UMONS – Université de Mons – Faculté des Sciences

Tame pairs of integers

par

Quentin Lambotte

Thèse de doctorat

Science mathématique

Directrice de thèse
Françoise Point
UMONS

Composition du jury
Alexis Bès
Université Paris-Est Créteil
Quentin Menet
UMONS
Christian Michaux
UMONS
Michel Rigo
ULiège
Christophe Troestler
UMONS
Frank Wagner
Université Lyon I

Année académique 2020–2021

Remerciements

Mes remerciements s'adressent en premier lieu à ma directrice de thèse, Françoise Point. C'est grâce à elle que je me suis lancé dans la recherche en théorie des modèles. Je me souviens encore de son cours sur le théorème de Morley et les théories ω-stables, où j'ai découvert pour la première fois l'élégance et la richesse de la théorie des modèles. Le mémoire sur les corps pseudofinis que j'ai réalisé sous sa direction n'a fait que confirmer ces impressions et mon envie d'en apprendre plus sur la théorie des modèles. Depuis la première heure de cours jusqu'à la fin de la rédaction de cette thèse, j'ai pu compter sur sa disponibilité, sa patience et ses nombreux conseils. Pour ces neuf années de collaboration et de découvertes, je lui suis profondément reconnaissant.

Je remercie Alexis Bès, Quentin Menet, Christian Michaux, Michel Rigo, Christophe Troestler et Frank Wagner d'avoir accepté de consacrer une partie de leur temps à l'évaluation de mon travail.

J'ai eu la chance de partager mon bureau avec Céline Nihoul. Je tiens à la remercier pour son soutien et son écoute durant les moments difficiles, pour ses nombreux conseils et partages d'expérience d'enseignement. Je tiens aussi à remercier Stéphanie Bridoux, grâce à qui je suis un bien meilleur enseignant. Merci à Christophe Troestler pour les nombreuses conversations enrichissantes et pour l'exploit de m'avoir réconcilié avec l'analyse. Enfin, merci à Quentin Brouette pour toutes nos discussions sur le monde et nos débats sur les moyens de le rendre meilleur.

Je ne serais jamais venu au bout de cette thèse sans le soutien inconditionnel de ma famille et de mes amis. Merci à tous pour votre présence dans ma vie. Je tiens à remercier spécialement Marie-Laure pour son soutien, sa bonne humeur et pour les nombreuses pauses Covid.

Enfin, merci à toi, Thibaut, qui a cru en moi jusqu'au bout, merci pour tes encouragements et ton aide précieuse lors de la rédaction de cette thèse. C'est peu dire que je n'y serais jamais arrivé sans toi.

Table of Contents

Introduction

The group of integers, $\mathscr{Z} = (\mathbf{Z}, +, 0)$, and the ordered group of integers, $\mathscr{Z}_< = (\mathbf{Z}, +, 0, <)$, are considered well behaved structures by model theorists. The first results that support this statement are due to M. Presburger who showed in [45] that they are both decidable, by providing a quantifier elimination result in each case after adding predicates for the non-trivial subgroups of $\mathbf{Z}$. Another consequence of this quantifier elimination is that $\mathscr{Z}$ has a superstable theory of Lascar rank 1 and $\mathscr{Z}_<$ has a dependent theory with dp-rank 1 (a property also known as *dp-minimality*).

