in Chapter 14 to prove the $\forall\Pi_4^0$ -conservation over $\text{RCA}_0 + \text{B}\Sigma_2^0$ of another statement from Ramsey theory: the ordered variable word theorem (OVW). OVW is a so-called *tree partition theorem*, i.e., a statement of the form

“For every finite coloring of the finite subtrees of an infinite tree-like structure, there exists an isomorphic sub-structure whose finite subtrees are monochromatic.”

Perhaps the simplest tree partition theorem is the tree theorem for singletons (TT^1), which says that for every finite coloring of $2^{<\mathbb{N}}$, there is a monochromatic subset $T \subseteq 2^{<\mathbb{N}}$ such that $(T, \leq)$ is isomorphic to $(2^{<\mathbb{N}}, \leq)$.

Tree partition theorems do not fit exactly into the formalism of Ramsey-like statements as defined in Definition 12.2.5. Nonetheless, applying the same approach by considering finitistic versions of the theorems defined in terms of α -largeness and α -largeness(θ) also produces conservation results.

Tree partition theorems play an important role in structural Ramsey theory. Many proofs of existence of *big Ramsey numbers* are reduced to higher-order versions of these theorems. For example, the existence of big Ramsey numbers for partitions of the rationals [33] or of the Rado graph [122] are both reduced to Milliken’s tree theorem [101].

Both the tree theorem (TT) and Milliken’s tree theorem (MTT) have been extensively studied from a reverse mathematical viewpoint (see [21, 24, 25, 30, 40, 112] for TT and [1] for MTT). The restrictions of TT and MTT to colorings of singletons are *computably true*, that is, every instance admits a solution computable in the instance. Thus, as with our study of the bounded Ramsey theorem for pairs BRT_2^2 in Chapter 11, their strength can be measured only in terms of the amount of induction necessary to prove them.

In this chapter, we study the first-order part of a stronger tree partition theorem, due to Carlson and Simpson [13], called the ordered variable word theorem¹. Fix a finite alphabet A and a distinguished variable symbol $\star$. A *word* over A is a finite

¹This statement is also called *Carlson-Simpson theorem* in combinatorics [35], but should be distinguished from the variable word theorem [100], also known as Carlson-Simpson lemma in reverse mathematics, which is a similar statement, but where each variable type is allowed to appear infinitely often.

sequence $w = a_0 \dots a_{k-1}$ where $a_i \in A$. Again, we write $w(i)$ for a_i and let $|w| = k$. A *left variable word* over A is a word w over $A \sqcup \{\star\}$ such that $w(0) = \star$. Given a left variable word w and a letter $a \in A$, we let $w[a]$ be the word of length $|w|$ obtained from w by substituting every occurrence of $\star$ by a . Beware, the notations $w(i)$ and $w[a]$ should not be confused: the former yields a letter, while the latter is a substitution.

Statement 15.1.1 (Ordered variable word theorem). OVW is the statement that: For every finite alphabet A and every finite coloring $f : A^{<\mathbb{N}} \rightarrow \ell$, there is a word c_0 over A and an infinite sequence of left variable words $w_1, w_2, \dots$ such that the following set is monochromatic:

$$\{c_0 w_1[a_1] w_2[a_2] \dots w_k[a_k] : k \in \mathbb{N}, a_1, \dots, a_k \in A\}$$

The ordered variable word theorem can be seen as a problem whose *instances* are pairs (A, f) , where A is a finite alphabet and f is a finite coloring of $A^{<\mathbb{N}}$. A *solution* to (A, f) is the giving of the sequence $c_0, w_1, w_2, \dots$ as above. The set $T := \{c_0 w_1[a_1] w_2[a_2] \dots w_k[a_k] : k \in \mathbb{N}, a_1, \dots, a_k \in A\}$ forms an f -homogeneous subtree of the original tree $A^{<\mathbb{N}}$ satisfying very strong structural properties. For example, for every $k \in \mathbb{N}$, there exists a single $n \in \mathbb{N}$ such that $\{c_0 w_1[a_1] w_2[a_2] \dots w_k[a_k] : a_1, \dots, a_k \in A\} \subseteq A^n$ (i.e., T is a *strong* subtree of $A^{<\mathbb{N}}$). Moreover, the paths between any two levels of T are all generated by instantiating the same left variable word.

The ordered variable word theorem was used as a pigeonhole principle by Carlson and Simpson to prove a dual version of Ramsey's theorem [13]. It was later used by Hubička [65] to give a simple proof of the existence of big Ramsey degrees for the universal triangle-free graph. See the monograph of Dodos and Kanellopoulos [34] for an extensive combinatorial analysis of variable word theorems. Friedman and Simpson [47] asked about its reverse mathematical strength.

Contrary to the tree theorem for singletons and Milliken's tree theorem for singletons, which are computably true, Miller and Solomon [100] proved that OVW admits a computable instance with no Δ_2 solutions. Therefore, OVW cannot be proved in WKL_0 , as there exists a model of WKL_0 containing only low (and hence Δ_2) sets by the low basis theorem (Theorem 4.2.7). Later, Liu, Monin, and Patey [97] constructed a computable instance of OVW whose solutions compute a $\emptyset'$ -DNC function, that is, a total function $f : \mathbb{N} \rightarrow \mathbb{N}$ such that $\forall e f(e) \neq \Phi_e^{\emptyset'}(e)$. On the other hand, Anglès d'Auriac et al. [2] proved that every computable instance of OVW admits a solution computable from any set PA over $\emptyset'$. The exact computable strength of OVW therefore lies between $\emptyset'$ -DNC degrees and PA degrees over $\emptyset'$. Note that the existence of the former is strictly weaker than ACA_0 , while the latter implies it. Whether OVW implies ACA_0 over RCA_0 was left open. We answer the question negatively:

Theorem 15.1.2. $\text{WKL}_0 + \text{OVW}$ does not imply ACA_0 .

Usually, separations from ACA_0 are done using the so-called *cone avoidance* property, like in Chapter 8, that is, by proving that for every non-computable set A , every computable instance of the problem admits a solution which does not compute A . In this chapter, we take a different approach, and prove the separation through a partial conservation result.

Theorem 15.1.3. $\text{WKL}_0 + \text{OVW}$ is $\forall\Pi_4^0$ -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$.

The tree theorem for singletons is strictly in between $\text{I}\Sigma_2^0$ and $\text{B}\Sigma_2^0$ over RCA_0 (see [30, 21]). It is unknown whether TT^1 is Π_1^1 -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$.

Chong, Wang, and Yang [24] obtained a partial conservation result by proving that TT^1 is $\forall\Pi_3^0$ -conservative over RCA_0 . Theorem 15.1.3 strengthens Chong, Wang, and Yang's result since TT^1 follows from OVW.

We defined largeness notions in terms of sets of integers, but the ordered variable word theorem is a statement about words and variable words. We bridge the two notions by defining an ordered X -variable word over a finite set $X \subseteq \mathbb{N}$.

Definition 15.1.4 (Ordered Y -variable word). For $Y = \{y_0, \dots, y_{n-1}\}$ a finite set, an *ordered Y -variable word over A* is a finite word w over the alphabet $A \sqcup \{x_j : j < n\}$ where for every $j < n$, the first occurrence of x_j is at position y_j and for every $j < n-1$ its last occurrence is before position y_{j+1} .

Definition 15.1.5 (Substitution). For w a Y -variable word over an alphabet A and $u \in A^{\leq \text{card } Y}$, the substitution $w[u]$ is defined as the word w where every occurrence of x_i for $i < |u|$ is replaced by $u(i)$ and cut just before the first occurrence of $x_{|u|}$.

Example 15.1.6. On the alphabet $A = \{a, b\}$, $w = abx_0ax_0bx_1bb$ is a $\{2, 6\}$ -variable word, $ax_0bx_1x_0ab$ is not a variable word, since there is an occurrence of x_0 after an occurrence of x_1 and aax_1b is not a variable word either since there is no occurrence of x_0 . $w[\varepsilon] = ab$ (where ε is the empty word), $w[a] = abaaab$ and $w[ba] = abbabbabb$.

15.2 Hales-Jewett and Graham-Rothschild theorems

The ordered variable word theorem belongs to a whole family of variable word statements, whose pigeonhole principle is the *Hales-Jewett theorem* [58]. In its simplest form, the Hales-Jewett theorem asserts the existence, for every finite alphabet A and every finite coloring of $A^{<\mathbb{N}}$, of a variable word w such that $\{w[a] : a \in A\}$ is monochromatic. Its generalization to multiple dimensions is known as the Graham-Rothschild theorem [53]. In this section, we define notions of largeness for these theorems, and obtain explicit quantitative bounds for them. This analysis will be reused in the next section to obtain quantitative bounds for the ordered variable word theorem.

Definition 15.2.1. For Y a set, a *combinatorial Y -space* over an alphabet A is a set of the form $S = \{w[u] : u \in A^{\text{card } Y}\}$ for some ordered Y -variable word w over A . We call w its *generating variable word*. The *dimension* of a combinatorial Y -space is $\text{card } Y$, and a *combinatorial Y -line* is a combinatorial Y -space of dimension 1.

Definition 15.2.2. For S a combinatorial X -space, a *combinatorial subspace S'* of S is a combinatorial Y -space included in S for some set Y .

A consequence of the definition is that if S' is a combinatorial Y -subspace of a combinatorial X -space S , then $Y \subseteq X$ and the generating word w' of S' is equal to $w[u]$ for w the generating word of S and u some I -variable word of length $\text{card } X$, where $X = \{x_0, \dots, x_n\}$ and $Y = \{x_i : i \in I\}$.

The original proof of the following theorem by Hales and Jewett [58] had an extremely fast-growing bound. A more recent proof by Shelah [125] uses only Σ_1^0 -induction, and yields a primitive recursive bound, hence is provable in IS_1^0 .