The study of decidability and definability in expansions of $\mathscr{Z}_<$ received a lot of attention, see [6] for a survey on the subject. A common line of research is the preservation of decidability and characterization of definability after adding new predicates. In this regard R. Büchi shed light in [10] on a connection between definability in $\mathscr{Z}_<$ and automata theory: a definable set in $\mathscr{Z}_<$ is k-recognizable for all $k \geq 2$. Furthermore, k-recognizability, for a fixed $k \geq 2$, is equivalent to definability in $(\mathbf{Z}, +, 0, <, V_k)$, where V_k is the function that sends $n \neq 0$ to the greatest power of k dividing n and 0 to 1. This result was first stated for $k = 2$ by R. Büchi in [10] with a predicate P_2 for the powers of 2 instead of V_2 but the proof was incorrect, as pointed out by R. McNaughton [34]: the power of expression of $(\mathbf{Z}, +, 0, <, P_2)$ is not sufficient. The characterization of k-recognizability was finally proved by V. Bruyère in [8] for $k = 2$ and a detailed proof of the general case can be found in [9]. Another important result concerning $(\mathbf{Z}, +, 0, <, V_k)$ is Cobham's Theorem [14]: given multiplicatively independent k, k', if $X \subset \mathbf{N}$ is both k-recognizable and k'-recognizable, then X is definable in $\mathscr{Z}_<$. This was generalized by A. Semenov in [48] to any set in $\mathbf{N}^n$. A proof of Cobham's and Semenov's theorems was given in [35] by C. Michaux and R. Villemaire and is based on a very useful result on definability in $\mathscr{Z}_<$: $R \subset \mathbf{Z}^n$ is not definable in $\mathscr{Z}_<$ if and only if there exists $L \subset \mathbf{Z}$ definable in $(\mathbf{Z}, +, 0, <, R)$ but not in $\mathscr{Z}_<$, see [35, Theorem 5.1].

A. Semenov investigated the preservation of decidability in expansions of the form $(\mathbf{Z}, +, 0, <, R)$ in [49], where R represents a *sparse* subset of $\mathbf{N}$, which roughly is a set enumerated by a fast growing sequence. Examples of decidable expansions of $\mathscr{Z}_<$ include $(\mathbf{Z}, +, 0, <, \{q^n \mid n \in \mathbf{N}\})$, for $q > 1$, $(\mathbf{Z}, +, 0, <, \mathrm{Fib})$, where Fib is enumerated by the Fibonacci sequence, and $(\mathbf{Z}, +, 0, <, \{n! \mid n \in \mathbf{N}\})$. More generally A. Semenov shows that for a sparse set R, $(\mathbf{Z}, +, 0, <, R)$ is decidable whenever R is effectively sparse and effectively eventually periodic modulo n for all $n > 1$, [49, Corollary 2].

The failure of decidability can happen in expansions of $\mathcal{Z}_<$ by a unary predicate. One important example is the expansion $\mathcal{Z}_{<,\mathbf{P}}$, where $\mathbf{P}$ is the set of prime numbers, which has been shown to be undecidable under Dickson's conjecture by P. Bateman, C. Jockusch and A. Woods [4]: multiplication is definable in $\mathcal{Z}_{<,\mathbf{P}}$.

The study of expansions of $\mathcal{Z}$ and $\mathcal{Z}_<$ in the context of S. Shelah's Classification Theory is surprisingly recent. A question of interest is the following: is there a classification of subsets R of $\mathbf{Z}$, such that the pair of integers $\mathcal{Z}_R = (\mathbf{Z}, +, 0, R)$ (resp. $\mathcal{Z}_{<,R} = (\mathbf{Z}, +, 0, <, R)$) is superstable (resp. dependent)? The first results on this subject are independently due to B. Poizat [44] and D. Palacín and R. Sklinos [37], where the pair $\mathcal{Z}_{\{q^n | n \in \mathbf{N}\}}$ is shown to be superstable of Lascar rank ω for any $q > 1$. Another example of a superstable pair given in [37] is $\mathcal{Z}_{\{n! | n \in \mathbf{N}\}}$. The examples of [37] were later generalized independently in two directions by G. Conant [16] and F. Point and the author [29]. These generalizations capture the fact that the examples of [37] are enumerated by fast growing sequences: in [16] it is shown that a pair $\mathcal{Z}_R$ is superstable as long as there exists a set X and a function $f : R \to X$ such that X is geometric (in the sense that $\{a/b \mid b \leq a \in X\}$ is closed and discrete) and $\sup\{|a - f(a)| \mid a \in R\} < \infty$ and in [29] it is shown that a pair $\mathcal{Z}_R$ is superstable whenever R is enumerated by a sequence (r_n) that has a *Kepler limit* θ in $\mathbf{R}^{>1} \cup \{\infty\}$ (that is θ is the limit of the successive quotients of (r_n)) and if θ is algebraic over $\mathbf{Q}$, then (r_n) is assumed to be a linear recurrence sequence whose minimal polynomial is the minimal polynomial of θ over $\mathbf{Q}$. Examples of dependent pairs of integers were given in [29], where it is shown in particular that $\mathcal{Z}_{<,\{q^n | n \in \mathbf{N}\}}$, $q > 1$, $\mathcal{Z}_{<,\{n! | n \in \mathbf{N}\}}$ and $\mathcal{Z}_{<,\mathrm{Fib}}$ are dependent.

In another direction, I. Kaplan and S. Shelah showed in [28] that the pair $\mathcal{Z}_{\mathbf{P} \cup -\mathbf{P}}$ is supersimple unstable and of Lascar rank 1, under Dickson's conjecture. This is in contrast with $\mathcal{Z}_{<,\mathbf{P}}$, which is considered wild, since it defines the multiplication by the result of P. Bateman, C. Jockusch and A. Woods.

It is interesting to note that, in the context of S. Shelah's Classification Theory, certain minimality properties of $\mathcal{Z}$ and $\mathcal{Z}_<$ are not preserved in pairs. For instance, any pair $\mathcal{Z}_R$ has Lascar rank at least ω, unless R is definable in $\mathcal{Z}$, see [37, Theorem 1]. For $\mathcal{Z}_<$, it is shown in [2, Proposition 6.6] that a pair $\mathcal{Z}_{<,R}$ is never dp-minimal, unless R is definable in $\mathcal{Z}$. Moreover, a pair $\mathcal{Z}_{<,R}$ is never strongly dependent, unless R is already definable in $\mathcal{Z}_<$, see [18, Corollary 2.20]. A similar statement holds for pairs $\mathcal{Z}_R$: the pairs $\mathcal{Z}_{\{q^n | n \in \mathbf{N}\}}$, $q > 1$, $\mathcal{Z}_{\{n! | n \in \mathbf{N}\}}$ and $\mathcal{Z}_{\mathrm{Fib}}$ are not dp-minimal [3, Corollary 5.35]. We note that the proofs of [2, Proposition 6.6] and [18, Corollary 2.20] are based on the theorem of C. Michaux and R. Villemaire [35, Theorem 5.1].

In this thesis, we investigate superstable and dependent pairs of integers.

Let $R \subset \mathbf{N}$ and let (r_n) enumerate R. We say that R is *regular* if (r_n) has a Kepler limit θ in $\mathbf{R}^{>1} \cup \{\infty\}$ and if θ is algebraic over $\mathbf{Q}$, then (r_n) is further assumed to be a linear recurrence sequence whose minimal polynomial is the minimal polynomial of θ over $\mathbf{Q}$. Our first result concerns pairs $\mathscr{Z}_R$, where R is regular.

Theorem A. *Let R be a regular set. Then $\mathscr{Z}_R$ is superstable of Lascar rank ω.*

We provide two proofs of this result. The first relies on the approach of D. Palacín and R. Sklinos in [37]. The main tool used is an abstract result of E. Casanovas and M. Ziegler [11], which in our context can be stated as follows: $\mathscr{Z}_R$ is superstable if $\mathscr{Z}_R$ is *bounded* and the induced structure R_{ind} is superstable. Here bounded means that any formula in the language of $\mathscr{Z}_R$ is equivalent to a *bounded formula*, that is a formula of the form $Q_1x_1 \in R \dots Q_nx_n \in R\, \varphi(\bar{x}, \bar{y})$, where $Q_i \in \{\exists, \forall\}$ for all $i \in \{1, \dots, n\}$ and $\varphi(\bar{x}, \bar{y})$ is a formula in the language of $\mathscr{Z}$. The induced structure R_{ind} is the trace on R of any $\{+, 0\}$-definable subset of $\mathbf{Z}^n$ without parameters.

The second proof of Theorem A is done by counting types after providing a quantifier elimination result. This quantifier elimination is done in an expanded language $\mathcal{L}$ where we add in particular a unary function S that is the successor function on R and the identity on $\mathbf{Z} \setminus R$ and predicates for formulas such as

$$\exists x_1, x_2 \in R\, (x_1 + 2x_2 = y \wedge D_2(x_1) \wedge D_5(x_2)),$$

where D_n is a predicate for the set $n\mathbf{Z}$. In the language $\mathcal{L}$, we define a theory T_R that axiomatizes $\mathscr{Z}_R$.