Theorem 15.2.3 (Hales-Jewett [58], Shelah [125]). IS_1^0 proves that there exists a primitive recursive function $HJ(k, \ell)$ such that for every finite set X with $\text{card } X \geq HJ(k, \ell)$, every combinatorial X -space S over an alphabet A of size k and every coloring $f : S \rightarrow \ell$, there exists some $x \in X$ and an f -homogeneous combinatorial $\{x\}$ -subspace of S .

As mentioned above, the Hales-Jewett theorem admits a multidimensional generalization, known as the Graham-Rothschild theorem [53]. Its proof follows from the Hales-Jewett theorem by elementary combinatorics. The original Graham-Rothschild theorem allows the variable kinds to be unordered, that is the first

occurrence of x_i must appear before the first occurrence of x_{i+1} , but the last occurrence of x_i might appear later. We are going to use a slightly modified version of the Graham-Rothschild theorem, due to Dodos et al. [35, Theorem 2.1], which requires the variable words to be ordered. The proof of its primitive recursive bound is an adaptation of Shelah's bound [125] for the original Graham-Rothschild theorem, which can be found in Dodos and Kanellopoulos [34, Theorem 2.9, Theorem 2.15].

Theorem 15.2.4 (Graham and Rothschild [53], Dodos and Kanellopoulos [34]). IS_1^0 proves that there exists a primitive recursive function $GR(k, d, m, \ell)$ such that for every finite set X with $\text{card } X \geq GR(k, d, m, \ell)$, every combinatorial X -space S over an alphabet A of size k and every coloring f of the m -dimensional subspaces of S with ℓ colors, there exists some d -dimensional subspace of S , all of whose m -dimensional subspaces have the same color.

Definition 15.2.5. For $\theta(x, y, z)$ a Σ_0^0 -formula, a set $X \subseteq_{fin} \mathbb{N}$ is said to be $\text{GR-}\alpha\text{-large}(\theta)$ if for every $\ell, k < \min X$, every combinatorial X -space S over an alphabet A of size k and every coloring $f : S \rightarrow \ell$, there exists some α -large(θ) subset $Y \subseteq X$ and an f -homogeneous combinatorial Y -subspace of S .

The following lemma states that if X is sufficiently large, then the size of X will be large enough to be able to apply the Hales-Jewett theorem for $\min X - 1$ colors and alphabet of size $\min X - 1$, and get a monochromatic combinatorial line.

As in the previous chapter, all occurrences of θ in the results of this chapter are assumed to be universally quantified over all Σ_0^0 -formulas within the considered theory.

Lemma 15.2.6. *There exists some $n_0 \in \omega$ such that IS_1^0 proves that if X is ω^{n_0} -large(θ) then X is $\text{GR-}\omega^0\text{-large}(\theta)$.*

Proof. For X to be $\text{GR-}\omega^0\text{-large}(\theta)$, it is sufficient that $\text{card } X \geq HJ(\min X - 1, \min X - 1)$ and that $\min X > b$ for b the largest first-order constant appearing in θ . By Theorem 15.2.3, $x \mapsto HJ(x, x)$ is primitive recursive, so by Theorem 13.2.5 (and by regularity of α -largeness), there exists some n_0 such that every ω^{n_0} -large set X satisfies $\text{card } X \geq HJ(\min X - 1, \min X - 1)$. Thus, every ω^{n_0} -large(θ) set is $\text{GR-}\omega^0\text{-large}(\theta)$. $\square$

In the following lemma, it might be helpful to see a combinatorial X -space over an alphabet A as the set of leaves of a tree isomorphic to $A^{\leq \text{card } X}$.

Lemma 15.2.7. IS_1^0 proves that for all $b \in \mathbb{N}$ and every finite set X , if X is ω^{2b+n_0+1} -large(θ) and $(x \mapsto x^{x^x})$ -sparse then X is $\text{GR-}\omega^b\text{-large}(\theta)$.

Proof. Fix some $b \in \mathbb{N}$. Let X be ω^{2b+n_0+1} -large(θ), let $\ell, k < \min X$, let S be a combinatorial X -space over an alphabet A of size k with generating variable word w , and let $f : S \rightarrow \ell$ be a coloring.

By Lemma 14.0.3, there exist pairwise θ -apart ω^{n_0} -large(θ) subsets of $X \setminus \{\min X\}$ $X_0 < \dots < X_{a-1}$ such that any $Y \in \prod_{i < a} X_i$ is ω^b -large(θ). For simplicity, we will assume that S is a combinatorial $(X_0 \cup \dots \cup X_{a-1})$ -space (this can always be done by taking a subspace of the original space).

The coloring f induces a coloring on every combinatorial X_{a-1} -subspace of S . Let n_{a-1} be the number of such subspaces, there is one of them for every possible value taken by the variables of $X_0 \cup \dots \cup X_{a-2}$, so

$$n_{a-1} = k^{\text{card } X_0 + \dots + \text{card } X_{a-2}} < k^{\max X_{a-2}}$$

We can consider the product of all those colorings $f' : S' \rightarrow \ell^{n_{a-1}}$ for S' an arbitrary combinatorial X_{a-1} -space. By $(x \mapsto x^{x^x})$ -sparsity of X ,

$$\ell^{n_{a-1}} < \ell^{k^{\max X_{a-2}}} < \min X_{a-1}$$

And since X_{a-1} is ω^{n_0} -large(θ), by Lemma 15.2.6, there is some $y_{a-1} \in X_{a-1}$ and an f' -homogeneous combinatorial $\{y_{a-1}\}$ -subspace of S' . By definition of f' , for every combinatorial X_{a-1} -subspace of S , there is an f -homogeneous combinatorial $\{y_{a-1}\}$ -subspace of it.

Consider S_{a-1} an arbitrary combinatorial $(X_0 \cup \dots \cup X_{a-2})$ -subspace of S (for example, by fixing an arbitrary value for the variables of X_{a-1}). There is a one-to-one correspondence between the elements of S_{a-1} and the combinatorial X_{a-1} -subspaces of S : every element of S_{a-1} corresponds to a possible value for the variables of $X_0 \cup \dots \cup X_{a-2}$ and therefore belongs to only one X_{a-1} -subspace of S and vice versa, every X_{a-1} -subspace of S contains only one element of S_{a-1} . So consider $f_{a-1} : S_{a-1} \rightarrow \ell$ a coloring that takes any element of S_{a-1} to the color of the f -homogeneous combinatorial $\{y_{a-1}\}$ -subspace of the corresponding combinatorial X_{a-1} -subspace.

We can then repeat the same argument and find some $y_{a-2} \in X_{a-2}$ such that, for every combinatorial X_{a-2} -subspace of S_{a-1} there is an f_{a-1} -homogeneous combinatorial $\{y_{a-2}\}$ -subspace of it. Therefore, for every combinatorial $(X_{a-2} \cup X_{a-1})$ -subspace of S , there is an f -homogeneous combinatorial $\{y_{a-2}, y_{a-1}\}$ -subspace of it.

Iterate the construction to find a sequence $y_0 \in X_0, \dots, y_{a-1} \in X_{a-1}$ such that, by letting $Y = \{y_0, \dots, y_{a-1}\}$, there is an f -homogeneous combinatorial Y -subspace of S . And Y is ω^b -large(θ) since $Y \in \prod_{i < a} X_i$. $\square$

15.3 Ordered variable word theorem

We now turn to the quantitative analysis of the target theorem of this chapter, namely, the ordered variable word theorem. The main result of this section is Corollary 15.3.9, which is then used to prove our conservation theorem.

Definition 15.3.1 (Ordered variable word Y -tree). For Y a set, an *OVW Y -tree* over an alphabet A is a set of the form $T = \{w[u] : u \in A^{\leq \text{card } Y}\}$ for some ordered Y -variable word w over A . We call w its *generating variable word*. The *dimension* of an OVW Y -tree T over A is $\text{card } Y$ and an *OVW Y -line* is an OVW Y -tree of dimension 1.

Definition 15.3.2. For T an OVW X -tree, an *OVW subtree* T' of T is an OVW Y -tree included in T for some set Y .

A consequence of the definition is that if T' is an OVW Y -subtree of an OVW X -tree T , then $Y \subseteq X$ and the generating word w' of T' is equal to $w[u]$ for w the generating word of T and u some I -variable word, where $X = \{x_0, \dots, x_n\}$ and $Y = \{x_i : i \in I\}$. We shall mainly consider two kinds of Y -subtrees: subtrees obtained by instantiating some variables of the generating word, which doesn't change its length, and subtrees obtained by truncating the generating word.

The following theorem is a finitary version of an infinitary theorem due to Carlson and Simpson [13]. The finitary version follows from its infinite version by compactness (by an argument similar to that in the proof of the finite Ramsey theorem (Theorem 6.0.3)). Dodos et al. [35, Theorem 4.1] gave an elementary proof of a higher order variant of its finitary version, with a primitive recursive bound. The finitary version of the ordered variable word follows easily from the Graham-Rothschild theorem, and thus can almost be considered as folklore. The proof below appears in Dodos and Kanellopoulos [34, Proposition 4.10].

Theorem 15.3.3 (Carlson and Simpson [13], Dodos and Kanellopoulos. [34]). $\aleph_1^0$ proves that there exists a primitive recursive bound $CS(k, d, \ell)$ such that for every $X \subseteq_{\text{fin}} \mathbb{N}$ with $\text{card } X \geq CS(k, d, \ell)$, every OVW X -tree T over an alphabet A of size

k and every coloring $f : T \rightarrow \ell$ there exists some subset $Y \subseteq X$ with $\text{card } Y = d$ and an f -homogeneous OVW Y -subtree of T .