Theorem B. *Let R be a regular set. Then T_R has quantifier elimination, is complete and superstable.*

As a corollary of the proof of Theorem A, we obtain preservation of ω-stability of expansions of $(\mathbf{Q}, +, 0)$ and $(\mathbf{R}, +, 0)$ by a regular set.

Theorem C. *Let R be a regular set. Then $(\mathbf{Q}, +, 0, R)$ and $(\mathbf{R}, +, 0, R)$ are ω-stable.*

We also point out an analogy between Theorem A and expansions of fields by a subgroup with the Mann property, as studied by L. van den Dries and A. Günaydın in [19]. This allows us to give a quick proof of a special case of a result of G. Conant on expansions of $\mathscr{Z}$ by a multiplicative submonoid, see [15, Theorem 3.1].

Theorem D. *Let $(M, \cdot, 1)$ be a submonoid of $(\mathbf{Z}, \cdot, 1)$ such that the subgroup of $(\mathbf{Q} \setminus \{0\}, \cdot, 1)$ generated by M has the Mann property. Then $\mathscr{Z}_M$ is superstable.*

We then investigate pairs of the form $\mathscr{Z}_{<,R}$, where R is a sparse set in the sense of A. Semenov, particular instances of such being regular sets. We revisit the work of A. Semenov and give a quantifier elimination result for pairs $\mathscr{Z}_{<,R}$, R sparse, in a language $\mathcal{L}_<$ similar to the one used in Theorem B. In this language, we define a theory $T_{<,R}$ which axiomatizes $\mathscr{Z}_{<,R}$.

Theorem E. *Let R be a sparse set. Then $T_{<,R}$ has quantifier elimination and is complete.*

This theorem allows to reduce the dependency of $\mathscr{Z}_{<,R}$ to the dependency of atomic formulas in the language $\mathcal{L}_<$. Using technology developed by A. Chernikov and P. Simon in [12], namely *honest definition over a predicate*, we are able to deduce the dependency of $\mathscr{Z}_{<,R}$ from the dependency of the theory $\mathrm{Th}(R, S, c, D_{n,k} \mid k < n \in \mathbf{N})$, where c is the smallest element in R, S is the successor function on R and $D_{n,k}$ is the set of elements in R that are equal to k modulo n.

Theorem F. *Let R be a sparse set. Then $\mathrm{Th}(\mathscr{Z}_{<,R})$ is dependent.*

Let us add that A. Chernikov and P. Simon provide in [12] an analogue of E. Casanovas and M. Ziegler's theorem [11] for dependent theories: $\mathscr{Z}_{<,R}$ is dependent if R_{ind} is dependent and $\mathscr{Z}_{<,R}$ is bounded, see [12, Corollary 2.6]. Theorem E implies that the pair $\mathscr{Z}_{<,R}$ is bounded when R is sparse. Furthermore, Theorem E implies that R_{ind} is essentially $\mathrm{Th}(R, S, c, D_{n,k} \mid k < n \in \mathbf{N})$. We however do not establish this in details and prove Theorem F by hand without the appeal of [12, Corollary 2.6].

Theorem F gives a proof of a result announced in [2, 3] that the pairs $\mathscr{Z}_{<,\{q^n \mid n \in \mathbf{N}\}}$, $q > 1$, $\mathscr{Z}_{<,\{n! \mid n \in \mathbf{N}\}}$ and $\mathscr{Z}_{<,\mathrm{Fib}}$ are dependent.

As in the case of expansions of $\mathscr{Z}$ by a regular set, we are able to extract from the proof of Theorem F the following results concerning expansions of $(\mathbf{Q}, +, 0, <)$ and $(\mathbf{R}, +, 0, <)$ by a sparse set.