Proof. Let $CS(k, d, \ell) := GR(k, d + 1, 1, \ell)$ and fix a set X with $\text{card } X \geq CS(k, d, \ell)$. Fix an OVW X -tree T over an alphabet A of size k and a coloring $f : T \rightarrow \ell$. Let w be the generating variable word of T and consider the combinatorial X -space S corresponding to the set of leaves of the tree T . A 1-dimensional subspace of S corresponds to an instantiation of all the kinds of variables in w except one, so define a coloring g that takes any 1-dimensional subspace of S to the color (by f) of its generating variable word cut before the first appearance of its only kind of variables. By definition of $GR(k, d + 1, 1, \ell)$, there is a $(d + 1)$ -dimensional subspace S' of S , all of whose 1-dimensional subspaces are of the same color for g . Let w' be the generating word of S' , and let w'' be w' cut before the first occurrence of its last variable: x_d , let Y be such that w'' is a Y -variable word ($\text{card } Y = d$). Then $\{w''[u] : u \in A^{\leq \text{card } Y}\}$ is an f -homogeneous OVW Y -subtree of T (The f -color of $w''[u]$ is the g -color of any 1-dimensional subspace of S' whose non-instantiated variable is the $|u|$ -th one). $\square$

Note that in the previous proof, we used the Graham-Rothschild theorem with colorings of 1-dimensional spaces to prove the Carlson-Simpson theorem which is about colorings of 0-dimensional OVW trees, that is, colorings of words. This can be generalized to prove a higher order version of the finite Carlson-Simpson theorem with primitive recursive bounds (see Dodos and Kanellopoulos [34, Theorem 4.21]). However, only the 0-dimensional version will be used in this chapter.

Definition 15.3.4. Let $\theta(x, y, z)$ be a Σ_0^0 -formula, a set $X \subseteq_{fin} \mathbb{N}$ is said to be OVW- α -large(θ) if for every $\ell, k < \min X$, every OVW X -tree T over an alphabet A of size k and every coloring $f : T \rightarrow \ell$, there is an α -large(θ) subset $Y \subseteq X$ and an f -homogeneous OVW Y -subtree of T .

Note that by definition of an ordered Y -variable word w , $|w| \geq \max Y$.

Lemma 15.3.5. *There exists some $n_1 \in \omega$ such that $|\Sigma_1^0|$ proves that if X is ω^{n_1} -large(θ) then X is OVW- $\omega^0 \cdot (\min X - 1)$ -large(θ).*

Proof. For X to be OVW- $\omega^0 \cdot (\min X - 1)$ -large(θ), it is sufficient to have $\text{card } X \geq CS(\min X - 1, \min X, \min X - 1)$ and have $X > b$ for b the largest first-order constant appearing in θ . By Theorem 15.3.3, the function $x \mapsto CS(x, x + 1, x)$ is primitive recursive, so by Theorem 13.2.5 (and by regularity of α -largeness), there exists some n_1 such that if X is ω^{n_1} -large(θ), then $\text{card } X \geq CS(\min X - 1, \min X, \min X - 1)$. Thus, every ω^{n_1} -large(θ) set is OVW- $\omega^0 \cdot (\min X - 1)$ -large(θ). $\square$

Lemma 15.3.6. $|\Sigma_1^0|$ proves that for every $b, r \in \mathbb{N}$, if the following statement holds:

- “Every ω^b -large(θ) $(x \mapsto x^{x^x})$ -sparse set is OVW- ω^r -large(θ)”

then the following statement holds:

- “Every ω^{2b+n_0+3} -large(θ) $(x \mapsto x^{x^x})$ -sparse set X is OVW- $\omega^r \cdot (\min X - 1)$ -large(θ)”.

Proof. Fix $r, b \in \mathbb{N}$ such that the first statement holds. Let X be ω^{2b+n_0+3} -large(θ) and $(x \mapsto x^{x^x})$ -sparse. Let us prove that X is OVW- $\omega^r \cdot (\min X - 1)$ -large(θ):

Consider $\ell, k < \min X$, an OVW X -tree T over an alphabet A of size k and a coloring $f : T \rightarrow \ell$. The set $X \setminus \{\min X\}$ is ω^{2b+n_0+2} -large(θ), so there exist pairwise θ -apart ω^{2b+n_0+1} -large(θ) subsets of $X \setminus (\min X \cup \min(X \setminus \{\min X\}))$, $X_0 < \dots < X_a$ with $a = \ell \times (\min X - 2) < (\min X)^2 < \min(X \setminus \{\min X\})$ (by $(x \mapsto x^{x^x})$ -sparsity). For simplicity, we will assume that T is an OVW $(X_0 \cup \dots \cup X_a)$ -tree (this can always be done by taking a subtree of the original tree).

Above every $\sigma \in A^{\min X_a} \cap T$ is an OVW X_a -subtree of T . The number n_a of such subtrees satisfies $n_a = k^{\text{card } X_0 + \dots + \text{card } X_{a-1}}$ (there is one subtree for each instantiation of the variables of $X_0 \cup \dots \cup X_{a-1}$), so $n_a \leq k^{\max X_{a-1}}$.

The coloring f induces a coloring on each of those trees, and since all these trees are isomorphic (not only isomorphic as trees, but they share the same structure as OVW trees: their generating words only differ for indexes less than $|\sigma|$), consider the product coloring $f' : T' \rightarrow \ell^{n_a}$ of all those colorings (for T' an arbitrary such X_a -subtree). By $(x \mapsto x^{x^x})$ -sparsity of X , $\ell^{n_a} \leq \ell^{k^{\max X_{a-1}}} < \min X_a$.

Since X_a is ω^{2b+n_0+1} -large(θ) and $(x \mapsto x^{x^x})$ -sparse, then it is OVW- ω^r -large(θ), so there exists some ω^r -large(θ) subset $Y_a \subseteq X_a$ and some f' -homogeneous OVW Y_a -subtree $S' \subseteq T'$. Therefore, by definition of f' , there exists an instantiation of the variables in $X_a \setminus Y_a$ making f -homogeneous the corresponding OVW Y_a -subtrees of T above each $\sigma \in A^{\min X_a} \cap T$. Let T'_a be the corresponding OVW $(X_0 \cup \dots \cup X_{a-1} \cup Y_a)$ -subtree of T . T'_a has the property that above every $\sigma \in T'_a \cap A^{\min X_a}$, there is a color $c_a^\sigma \in \ell$ such that all the τ in the OVW Y_a -tree above σ have that same color c_a^σ .

Let $T_a = T \cap A^{\leq \min X_a}$. Note that T_a is an OVW $(X_0 \cup \dots \cup X_{a-1})$ -subtree of T . Let S_a be the corresponding combinatorial $(X_0 \cup \dots \cup X_{a-1})$ -space, every element of S_a corresponds to one instantiation of the variables of $X_0 \cup \dots \cup X_{a-1}$ and therefore to exactly one of the OVW X_a -subtrees considered before. So, consider the coloring $f_a : S_a \rightarrow \ell$ sending any element of S_a to the color c_a^σ of its corresponding OVW X_a -subtree.

There are at most $k^{\text{card}(X_0 \cup \dots \cup X_{a-2})} \leq k^{\max X_{a-2}}$ combinatorial X_{a-1} -subspaces of S_a (one for each instantiation of the variables in $X_0 \cup \dots \cup X_{a-2}$) and f_a induces a coloring on each of them. So, by using the same trick of considering the product coloring, we can apply Lemma 15.2.7 to get an ω^b -large(θ) subset $Z_{a-1} \subseteq X_{a-1}$ such that there exists an instantiation of the variables in $X_{a-1} \setminus Z_{a-1}$ making f_a -homogeneous the corresponding combinatorial Z_{a-1} -subspaces of each combinatorial X_{a-1} -subspace of S_a .

Since Z_{a-1} is ω^b -large(θ), then it is OVW- ω^r -large(θ), so there exists some ω^r -large(θ) subset $Y_{a-1} \subseteq Z_{a-1}$ and some instantiation of the variables in $Z_{a-1} \setminus Y_{a-1}$ making f -homogeneous the corresponding OVW Y_{a-1} -subtrees of T_a above each $\sigma \in A^{\min X_{a-1}} \cap T_a$ (again by considering a product coloring).

Let T'_{a-1} be the corresponding OVW $(X_0 \cup \dots \cup X_{a-2} \cup Y_{a-1} \cup Y_a)$ -subtree of T . T'_{a-1} has the property that for every $\sigma \in T'_{a-1} \cap A^{\min X_{a-1}}$ there is a tuple of colors $(c_{a-1}^\sigma, c_a^\sigma)$ such that all the τ in the OVW $(Y_{a-1} \cup Y_a)$ -tree above σ have color c_{a-1}^σ if $|\tau| \in Y_{a-1}$ (by construction of Y_{a-1}) and color c_a^σ if $|\tau| \in Y_a$ (by construction of Z_{a-1}).

We can then iterate the construction and obtain a sequence of ω^r -large(θ) subsets $Y_i \subseteq X_i$ and of OVW $(X_0 \cup \dots \cup X_{i-1} \cup Y_i \cup \dots \cup Y_a)$ -subtrees T'_i of T , where each T'_i has the property that for every $\sigma \in T'_i \cap A^{\min X_i}$ there is a $(a-i+1)$ -tuple of colors $(c_i^\sigma, \dots, c_a^\sigma)$ such that all the τ in the OVW $(Y_i \cup \dots \cup Y_a)$ -tree above σ have color c_j^σ if $|\tau| \in Y_j$. At each step of the construction, we consider the tree $T_i = T \cap A^{\leq \min X_i}$, the corresponding combinatorial subspace S_i , a coloring $f_i : S_i \rightarrow \ell^{a+1-i}$ (The coloring f_i gives the tuple of colors corresponding to the OVW $(Y_{i+1} \cup \dots \cup Y_a)$ -subtree above every element of S_i), an f_i -homogeneous combinatorial Z_i -subspace of S_i for some $Z_i \subseteq X_i$ ω^b -large(θ), and finally an ω^r -large(θ) subset $Y_i \subseteq Z_i$ satisfying the desired properties and T'_i , the corresponding OVW $(X_0 \cup \dots \cup X_{i-1} \cup Y_i \cup \dots \cup Y_a)$ -subtree of T .