Theorem G. *Let R be a sparse set. Then*

1. *if $(\mathbf{Q}, +, 0, <, R)$ is bounded, then $(\mathbf{Q}, +, 0, <, R)$ is dependent;*
2. *if $(\mathbf{R}, +, 0, <, R)$ is bounded, then $(\mathbf{R}, +, 0, <, R)$ is dependent;*
3. *if $(\mathbf{R}, +, 0, \lfloor \cdot \rfloor, <, R)$ is bounded, then $(\mathbf{R}, +, 0, \lfloor \cdot \rfloor, <, R)$ is dependent, where $\lfloor \cdot \rfloor$ is the integer part function.*

This thesis is organized in five chapters, each beginning with an introduction. The first chapter collects preliminary material on stability theory, dependent theories and linear recurrence sequence. In a first section, we recall what is the Lascar rank and how it can be calculated using the notion of generic type in a stable group. In the next section, we define what is a dependent theory and we give various equivalent definitions as well as various examples relevant to this thesis. We end with a section in which we collect several facts on linear recurrence sequences, such as a description using exponential polynomials and a necessary and sufficient condition for the existence of a Kepler limit.

The second chapter details the context in which this thesis takes place. In the first section, we recall the main technical tools used in the study of pairs both in the stable and dependent case, namely the results of E. Casanovas and M. Ziegler and A. Chernikov and P. Simon. In the second section, we provide a quick review of recent results on tame pairs of integers, with an emphasis on pairs $\mathscr{Z}_R$, where R is enumerated by a linear recurrence sequence, where interesting phenomena occur.

The third chapter gives the first proof of Theorem A, as well as the proofs of Theorem C and Theorem D. The first five sections of Chapter 3 contain the proof of Theorem A, see Theorem 3.5.1. We first show that R_{ind} is superstable for any regular set R. To do this, we only need to check that the induced structure on R by equations in $\mathscr{Z}_R$ is superstable. Assuming (r_n) enumerates R, we analyze sets of the form $X_{\bar{a}} = \{(r_{n_1}, \ldots, r_{n_k}) \in R^k \mid a_1 r_{n_1} + \cdots + a_k r_{n_k} = 0\}$, where $k \geq 1$ and $\bar{a} \in \mathbf{Z}$. We show that for all $\bar{a} \in \mathbf{Z}^k$ there is $c_{\bar{a}} \in \mathbf{N}$ such that if $(r_{n_1}, \ldots, r_{n_k}) \in X_{\bar{a}}$, then

$$\max\{|n_i - n_j| \mid 1 \leq i, j \leq k\} \leq c_{\bar{a}},$$

unless there exists $I \subsetneq \{1, \ldots, k\}$ such that $\sum_{i \in I} a_i r_{n_i} = 0$. This is the content of Proposition 3.3.1 and its proof relies on the following property: sets of the form $\{r_n \in R \mid a'_0 r_n + a'_1 r_{n+1} + \cdots + a'_\ell r_{n+\ell} = 0\}$ are either finite or R, where $\bar{a}' \in \mathbf{Z}^{\ell+1}$ and $\ell \in \mathbf{N}$. The analysis of the sets $X_{\bar{a}}$ allows us to show that the induced structure on R by equations is definably interpreted in the superstable structure $(\mathbf{N}, S, S^{-1}, 0)$, where $S(n) = n + 1$, $S^{-1}(n + 1) = n$ and $S^{-1}(0) = 0$. This will be enough to conclude that R_{ind} is superstable, using an observation made by G. Conant that $(\mathbf{N}, S, S^{-1}, 0)$ is monadically superstable: the expansion of $(\mathbf{N}, S, S^{-1}, 0)$ by unary predicates for *any* subset of $\mathbf{N}$ is superstable.

We then show that $\mathscr{Z}_R$ is bounded when R is regular. Recall that a subset of $\mathbf{N}$ is *piecewise syndetic* if it contains arbitrarily long sequences with bounded gaps. We use again Proposition 3.3.1 to show that we cannot bound the length of expansions in base R of natural numbers. In other words, we show that any set of the form

$$\{z \in \mathbf{Z} \mid z = a_1 r_{n_1} + \cdots + a_k r_{n_k} \text{ for some } (r_{n_1}, \ldots, r_{n_k}) \in R^k\} \cap \mathbf{N}$$

is not piecewise syndetic. This allows us to prove that R is bounded.

Chapter 3 ends with three sections: one where we take time to compare Theorem A with the main result of G. Conant in [16] and the last two sections are respectively on Theorems D and C, see respectively Theorems 3.7.2 and 3.8.1.

Theorem B is proved in the fourth chapter and we use it to give another proof of Theorem A. The proof of Theorem B is based on the content of Chapter 3 and is done in three steps, see Theorem 4.1.1, Corollary 4.1.3 and Theorem 4.5.1. The first step consist in an in-depth analysis of equations in non-standard models of T_R. The second step consist in a construction of algebraically prime models in T_R. The third and final step consist in showing that T_R is 1-existentially closed, that is existentially closed for existential formulas with at most one existential quantifier. This will be enough to conclude that T_R has quantifier elimination. We then show that T_R is superstable by counting types. We end Chapter 4 with a brief account of the decidability of T_R.

The final chapter is concerned with pairs $\mathscr{Z}_{<,R}$, with R sparse, where Theorems E, F and G are proved. The main step towards the proof of Theorem E (see Theorem 5.5.3 and Corollary 5.5.4) consists in showing that the negation of certain bounded existential formulas is equivalent to a bounded existential formula. For instance, the negation of the formula

$$\exists \bar{x} \in R \left(\bigwedge_{i=1}^{n} \sum_{j=1}^{m} a_{ij} x_j > y_i \wedge \varphi(\bar{x}) \right), \tag{1}$$

where $\bar{a}_1, \ldots, a_n \in \mathbf{Z}^m$ and $\varphi(\bar{x})$ is any formula in $\{S, S^{-1}, c, <\}$ with quantifiers relativized to R, is equivalent to a disjunction of existential formulas of the same shape. Using Theorem E, we show that the dependency of T_R reduces to the dependency of existential formulas such as the formula (1). Then using the fact that those formulas have honest definitions over R by formulas in the language $\{S, S^{-1}, c, <\}$ with quantifiers relativized to R, we obtain Theorem F, see Theorem 5.6.7. We end Chapter 5 with a proof of Theorem G, see Theorem 5.7.1.

Part of the content of Chapters 3 and 4 as well as a special case of Theorem F, under the assumption that R is ultimately periodic modulo n for all $n > 1$, have been published in the following paper:

> Quentin Lambotte and Françoise Point. "On expansions of $(\mathbf{Z}, +, 0)$". In: *Ann. Pure Appl. Logic* 171.8 (2020), pp. 1–36.

CHAPTER

1

Preliminaries

We present in this chapter the material we will need on stable and dependent theories and on linear recurrence sequence. We assume basic knowledge in model theory, specifically the content of [53, Chapters 1-5] or [33, Chapters 1-5].

This chapter is organized as follows. In Section 1.1, we recall several notions from stability theory. Specifically, the notions of λ-stable and superstable theories are introduced for arbitrary theories and Lascar rank and generic types are introduced for theories whose models are superstable abelian groups, possibly with extra structure (by extra structure we mean, for instance, that of a field). The focus on the superstable case, while not the most general, allows to present the notion of Lascar rank without the appeal to the abstract notions of dividing and forking for formulas and types.

In Section 1.2, we recall the definition of a dependent theory and several tools and criteria that are helpful to show whether or not a complete theory is dependent. We also take some time to illustrate this notion with several examples that are relevant for Chapter 5.

In the last section, we give a quick account of elementary properties of linear recurrence sequences, as these well behaved sequences give instances of tame expansions of both $(\mathbf{Z}, +, 0)$ and $(\mathbf{Z}, +, 0, <)$. Among other things, we recall that these sequences can be represented as sums of exponentials with polynomial coefficients and that they satisfy a strong regularity property: the set of indices where those sequences take the value 0 is a union of finite number of sets of the form $a + b\mathbf{N}$, $a, b \in \mathbf{N}$. This last property is known as the Skolem-Mahler-Lech Theorem.

We end this introduction with the notations and conventions that we use throughout this text.

The set of natural numbers, of integers and of real numbers will be denoted respectively $\mathbf{N}$, $\mathbf{Z}$ and $\mathbf{R}$. When X is one of the above sets and $a \in X$, the notations $X^{>a}$, $X^{\geq a}$ and $X^{>a}_{\infty}$ refer respectively to the sets $\{x \in X \mid x > a\}$, $\{x \in X \mid x \geq a\}$ and $X^{>a} \cup \{\infty\}$. For a natural number n, the set $\{1, \ldots, n\}$ will be denoted $[n]$. The cardinality of a set A will be denoted by $|A|$. Likewise, the length of a tuple $\bar{x}$ will be denoted $|\bar{x}|$. For $x \in \mathbf{R}$, $\lfloor x \rfloor$ is the integer part of x, that is $\max\{n \mid n \leq x, n \in \mathbf{Z}\}$. Given a sequence $(r_n)_{n \in I}$, $I \subset \mathbf{N}$, of complex numbers, we usually write (r_n) in case

$I = \mathbf{N}$ instead of $(r_n)_{n\in\mathbf{N}}$.