Therefore, the color of every $\tau \in T'_0$ is determined by the index $j \leq a$ such that $|\tau| \in Y_j$. By the finite pigeonhole principle, there is one color that appears at least $(\min X - 1)$ times, since $a+1 = \ell \times (\min X - 2) + 1$. Let $i_0 < \dots < i_{\min X - 2} \leq a$ be indices witnessing the pigeonhole principle, and let $Y = Y_{i_0} \cup \dots \cup Y_{i_{\min X - 2}}$. The set Y is $\omega^r \cdot (\min X - 1)$ -large(θ) and any OVW Y -subtree of T'_0 is f -homogeneous. $\square$

Lemma 15.3.7. $|\Sigma_1^0$ proves that for every $b, r \in \mathbb{N}$, if the following statement holds:

- “Every ω^b -large(θ) $(x \mapsto x^{x^x})$ -sparse set X is $\text{OVW-}\omega^r \cdot (\min X - 1)$ -large(θ)”

then the following statement holds:

- “Every ω^{2b+n_0+2} -large(θ) $(x \mapsto x^{x^x})$ -sparse set is $\text{OVW-}\omega^{r+1}$ -large(θ)”.

Proof. Fix $r, b \in \mathbb{N}$ such that the first statement holds. Let X be ω^{2b+n_0+2} -large(θ) and $(x \mapsto x^{x^x})$ -sparse. Let us prove that X is $\text{OVW-}\omega^{r+1}$ -large(θ):

Consider $\ell, k < \min X$, an OVW X -tree T over an alphabet A of size k and a coloring $f : T \rightarrow \ell$.

Let $X_0 < \dots < X_\ell$ be ω^{2b+n_0+1} -large(θ) subsets of X . By the same construction as in the proof of Lemma 15.3.6, but replacing $\text{OVW-}\omega^r$ -largeness(θ) by $\text{OVW-}\omega^r \cdot (\min X_i - 1)$ -largeness(θ), we obtain for each $i \leq \ell$ an $\omega^r \cdot (\min X_i - 1)$ -large(θ) subset $Y_i \subseteq X_i$, and an OVW $(Y_0 \cup \dots \cup Y_\ell)$ -tree T'_0 such that the color of every $\rho \in T'_0$ is determined by the index $j \leq \ell$ such that $|\rho| \in Y_j$. Since $\ell + 1 > \ell$, there are indices $i < j$ such that all the $\rho \in T'_0$ with $|\rho| \in Y_i \cup Y_j$ have the same color. Take some $y \in Y_i$, since $y \leq \min Y_j - 1$ and Y_j is $\omega^r \cdot (\min X_j - 1)$ -large(θ), then $\{y\} \cup Y_j$ is ω^{r+1} -large(θ) and any OVW $(\{y\} \cup Y_j)$ -subtree of T'_0 is f -homogeneous. $\square$

Lemma 15.3.8. $|\Sigma_1^0$ proves that for every $b, r \in \mathbb{N}$, if the following statement holds:

- “Every ω^b -large(θ) $(x \mapsto x^{x^x})$ -sparse set is $\text{OVW-}\omega^r$ -large(θ)”

then the following statement holds:

- “Every ω^{4b+3n_0+8} -large(θ) $(x \mapsto x^{x^x})$ -sparse set is $\text{OVW-}\omega^{r+1}$ -large(θ)”.

Proof. Immediate by Lemma 15.3.6 and Lemma 15.3.7. $\square$

Corollary 15.3.9. $|\Sigma_1^0$ proves that for every $b \in \mathbb{N}$, every $\omega^{O(4^b)}$ -large(θ) $(x \mapsto x^{x^x})$ -sparse set is $\text{OVW-}\omega^b$ -large(θ).

Proof. The $O(4^b)$ notation can actually be replaced by a complicated explicit bound $p(b)$, computed from the bounds in Lemma 15.3.5, Lemma 15.3.7, and Lemma 15.3.8. One can prove the explicit bound version of the statement by a simple Π_1^0 -induction over b , using Lemma 15.3.5 and Lemma 15.3.7 for the base case, and Lemma 15.3.8 for the induction step. The inductive formula is of the form “for every finite set F which is $\omega^{p(b)}$ -large(θ) and $(x \mapsto x^{x^x})$ -sparse, F is $\text{OVW-}\omega^b$ -large(θ)”. Since being $\omega^{p(b)}$ -large(θ), $(x \mapsto x^{x^x})$ -sparse and $\text{OVW-}\omega^b$ -large(θ) are Σ_0^0 properties, the overall formula is Π_1^0 . $\square$

Theorem 15.1.3. $\text{WKL}_0 + \text{OVW}$ is $\forall\Pi_4^0$ -conservative over $\text{RCA}_0 + \text{B}\Sigma_2^0$.

Proof. As stated in the introduction of this chapter, OVW is not an RT-like statement, and we cannot apply Theorem 12.3.7 with the bounds obtained in Corollary 15.3.9. Nevertheless, a similar argument can be applied in our case, as it was done in [87, Theorem 1.4]. $\square$

In particular, OVW is finitistically reducible.

Corollary 15.3.10. $\text{WKL}_0 + \text{OVW}$ is $\forall\Pi_3^0$ -conservative over RCA_0 and Π_2 -conservative over PRA.

And we obtain the following separation result.

Corollary 15.3.11. $\text{WKL}_0 + \text{OVW}$ does not imply ACA_0 .

Index

Symbols

A'	37	Σ_n^0	43, 103
$A'[s]$	37	$\Sigma_n^0(C_0, \dots)$	103
$A^{(\omega)}$	38	$\Vdash$	113, 116, 119
$A^{(n)}$	38	$\Vdash_i$	124
$A^{<\mathbb{N}}$	35	α^*	58
$A^{<n}$	35	$\mathfrak{C}_0$	101
$A^{\leq n}$	35	ω	101
$[A]^n$	53	$\bigoplus$	37, 111
$[A]^{\omega}$	81	$\equiv_T$	37
$[A]^{<\mathbb{N}}$	81	$\subseteq_{\text{fin}}$	36
$[A]^{\leq \omega}$	91	$\forall \Gamma$	146
$[T]$	39	$\forall \Pi_4^0$	51
$[\sigma]$	40	$\forall \Pi_5^0$	108
$[b]^n$	53	$\forall^\infty$	119
$\Delta_n(C_0, \dots)$	103	$\langle \dots \rangle$	28
$\Delta_n(M)$	103	$\leq_T$	37
$\Delta_n(\mathcal{M})$	103	$\leq_{\text{lex}}$	90
Δ_n^0	103	$\leq_{\text{SW}}$	84
$\Delta_n^0(C_0, \dots)$	103	$\ll$	41, 156
$\mathcal{M}[G]$	105	0	37
Φ_e	35	ω_n	33
$\Phi_e(x) \downarrow = v$	35	ω_n^α	33
$\Phi_e(x) \downarrow$	35	$\ominus$	64
$\Phi_e(x) \uparrow$	35	$\oplus$	37, 64, 111
$\Phi_e(x)[s] \downarrow$	36	$\otimes$	111
$\Phi_e(x)[s] \uparrow$	36	$\bar{p}$	64
$\Phi_e(x)[s]$	36	$\leq$	36
Φ_e^A	36	$? \vdash$	94, 124
$\Pi_1^1\text{-CA}_0$	49	$\star$	179
Π_n	28	$\subseteq^*$	58
$\Pi_n(C_0, \dots)$	103	$\uparrow$	37
$\Pi_n(M)$	103	$\models$	45, 46
$\Pi_n(X_0, \dots)$	43	ε_0	33
$\Pi_n(\mathcal{M})$	103	$\{\alpha\}(X)$	157
Π_n^0	43, 103	$\{\alpha\}(x)$	156
$\Pi_n^0(C_0, \dots)$	103	h^A	162
$\Rightarrow_x$	158	$h_\alpha(x) \downarrow$	157
Σ_n	28	$h_\alpha(x) \uparrow$	157
$\Sigma_n(C_0, \dots)$	103	$w[a]$	180
$\Sigma_n(M)$	103	$w[u]$	181
$\Sigma_n(X_0, \dots)$	43	$2_n(a)$	136
$\Sigma_n(\mathcal{M})$	103		

A

ACA_0	49
ACA_0^+	83
ACA'_0	96
ADS.....	56
apart	
θ -.....	151
<i>ARITH</i>	49
arithmetic	
first-order.....	27
Peano.....	27
Robinson.....	29
second-order.....	43
true.....	39
ATR_0	49

B

B -block Z -immune with cost n ...	136
b -bounded.....	109
b -constrained.....	91
bad	
(f, k) -.....	73
f -.....	71
basis.....	69
$BC\Delta_n$	32
$BC\Pi_n$	32
$BC\Sigma_n$	32
Big Five.....	49
block.....	73
BME.....	109
$BME_\star$	109
bounded	
b -.....	109
k -.....	79
comprehension.....	32
formula.....	27
induction.....	27
M -.....	48
monotone enumeration.....	109
quantifier.....	27
set.....	48
$B\Pi_n$	31
$B\Pi_n^0$	45
BRT_2^2	70, 131
$B\Sigma_n$	31
$B\Sigma_n^0$	45