Capital letters I, J and K will refer to (usually non-empty) sets of indices. The power set of X will be denoted $\mathfrak{P}(X)$. Capital letters will refer to sets and small letters will refer to elements of a given set. For a tuple $\bar{a}$ of length n and $I \subset [n]$, $\bar{a}_I$ refers to the tuple $(a_i \mid i \in I)$. For $n \in \mathbf{N}^{>0}$, we let $\mathsf{Part}([n])$ be the set of (ordered) partitions $\bar{I} = (I_1, \ldots, I_\ell)$ of $[n]$.

A first-order language will be denoted by the letter $\mathcal{L}$, possibly with a subscript. An $\mathcal{L}$-structure will be referred to by a round letter and its domain by the corresponding capital letter. For instance $\mathscr{M}$ is an $\mathcal{L}$-structure whose domain is M. For an element a of M and $A \subset M$, the notations $\mathrm{acl}^{\mathcal{L}}(a/A)$, $\mathrm{tp}^{\mathcal{L}}(a/A)$ mean respectively the algebraic closure and the type of a over A in $\mathscr{M}$. The set of n-types over A is denoted $S_n^{\mathcal{L}}(A)$, or $S_n(A)$ if the ambient language is clear. Likewise, $S^{\mathcal{L}}(A)$ denotes

$$\bigcup_{n\in\mathbf{N}} S_n^{\mathcal{L}}(A)$$

and we use also the notation $S(A)$ if the ambient language is clear. If $R \in \mathcal{L}$ is a n-ary predicate symbol, the set $\{\bar{a} \in M^n \mid \mathscr{M} \models R(\bar{a})\}$ will be denoted $R(M^n)$ or simply R when there is no confusion.

We make the following (usual) abuse of notations. When R is a unary predicate symbol, expressions of the form $\exists x \in R\, \varphi(x)$ and $\forall x \in R\, \varphi(x)$ respectively mean $\exists x\, (R(x) \wedge \varphi(x))$ and $\forall x\, (R(x) \to \varphi(x))$.

For each $n \in \mathbf{N}^{>1}$, let D_n be a unary predicate. We let $\mathcal{L}_g = \{+, -, 0, D_n \mid n > 1\}$ and $\mathcal{L}_S = \{S, S^{-1}, c\}$, where S and S^{-1} are unary function symbols and c is a constant symbol. An abelian group $(G, +, -, 0)$ will always be expanded to an $\mathcal{L}_g$-structure as follows: for each $n \in \mathbf{N}^{>1}$, the symbol D_n is interpreted as the set $\{x \in G \mid (G, +, -, 0) \models \exists y\, x = ny\}$. When we work in $\mathbf{Z}$, we sometimes use the notation $x \equiv_n y$ instead of $D_n(x - y)$. The group $(\mathbf{Z}, +, -, 0)$, considered as an $\mathcal{L}_g$-structure, will be denoted by $\mathscr{Z}$. When working in $\mathrm{Th}(\mathscr{Z})$, an expression of the form $x > c$, where $c \in \mathbf{N}$, is an abbreviation for $\bigwedge_{i=0}^{c} x \neq i$.

Given a ring K, $K[X]$ is the ring of polynomials in the variable X with coefficients in K. In what follows, given $P \in \mathbf{C}[X]$, when we say that P is irreducible, we mean irreducible over $\mathbf{Q}$. And for $z \in \mathbf{C}$, when we say that z is algebraic, we mean algebraic over $\mathbf{Q}$.

1.1 Lascar rank and generics in superstable abelian groups

We recall in this section the necessary stability theoretic material needed for the proof of the main results of Chapter 3. We first recall the definition of a superstable theory.