C

c.e.....	36
Cantor normal form.....	33, 156
short.....	156
Cantor space.....	36

Catalan numbers.....	87
class	
Π_1^0	40
clique.....	55
C_n	87
coded	
M -.....	146
$Cod(M/I)$	146
cofinal.....	132
COH.....	58
cohesive.....	58
$\vec{R}$ -.....	58
p -.....	58
set principle.....	58
collection.....	31
coloring	
b -constrained.....	91
k -constrained.....	91
limit.....	93
progressive.....	91
stable.....	57
transitive.....	56
combinatorial	
Y -line.....	181
Y -space.....	181
subspace.....	181
combinatorial Y -space	
dimension.....	181
<i>COMP</i>	48
compatible	
for COH.....	113
for SADS.....	124
for SEM.....	119
for WKL.....	116
complete	
Σ_1 -.....	30
theory.....	20
comprehension	
Δ_1^0 -.....	48
Π_1^1 -.....	49
arithmetical.....	49
bounded.....	32
scheme.....	48
computable	
function.....	35
relative to a set.....	36
set.....	36
computably enumerable.....	36
approximation.....	37
computably true.....	54
Con.....	20
condition.....	91
extension .. 92, 113, 116, 119, 124	
for COH.....	113
for SADS.....	124
for SEM.....	119

for WKL	115
good	91
Mathias	113
cone avoidance	80
strong	80
conservation	21
consistent theory	20
constrained	
b -	91
k -	91
$\text{coRT}_2^2(p)$	63
$\text{coSRT}_2^2(p)$	64
$CS(k, d, \ell)$	183
cut	104, 146
proper	104, 146
semi-regular	147

D

$\mathcal{D}$	37
D_2^2	57
decidable theory	20
degree	
PA	39
DNC	66
DNC_2	40
low	41
low_n	41
Turing	37
dimension	
of a k -fractal	69
of a combinatorial Y -space	181
of an OVW Y -tree	183
direct sum	64
DNC	
degree	66
function	66
DNC_2	
degree	40
function	40
dominate	38

E

effective immunity	136
effectively closed	40
elementary	29
EM	56
exactly $(\omega + 1)$ -large	89
exactly ω -large	81
extends	119
extension	92
ω -	102
external induction	101

F

f -avoid	63
f -bad	71
f -free	83
f -good	71
f -rainbow	83
f -realize	63
f -thin	83
fallow	
Erdős-Moser theorem	167
f -	167
set	167
fEM	167
fEM_k	167
f_G	148
finitistically reducible	22
first-order	
arithmetic	27
model	45
part	45, 46
PRA	27
structure	45
$\mathcal{F}_k$	69
force	94
forcing	
question for $\text{PFS}^{\text{!}\omega}$	94
question for SADS	124
relation for COH	113
relation for SADS	124
relation for SEM	119
relation for WKL	116
set theory	22
formula	
Π_n	28
$\Pi_n(X_0, \dots)$	43
Π_n^0	43
Π_n^1	44
Σ_n	28
$\Sigma_n(X_0, \dots)$	43
Σ_n^0	43
Σ_n^1	44
$\forall\Gamma$	146
$\forall\Pi_4^0$	51
$\forall\Pi_5^0$	108
r - Π_2^1	58
arithmetical	27
bounded	27
universal Π_n	38
universal Σ_n	38
fractal	
decomposition	73
dimension	69
k -	69
free	
f -	83

set	79
set theorem	79
$FS^{<\mathbb{N}}$	81
$FS^{! \omega}$	83
FS^n	79
function	
Z -escaping	66
DNC	66
cofinal	132
hyperimmune	65
primitive recursive	23
fundamental sequence	156

G

good	
(f, k) -	73
condition	91
f -	71
$GR\text{-}\alpha\text{-large}(\theta)$	182
$GR(k, d, m, \ell)$	182
g -stabilize	92

H

h_α	157
Henkin	
semantics	46
structure	46
Hilbert's program	19
$HJ(k, \ell)$	181
homogeneous	53
hyperimmune	
function	65

I

ideal	
jump	49
Scott	49
Turing	48
$I\Delta_n$	31
$I\Delta_n^0$	45
immune	136
B -block Z -immune with cost n	
136	
Z -immune	136
h -effectively Z -immune	136
independence	20
induction	
bounded	27
external	101
first-order	30
internal	101

quantifier-free	25
second-order	45
internal induction	101
intuitionism	19, 25
Π_n	30
Π_n^0	45
Σ_n	30
Σ_n^0	45

J

join	37
jump	37

K

k -bounded	79
k -constrained	91
k -fractal	69
of dimension n	69
Kleene normal form	
for Π_1^0 -formulas	44
for Σ_1^0 -formulas	44
for arithmetical formulas	44

L

$\mathcal{L}_g$	166
largeness	
$GR\text{-}\alpha\text{-large}(\theta)$	182
$\mathcal{L}$ -large	147
$OVW\text{-}\alpha\text{-large}(\theta)$	184
$P\text{-}\mathcal{L}$ -large	148
$P\text{-}\alpha\text{-large}(\theta)$	151
α -large	148
$\alpha\text{-large}(\theta)$	151
α -small	157
ω - g -large	170
ω -large	145
ω -large	81
$m\text{-dense}(\theta, P_0, \dots, P_i)$	151
$m\text{-dense}(P_0, \dots, P_{i-1})$	149
at most α -large	155, 157
exactly $(\omega + 1)$ -large	89
exactly α -large	157
exactly ω -large	81
notion	147
quasi-exactly ω -large	81
left variable word	180
lemma	
Shoenfield's limit	38
weak König's	39
limit coloring	93
$\mathcal{L}_{Z_2}$	43

M

M -coded	146
m -dense($P_0, \dots, P_{i-1}$)	149
Mathias condition	113
maximal	
(σ, a) -tree	120
minimal interval	119
model	
ω -model	46
first-order	45
non-standard	46
of a sentence	45, 46
of a theory	45, 46
second-order	46
topped	106
modulus	37
monotone enumeration	109
b -bounded	109
bounded	109
MTT	179

N

natural	
product	111
sum	111
node	
(σ, a) -	120
non-standard	
element	46
formula	102
model	46

O

ordinal	32
natural product	111
natural sum	111
notation	33
overspill principle	104
OVW	
Y -line	183
Y -tree	183
Y -tree dimension	183
subtree	183
OVW	180
OVW- α -large(θ)	184

P

PA	27
degree	39
over	41

relative to	41
theory	27
PA ⁻	30
pattern	63
RT ₂ ²	63
dual	64
size	63
permutation	64
separable	65
PFS _k ^{!ω}	91
PFS _k ^{!ω}	91
PH	145
piecewise-coded	104
pigeonhole principle	
for α -large sets	163, 164
for Δ_2^0 sets	57
infinite	53
P- $\mathcal{L}$ -large	148
PRA	20
first-order	27
quantifier-free	24
preservation of ω hyperimmunities ..	65
primitive recursive	
arithmetic	20
function	23
relation	23
set	23
principle	
ascending/descending sequence ..	56
bounded monotone enumeration	
109	
cohesive set	58
infinite pigeonhole	53
Paris-Harrington	145
stable ascending/descending	
sequence	58
problem	
Π_2^1	44
RT-like	148
instance	44
solution	44
progressive	91
free set theorem	91
proof	
size	21
proper cut	104
pseudo-norm	156
psn(α)	156

Q

quantifier	
-free induction	25
bounded	27
quasi-exactly ω -large	81

R

$\vec{R}$ -cohesive	58
rainbow	
f -.....	83
Ramsey theorem	79
set	79
Ramsey theorem	
finite	53
infinite	53
RCA_0	48
RCA_0^*	104
realize	
f -.....	63
relativization	39
reservoir	91
reverse mathematics	47
$R_k(d)$	53
$r\text{-}\Pi_2^1$	58
$\text{RRT}_k^{<\mathbb{N}}$	81
$\text{RRT}_k^{l\omega}$	83
$\text{RRT}_k^{l\omega}$	83
$\text{RRT}_2^{l(\omega+1)}$	89
RRT_k^n	79
RT	53
$\text{RT}_2^2(p)$	63
$\text{RT}^{<\mathbb{N}}$	81
$\text{RT}_k^{l\omega}$	82
RT_k^n	53
RT_k^n	53

S

SADS	58
satisfiability	
for (σ, a) -node	120
for (σ, a) -tree	120
for (σ^0, σ^1) -tree	125
for node	125
scheme	
BC Γ (bounded comprehension)	32
B Γ (collection)	31
I Γ (induction)	30
comprehension	48
second-order	
arithmetic	20, 43
formula	43
model	46
part	46
quantifier	43
structure	46
theory	45
SEM	58
semantic	

full	46
Henkin	46
semi-regular cut	147
separable permutation	65
set	
B -block Z -immune with cost n	
136	
Δ_n	38
$\Delta_n(C_0, \dots)$	43, 103
$\Delta_n(M)$	103
$\Delta_n(\mathcal{M})$	103
Δ_n^0	103
$\Delta_n^0(C_0, \dots)$	103
Π_n	38
$\Pi_n(C_0, \dots)$	43, 103
$\Pi_n(M)$	103
$\Pi_n(\mathcal{M})$	103
Π_n^0	103
$\Pi_n^0(C_0, \dots)$	103
Σ_n	38
$\Sigma_n(C_0, \dots)$	43, 103
$\Sigma_n(M)$	103
$\Sigma_n(\mathcal{M})$	103
Σ_n^0	103
$\Sigma_n^0(C_0, \dots)$	103
f -fallow	167
h -effectively Z -immune	136
m -dense($P_0, \dots, P_{i-1}$)	149
arithmetical	38
finite	32
free	83
immune	136
M -bounded	48
M -coded	146
M -finite	48
$\mathcal{M}$ -infinite	48
M -unbounded	48
piecewise-coded	104
quantifier	43
rainbow	79
thin	79
variable	43
set theory	19
Shore blocking	111
short Cantor normal form	156
skew sum	64
sparse	
$\mathcal{L}$ -sparse	165
α -sparse	165
g -sparse	166
split pair	124
SRT_k^2	57
stabilize	
g -	92
statement	
ADS	56

- BME 109
 BME_★ 109
 BRT₂² 70, 131
 CH 22
 COH 58
 D₂² 57
 EM 56
 FSⁿ 79
 FS^{l_ω} 83
 FS^{<N} 81
 MTT 179
 OVW 180
 PFS^{l_ω} 91
 PFS^{l_ω}_k 91
 PH 145
 RRTⁿ 79
 RRT^{l_ω}_{2^(ω+1)} 89
 RRT^{l_ω} 83
 RRT^{<N}_k 81
 RRTⁿ_k 79
 RRT^{l_ω}_k 83
 RT 53
 RT₂²(p) 63
 RTⁿ 53
 RTⁿ_k 53
 RT^{l_ω}_k 82
 RT^{<N} 81
 SADS 58
 SEM 58
 SRT_k² 57
 TSⁿ 79
 TSⁿ_k 87
 TS^{l_ω} 83
 TS^{<N} 81
 TT 179
 TT¹ 179
 WF⁰(α) 45
 WKL 39
 coRT₂²(p) 63
 coSRT₂²(p) 64
 fEM 167
 fEM_k 167
 trRT_k² 167
 exp 30
 stem 91
 strong cone avoidance 80
 strong Weihrauch reducibility 84
 sub-pattern 63
 submodel
 ω- 102
- T**
 theorem
 amalgamation 177
- bounded Ramsey 70, 131
 Cantor normal form 33
 compactness 47
 completeness 47
 cone avoidance basis 40
 Erdős-Moser 56
 fallow Erdős-Moser 167
 free set 79
 Graham-Rothschild 182
 Gödel's incompleteness 21
 Hales-Jewett 181
 large free set 83
 large rainbow Ramsey 83
 large Ramsey 82
 large thin set 83
 low basis 41
 Löwenheim-Skolem 47
 Milliken's tree 179
 ordered variable word 180
 Post 38
 progressive free set 91
 rainbow Ramsey 79
 Ramsey (finite version) 53
 Ramsey (infinite version) 53
 soundness 47
 stable Erdős-Moser 58
 stable Ramsey 57
 thin set 79
 thin set for finite colorings 87
 transitive Ramsey 167
 tree 179
- theory
 ACA₀ 49
 ACA₀⁺ 83
 ATR₀ 49
 PA 27
 PA⁻ 30
 PRA 20
 Π₁¹-CA₀ 49
 RCA₀ 48
 RCA₀^{*} 104
 WKL₀ 49
 WKL₀^{*} 107
 ZF 22
 ZFC 20
 ACA₀['] 96
 Q 29
 Z₂ 20
 complete 20
 consistent 20
 decidable 20
- thin
 f- 83
 set 79
 set theorem 79
 topped 106

tournament	56
transitive	
coloring	56
set	56
tournament	56
tree	
(σ, a) -	120
(σ^0, σ^1) -	125
OVW Y -tree	183
partition theorem	179
theorem	179
trRT_k^2	167
$\text{TS}_{<\mathbb{N}}$	81
$\text{TS}^{! \omega}$	83
TS^n	79
TS_k^n	87
TT	179
TT^1	179
Turing	
degree	37
equivalence	37
functional	36
ideal	48
index	35
jump	37
reducibility	37

U

unbounded	
M -	48

universal	
(σ, a) -tree	120
Π_n -formula	38
Σ_n -formula	38
use property	37

V

valid	
(σ, a) -node	120
node	125

W

W_e	36
$\text{WF}^0(\alpha)$	45
$\text{WF}(\alpha)$	33
WKL	39
WKL_0	49
WKL_0^*	107
word	179
generating variable	181, 183
left variable	180
ordered Y -variable	181

Z

Z_2	20
ZF	22
ZFC	20

Bibliography

- [1] Paul-Elliot Anglès d'Auriac, Peter A. Cholak, Damir D. Dzhafarov, Benoît Monin, and Ludovic Patey. Milliken's tree theorem and its applications: a computability-theoretic perspective. *Mem. Amer. Math. Soc.*, 293(1457):vi+118, 2024.
- [2] Paul-Elliot Anglès d'Auriac, Lu Liu, Bastien Mignoty, and Ludovic Patey. Carlson-Simpson's lemma and applications in reverse mathematics. *Ann. Pure Appl. Logic*, 174(9):Paper No. 103287, 16, 2023.
- [3] Eric P. Astor, Laurent Bienvenu, Damir Dzhafarov, Ludovic Patey, Paul Shafer, Reed Solomon, and Linda Westrick. The weakness of typicality. *arXiv preprint arXiv:2609.00884*, 2026.
- [4] Jeremy Avigad. Formalizing forcing arguments in subsystems of second-order arithmetic. *Ann. Pure Appl. Logic*, 82(2):165–191, 1996.
- [5] Jeremy Avigad. *Mathematical logic and computation*. Cambridge University Press, Cambridge, 2023.
- [6] David R Belanger. Conservation theorems for the Cohesiveness Principle. *arXiv preprint arXiv:2212.13011*, 2022.
- [7] Laurent Bienvenu, Ludovic Patey, and Paul Shafer. On the logical strengths of partial solutions to mathematical problems. *Trans. London Math. Soc.*, 4(1):30–71, 2017.
- [8] Teresa Bigorajska and Henryk Kotlarski. A partition theorem for α -large sets. *Fund. Math.*, 160(1):27–37, 1999.
- [9] Teresa Bigorajska and Henryk Kotlarski. Partitioning α -large sets: some lower bounds. *Trans. Amer. Math. Soc.*, 358(11):4981–5001, 2006.
- [10] Prosenjit Bose, Jonathan F. Buss, and Anna Lubiw. Pattern matching for permutations. *Inform. Process. Lett.*, 65(5):277–283, 1998.
- [11] Andrey Bovykin and Andreas Weiermann. The strength of infinitary Ramseyan principles can be accessed by their densities. *Ann. Pure Appl. Logic*, 168(9):1700–1709, 2017.
- [12] Patrick Caldon and Aleksandar Ignjatović. On mathematical instrumentalism. *J. Symbolic Logic*, 70(3):778–794, 2005.
- [13] Timothy J. Carlson and Stephen G. Simpson. A dual form of Ramsey's theorem. *Adv. in Math.*, 53(3):265–290, 1984.
- [14] Lorenzo Carlucci, Oriola Gjetaj, Quentin Le Houérou, and Ludovic Levy Patey. Ramsey-like theorems for the Schreier barrier. *The Journal of Symbolic Logic*, page 1–29, 2025.

- [15] Lorenzo Carlucci and Konrad Zdanowski. The strength of Ramsey’s theorem for coloring relatively large sets. *J. Symb. Log.*, 79(1):89–102, 2014.
- [16] Julien Cervelle, William Gaudelier, and Ludovic Patey. The reverse mathematics of CAC for trees. *J. Symb. Log.*, 89(3):1189–1211, 2024.
- [17] Peter Cholak and Ludovic Patey. Thin set theorems and cone avoidance. *Trans. Amer. Math. Soc.*, 373(4):2743–2773, 2020.
- [18] Peter A. Cholak, Mariagnese Giusto, Jeffrey L. Hirst, and Carl G. Jockusch, Jr. Free sets and reverse mathematics. In *Reverse mathematics 2001*, volume 21 of *Lect. Notes Log.*, pages 104–119. Assoc. Symbol. Logic, La Jolla, CA, 2005.
- [19] Peter A. Cholak, Carl G. Jockusch, and Theodore A. Slaman. On the strength of Ramsey’s theorem for pairs. *The Journal of Symbolic Logic*, 66(1):1–55, 2001.
- [20] C. T. Chong, Steffen Lempp, and Yue Yang. On the role of the collection principle for Σ_2^0 -formulas in second-order reverse mathematics. *Proc. Amer. Math. Soc.*, 138(3):1093–1100, 2010.
- [21] C. T. Chong, Wei Li, Wei Wang, and Yue Yang. On the strength of Ramsey’s theorem for trees. *Adv. Math.*, 369:107180, 39, 2020.
- [22] C. T. Chong, Theodore A. Slaman, and Yue Yang. Π_1^1 -conservation of combinatorial principles weaker than Ramsey’s theorem for pairs. *Adv. Math.*, 230(3):1060–1077, 2012.
- [23] C. T. Chong, Theodore A. Slaman, and Yue Yang. The inductive strength of Ramsey’s Theorem for Pairs. *Adv. Math.*, 308:121–141, 2017.
- [24] Chi Tat Chong, Wei Wang, and Yue Yang. Conservation strength of the infinite pigeonhole principle for trees. *Israel Journal of Mathematics*, pages 1–24, 2023.
- [25] Jennifer Chubb, Jeffrey L. Hirst, and Timothy H. McNicholl. Reverse mathematics, computability, and partitions of trees. *J. Symbolic Logic*, 74(1):201–215, 2009.
- [26] Peter Clote. A recursion theoretic analysis of the clopen Ramsey theorem. *The Journal of symbolic logic*, 49(2):376–400, 1984.
- [27] Peter Clote, Petr Hájek, and Jeff B. Paris. On some formalized conservation results in arithmetic. *Archive for Mathematical Logic*, 30:201–218, 1990.
- [28] Paul Cohen. The independence of the continuum hypothesis. *Proc. Nat. Acad. Sci. U.S.A.*, 50:1143–1148, 1963.
- [29] Chris J Conidis. Computability and combinatorial aspects of minimal prime ideals in noetherian rings, 2024.
- [30] Jared Corduan, Marcia J. Groszek, and Joseph R. Mileti. Reverse mathematics and Ramsey’s property for trees. *J. Symbolic Logic*, 75(3):945–954, 2010.
- [31] William Craig. On axiomatizability within a system. *J. Symbolic Logic*, 18:30–32, 1953.
- [32] Barbara F. Csima and Joseph R. Mileti. The strength of the rainbow Ramsey theorem. *J. Symbolic Logic*, 74(4):1310–1324, 2009.
- [33] Denis Campau Devlin. *Some partition theorems and ultrafilters on omega*. ProQuest LLC, Ann Arbor, MI, 1980. Thesis (Ph.D.)—Dartmouth College.

- [34] Pandelis Dodos and Vassilis Kanellopoulos. *Ramsey theory for product spaces*, volume 212 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2016.
- [35] Pandelis Dodos, Vassilis Kanellopoulos, and Konstantinos Tyros. A density version of the Carlson-Simpson theorem. *J. Eur. Math. Soc. (JEMS)*, 16(10):2097–2164, 2014.
- [36] François G. Dorais, Damir D. Dzhafarov, Jeffrey L. Hirst, Joseph R. Mileti, and Paul Shafer. On uniform relationships between combinatorial problems. *Trans. Amer. Math. Soc.*, 368(2):1321–1359, 2016.
- [37] Rodney G. Downey and Denis R. Hirschfeldt. *Algorithmic randomness and complexity*. Theory and Applications of Computability. Springer, New York, 2010.
- [38] Damir D. Dzhafarov. Strong reductions between combinatorial principles. *J. Symb. Log.*, 81(4):1405–1431, 2016.
- [39] Damir D. Dzhafarov and Carl Mummert. *Reverse mathematics—problems, reductions, and proofs*. Theory and Applications of Computability. Springer, Cham, [2022] ©2022.
- [40] Damir D. Dzhafarov and Ludovic Patey. Coloring trees in reverse mathematics. *Adv. Math.*, 318:497–514, 2017.
- [41] P. Erdős and L. Moser. On the representation of directed graphs as unions of orderings. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 9:125–132, 1964.
- [42] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.
- [43] V. Farmaki and S. Negrepontis. Schreier sets in Ramsey theory. *Trans. Amer. Math. Soc.*, 360(2):849–880, 2008.
- [44] António M. Fernandes, Fernando Ferreira, and Gilda Ferreira. Analysis in weak systems. In *Logic and computation*, volume 33 of *Tributes*, pages 231–261. Coll. Publ., [London], 2017.
- [45] Fernando Ferreira and Gilda Ferreira. Harrington’s conservation theorem redone. *Arch. Math. Logic*, 47(2):91–100, 2008.
- [46] Marta Fiori-Carones, Leszek Aleksander Kołodziejczyk, Tin Lok Wong, and Keita Yokoyama. An isomorphism theorem for models of weak König’s lemma without primitive recursion. *J. Eur. Math. Soc. (JEMS)*, 28(7):3203–3233, 2026.
- [47] Harvey Friedman and Stephen G. Simpson. Issues and problems in reverse mathematics. In *Computability theory and its applications (Boulder, CO, 1999)*, volume 257 of *Contemp. Math.*, pages 127–144. Amer. Math. Soc., Providence, RI, 2000.
- [48] Harvey M. Friedman. Systems on second order arithmetic with restricted induction i, ii. *J. Symb. Logic*, 41:557–559, 1976.
- [49] Gerhard Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Math. Ann.*, 112(1):493–565, 1936.
- [50] Kurt Gödel. Die Vollständigkeit der Axiome des logischen Funktionenkalküls. *Monatshefte für Mathematik und Physik*, 37:349–360, 1930.
- [51] Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme i. *Monatshefte für Mathematik und Physik*, 38:173–198, 1931.

- [52] Kurt Gödel. Über die Länge von Beweisen. *Ergebnisse eines mathematischen Kolloquiums*, 7:23–24, 1936.
- [53] Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer. *Ramsey theory*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, 2013. Paperback edition of the second (1990) edition [MR1044995].
- [54] Noam Greenberg and Joseph S. Miller. Lowness for Kurtz randomness. *J. Symbolic Logic*, 74(2):665–678, 2009.
- [55] Marcia Groszek and Theodore Slaman. On Turing reducibility, 1994.
- [56] Petr Hájek. Interpretability and fragments of arithmetic. In *Arithmetic, proof theory, and computational complexity (Prague, 1991)*, volume 23 of *Oxford Logic Guides*, pages 185–196. Oxford Univ. Press, New York, 1993.
- [57] Petr Hájek and Pavel Pudlák. *Metamathematics of first-order arithmetic*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1998. Second printing.
- [58] A. W. Hales and R. I. Jewett. Regularity and positional games. *Trans. Amer. Math. Soc.*, 106:222–229, 1963.
- [59] Godfrey H. Hardy. A theorem concerning the infinite cardinal numbers. *Quarterly journal of Mathematics*, 35:87–94, 1904.
- [60] Leon Henkin. Completeness in the theory of types. *J. Symbolic Logic*, 15:81–91, 1950.
- [61] Denis R. Hirschfeldt. *Slicing the truth*, volume 28 of *Lecture Notes Series. Institute for Mathematical Sciences. National University of Singapore*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2015. On the computable and reverse mathematics of combinatorial principles, Edited and with a foreword by Chitát Chong, Qi Feng, Theodore A. Slaman, W. Hugh Woodin and Yue Yang.
- [62] Denis R. Hirschfeldt, Carl G. Jockusch, Jr., Bjørn Kjos-Hanssen, Steffen Lempp, and Theodore A. Slaman. The strength of some combinatorial principles related to Ramsey’s theorem for pairs. In *Computational prospects of infinity. Part II. Presented talks*, volume 15 of *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.*, pages 143–161. World Sci. Publ., Hackensack, NJ, 2008.
- [63] Denis R. Hirschfeldt and Richard A. Shore. Combinatorial principles weaker than Ramsey’s theorem for pairs. *J. Symbolic Logic*, 72(1):171–206, 2007.
- [64] Jeffry Lynn Hirst. *Combinatorics in subsystems of second order arithmetic*. ProQuest LLC, Ann Arbor, MI, 1987. Thesis (Ph.D.)—The Pennsylvania State University.
- [65] Jan Hubička. Big Ramsey degrees using parameter spaces. *Adv. Math.*, 478:Paper No. 110386, 27, 2025.
- [66] Aleksandar Djordje Ignjatovic. *Fragments of first and second-order arithmetic and length of proofs*. ProQuest LLC, Ann Arbor, MI, 1990. Thesis (Ph.D.)—University of California, Berkeley.
- [67] Hiroyuki Ikari. *Ramsey’s theorem in systems of arithmetic and proof transformations*. PhD thesis, Tohoku University, 2024.

- [68] C. G. Jockusch, Jr., M. Lerman, R. I. Soare, and R. M. Solovay. Recursively enumerable sets modulo iterated jumps and extensions of Arslanov's completeness criterion. *J. Symbolic Logic*, 54(4):1288–1323, 1989.
- [69] Carl Jockusch and Frank Stephan. A cohesive set which is not high. *Math. Logic Quart.*, 39(4):515–530, 1993.
- [70] Carl G. Jockusch, Jr. Ramsey's theorem and recursion theory. *J. Symbolic Logic*, 37:268–280, 1972.
- [71] Carl G. Jockusch, Jr. and Robert I. Soare. Π_1^0 classes and degrees of theories. *Trans. Amer. Math. Soc.*, 173:33–56, 1972.
- [72] Jussi Ketonen and Robert Solovay. Rapidly growing Ramsey functions. *Ann. of Math. (2)*, 113(2):267–314, 1981.
- [73] L. A. S. Kirby and J. B. Paris. Initial segments of models of Peano's axioms. In *Set theory and hierarchy theory, V (Proc. Third Conf., Bierutowice, 1976)*, Lecture Notes in Math., Vol. 619, pages 211–226. Springer, Berlin, 1977.
- [74] Bjørn Kjos-Hanssen. Infinite subsets of random sets of integers. *Math. Res. Lett.*, 16(1):103–110, 2009.
- [75] Bjørn Kjos-Hanssen, Wolfgang Merkle, and Frank Stephan. Kolmogorov complexity and the recursion theorem. *Trans. Amer. Math. Soc.*, 363(10):5465–5480, 2011.
- [76] S. C. Kleene. Recursive predicates and quantifiers. *Transactions of the American Mathematical Society*, 53(1):41–73, 1943.
- [77] Stephen Cole Kleene. *Introduction to metamathematics*. D. Van Nostrand Co., Inc., New York, 1952.
- [78] Leszek Aleksander Kołodziejczyk, Tin Lok Wong, and Keita Yokoyama. Ramsey's theorem for pairs, collection, and proof size. *Journal of Mathematical Logic*, page 2350007, 2023.
- [79] Leszek Aleksander Kołodziejczyk and Keita Yokoyama. Some upper bounds on ordinal-valued Ramsey numbers for colourings of pairs. *Selecta Math. (N.S.)*, 26(4):Paper No. 56, 18, 2020.
- [80] Leszek Aleksander Kołodziejczyk and Keita Yokoyama. In search of the first-order part of Ramsey's theorem for pairs. In *Connecting with computability*, volume 12813 of *Lecture Notes in Comput. Sci.*, pages 297–307. Springer, Cham, [2021] ©2021.
- [81] Henryk Kotlarski. *A model-theoretic approach to proof theory*, volume 51 of *Trends in Logic—Studia Logica Library*. Springer, Cham, 2019. Edited by Zofia Adamowicz, Teresa Bigorajska and Konrad Zdanowski.
- [82] Henryk Kotlarski, Bożena Piekart, and Andreas Weiermann. More on lower bounds for partitioning α -large sets. *Ann. Pure Appl. Logic*, 147(3):113–126, 2007.
- [83] Alexander P. Kreuzer. Primitive recursion and the chain antichain principle. *Notre Dame J. Form. Log.*, 53(2):245–265, 2012.
- [84] Alexander P. Kreuzer and Keita Yokoyama. On principles between Σ_1 - and Σ_2 -induction, and monotone enumerations. *J. Math. Log.*, 16(1):1650004, 21, 2016.

- [85] Quentin Le Hou  rou, Ludovic Levy Patey, and Keita Yokoyama. Conservation of Ramsey’s theorem for pairs and well-foundedness. *Trans. Amer. Math. Soc.*, 378(3):2157–2186, 2025.
- [86] Quentin Le Hou  rou, Ludovic Patey, and Keita Yokoyama. Corrigendum to: Π_4^0 conservation of Ramsey’s theorem for pairs, 2026.
- [87] Quentin Le Hou  rou and Ludovic Levy Patey. Π_4^0 conservation of the ordered variable word theorem. *The Journal of Symbolic Logic*, page 1–16, 2025.
- [88] Quentin Le Hou  rou, Ludovic Levy Patey, and Keita Yokoyama. Π_4^0 conservation of Ramsey’s theorem for pairs. *Journal of the London Mathematical Society*, 113(1):e70419, 2026.
- [89] Quentin Le Hou  rou, Ahmed Mimouni, and Ludovic Patey. The reverse mathematics of the pigeonhole hierarchy. *arXiv preprint arXiv:2407.01236*, 2026.
- [90] Quentin Le Hou  rou and Ludovic Patey. Ramsey-like theorems for separable permutations. *arXiv preprint arXiv:2507.07606*, 2025.
- [91] Quentin Le Hou  rou and Ludovic Patey. Largeness notions and polytime translation for $\forall\Sigma_3^0$ -consequences of RT_2^2 . *arXiv preprint arXiv:2602.11906*, 2026.
- [92] Quentin Le Hou  rou and Ludovic Patey. Partition theorems for Ketonen-Solovay largeness: Hardy-like version. *arXiv preprint arXiv:2602.08778*, 2026.
- [93] Quentin Le Hou  rou and Ludovic Patey. The reverse mathematics of bounded Ramsey’s theorem for pairs. *The Journal of Symbolic Logic*, page 1–26, 2026.
- [94] Manuel Lerman, Reed Solomon, and Henry Towsner. Separating principles below Ramsey’s theorem for pairs. *J. Math. Log.*, 13(2):1350007, 44, 2013.
- [95] Jiayi Liu. RT_2^2 does not imply WKL_0 . *J. Symbolic Logic*, 77(2):609–620, 2012.
- [96] Lu Liu. Cone avoiding closed sets. *Trans. Amer. Math. Soc.*, 367(3):1609–1630, 2015.
- [97] Lu Liu, Benoit Monin, and Ludovic Patey. A computable analysis of variable words theorems. *Proc. Amer. Math. Soc.*, 147(2):823–834, 2019.
- [98] Lu Liu and Ludovic Patey. The reverse mathematics of the thin set and Erd  s-Moser theorems. *J. Symb. Log.*, 87(1):313–346, 2022.
- [99] Joseph Roy Mileti. *Partition theorems and computability theory*. ProQuest LLC, Ann Arbor, MI, 2004. Thesis (Ph.D.)–University of Illinois at Urbana-Champaign.
- [100] Joseph S. Miller and Reed Solomon. Effectiveness for infinite variable words and the dual Ramsey theorem. *Arch. Math. Logic*, 43(4):543–555, 2004.
- [101] Keith R. Milliken. A Ramsey theorem for trees. *J. Combin. Theory Ser. A*, 26(3):215–237, 1979.
- [102] Ahmed Mimouni and Ludovic Patey. Ramsey-like theorems and immunities. *arXiv preprint arXiv:2508.15597*, 2026.
- [103] Benoit Monin and Ludovic Patey. SRT_2^2 does not imply RT_2^2 in ω -models. *Adv. Math.*, 389:Paper No. 107903, 32, 2021.
- [104] Antonio Montalb  n. Open questions in reverse mathematics. *Bulletin of Symbolic Logic*, 17(03):431–454, 2011. Publisher: Cambridge Univ Press.

- [105] André Nies. *Computability and randomness*, volume 51 of *Oxford Logic Guides*. Oxford University Press, Oxford, 2009.
- [106] J. B. Paris. Some conservation results for fragments of arithmetic. In *Model theory and arithmetic (Paris, 1979–1980)*, volume 890 of *Lecture Notes in Math.*, pages 251–262. Springer, Berlin, 1981.
- [107] J. B. Paris and L. A. S. Kirby. Σ_n -collection schemas in arithmetic. In *Logic Colloquium '77 (Proc. Conf., Wrocław, 1977)*, volume 96 of *Studies in Logic and the Foundations of Mathematics*, pages 199–209. North-Holland, Amsterdam-New York, 1978.
- [108] Jeff Paris and Leo Harrington. A mathematical incompleteness in Peano Arithmetic. In Jon Barwise, editor, *Handbook of mathematical logic*, pages 90–1133. North-Holland, 1977.
- [109] Charles Parsons. On a number theoretic choice schema and its relation to induction. In *Intuitionism and Proof Theory (Proc. Conf., Buffalo, N.Y., 1968)*, Stud. Logic Found. Math., pages 459–473. North-Holland, Amsterdam-London, 1970.
- [110] Ludovic Patey. Somewhere over the rainbow Ramsey theorem for pairs. *arXiv preprint arXiv:1501.07424*, 2015.
- [111] Ludovic Patey. *The reverse mathematics of Ramsey-type theorems*. PhD thesis, Université Paris Diderot, 2016. arXiv preprint arXiv:1601.04428.
- [112] Ludovic Patey. The strength of the tree theorem for pairs in reverse mathematics. *J. Symb. Log.*, 81(4):1481–1499, 2016.
- [113] Ludovic Patey. The weakness of being cohesive, thin or free in reverse mathematics. *Israel J. Math.*, 216(2):905–955, 2016.
- [114] Ludovic Patey. Partial orders and immunity in reverse mathematics. *Computability*, 7(4):323–339, 2018.
- [115] Ludovic Patey and Keita Yokoyama. The proof-theoretic strength of Ramsey’s theorem for pairs and two colors. *Adv. Math.*, 330:1034–1070, 2018.
- [116] E.L. Post. Degrees of recursive unsolvability: preliminary report (abstract). *Bull. Amer. Math. Soc.*, 54:641–642, 1948.
- [117] Pavel Pudlák. Cuts, consistency statements and interpretations. *J. Symbolic Logic*, 50(2):423–441, 1985.
- [118] Pavel Pudlák and Vojtěch Rödl. Partition theorems for systems of finite subsets of integers. *Discrete Math.*, 39(1):67–73, 1982.
- [119] F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, s2-30(1):264–286, 1930.
- [120] Z. Ratajczyk. A combinatorial analysis of functions provably recursive in $I\Sigma_n$. *Fund. Math.*, 130(3):191–213, 1988.
- [121] Saeed Salehi. On rudimentarity, primitive recursivity and representability. *Reports on Mathematical Logic*, 55:73–85, 2020.
- [122] N. W. Sauer. Coloring subgraphs of the Rado graph. *Combinatorica*, 26(2):231–253, 2006.
- [123] Dana Scott. Algebras of sets binumerable in complete extensions of arithmetic. In *Proc. Sympos. Pure Math., Vol. V*, pages 117–121. Amer. Math. Soc., Providence, RI, 1962.

- [124] David Seetapun and Theodore A. Slaman. On the strength of Ramsey's theorem. *Notre Dame J. Formal Logic*, 36(4):570–582, 1995. Special Issue: Models of arithmetic.
- [125] Saharon Shelah. Primitive recursive bounds for van der Waerden numbers. *J. Amer. Math. Soc.*, 1(3):683–697, 1988.
- [126] Wilfried Sieg. Fragments of arithmetic. *Annals of Pure and Applied Logic*, 28(1):33–71, 1985.
- [127] Stephen G. Simpson. Partial realizations of Hilbert's Program. *J. Symbolic Logic*, 53(2):349–363, 1988.
- [128] Stephen G. Simpson. *Subsystems of second order arithmetic*. Perspectives in Logic. Cambridge University Press, Cambridge; Association for Symbolic Logic, Poughkeepsie, NY, second edition, 2009.
- [129] Stephen G. Simpson and Rick L. Smith. Factorization of polynomials and Σ_1^0 induction. *Ann. Pure Appl. Logic*, 31(2-3):289–306, 1986. Special issue: second Southeast Asian logic conference (Bangkok, 1984).
- [130] Stephen G. Simpson and Keita Yokoyama. Very weak fragments of weak König's lemma, 2021.
- [131] T. Skolem. *Begründung der elementaren Arithmetik durch die rekurrierende Denkweise ohne Anwendung scheinbarer Veränderlichen mit unendlichem Ausdehnungsbereich*. Skrifter udgivne af Videnskabsselskabet i Christiania. 1.1923,6. Dybwad in Komm., 1923.
- [132] Theodore A. Slaman. Σ_n -bounding and Δ_n -induction. *Proc. Amer. Math. Soc.*, 132(8):2449–2456, 2004.
- [133] Theodore A. Slaman and Keita Yokoyama. The strength of Ramsey's theorem for pairs and arbitrarily many colors. *J. Symb. Log.*, 83(4):1610–1617, 2018.
- [134] Giovanni Solda. *Calibrating the complexity of combinatorics: reverse mathematics and Weihrauch degrees of some principles related to Ramsey's theorem*. PhD, University of Leeds, 2021.
- [135] Clifford Spector. On degrees of recursive unsolvability. *Ann. of Math. (2)*, 64:581–592, 1956.
- [136] R. Statman. Lower bounds on Herbrand's theorem. *Proc. Amer. Math. Soc.*, 75(1):104–107, 1979.
- [137] W. W. Tait. Finitism. *Journal of Philosophy*, 78(9):524–546, 1981.
- [138] Stevo Todorćević. *Introduction to Ramsey Spaces*. Princeton University Press, Princeton, 2010.
- [139] Henry Towsner and Keita Yokoyama. Erdős-Moser and $I\Sigma_2$. *Israel J. Math.*, 263(2):843–870, 2024.
- [140] Wei Wang. Some logically weak Ramseyan theorems. *Adv. Math.*, 261:1–25, 2014.
- [141] Nik Weaver. *Forcing for mathematicians*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2014.
- [142] Keita Yokoyama. On Π_1^1 conservativity of Π_2^1 theories in second order arithmetic. In C. T. Chong et al., editor, *Proceedings of the 10th Asian Logic Conference*, pages 375–386. World Scientific, 2009.