Definition 1.1.1. Let T be a complete $\mathcal{L}$-theory, $\mathcal{L}$ a countable language, and λ a cardinal.

1. We say that T is *λ-stable* if for all $\mathscr{M} \models T$ and all $A \subset M$, if $|A| = \lambda$, then $|S_1(A)| = \lambda$.
2. We say that T is *superstable* if it is λ-stable for all $\lambda \geq 2^{\aleph_0}$.

For an $\mathcal{L}$-structure $\mathscr{M}$, we say that $\mathscr{M}$ is λ-stable (resp. superstable) if $\mathrm{Th}(\mathscr{M})$ is λ-stable (resp. superstable).

The central example of superstable theory in this thesis is $T = \mathrm{Th}(\mathscr{Z})$ (see [46, Theorem 15.4.4]) and this can be seen using quantifier elimination in T (see [46, Theorem 15.2.1]).

In these preliminaries, we always work in a language $\mathcal{L}$ containing $\{+, -, 0\}$ and a complete $\mathcal{L}$-theory whose models are infinite abelian groups with respect to $\{+, -, 0\}$. We further assume that T is superstable and we fix $\mathscr{G}$ a monster model of T. As we mentioned earlier, these are extra conditions with regard to the general theory, a treatment of which can be found in the following references: [43] for stable groups and [39], [53, Chapter 8] or [42] for general stability theory.

Let us start by recalling the definition of Lascar rank.

Definition 1.1.2 ([42, P. 438])**.** Let $A \subset G$. We define the rank U of a type in $S(A)$ as follows:

1. $U(p) \geq 0$ for all $p \in S(A)$;
2. if α is a limit ordinal, $U(p) \geq \alpha$ if and only if $U(p) \geq \beta$ for all $\beta \leq \alpha$;
3. $U(p) \geq \alpha + 1$ if and only if for all cardinal λ, there exists $B \subset G$, such that $A \subset B$ and p has at least λ extensions q in $S(B)$ with $U(q) \geq \alpha$.

We set $U(p) = \infty$ if $U(p) \geq \alpha$ for all ordinal α and otherwise $U(p)$ is $\sup\{\alpha \mid U(p) \geq \alpha\}$. If $p = \mathrm{tp}\,(\bar{a}/A)$, we write $U(\bar{a}/A)$ instead of $U(\mathrm{tp}\,(\bar{a}/A))$.

A type $p \in S_1(A)$ has U-rank 0 if and only if it is algebraic (that is, has finitely many realizations). From the definition, we have that if $q \in S_1(B)$ extends $p \in S_1(A)$, then $U(q) \leq U(p)$ (p has at least the same number of extensions of q to any set that contains B). An equivalent definition of a superstable theory is that the rank of any type is an ordinal.

We will need the following property of the U-rank, which is known as Lascar's inequality. In the following statement, given ordinals α and β, $\alpha + \beta$ is the ordinary sum while $\alpha \oplus \beta$ is defined as follows: if $\alpha = \omega^{\gamma_1} n_1 + \cdots + \omega^{\gamma_k} n_k$ and $\beta = \omega^{\gamma_1} n'_1 + \cdots + \omega^{\gamma_k} n'_k$, where $\bar{n}, \bar{n}' \in \mathbf{N}^k$ and $0 \leq \gamma_{i+1} < \gamma_i$, $i \in [k-1]$, then

$$\alpha \oplus \beta = \omega^{\gamma_1}(n_1 + n'_1) + \cdots + \omega^{\gamma_k}(n_k + n'_k).$$

Proposition 1.1.3 ([42, Théorème 19.04]). *Let $A \subset G$. Let $\bar{a}, \bar{b}$ be finite tuples of elements in G. Then*

$$U(\bar{b}/A, \bar{a}) + U(\bar{a}/A) \leq U(\bar{a}, \bar{b}/A) \leq U(\bar{b}/A, \bar{a}) \oplus U(\bar{a}/A).$$

In particular, if $U(\bar{a}/A) < \omega$ and $U(\bar{b}/A) < \omega$, then $U(\bar{a}, \bar{b}/A) = U(\bar{a}/\bar{b}A) + U(\bar{b}/A)$.

Let us analyze the notion of Lascar rank in the theory of $\mathscr{Z}$.

Example 1.1.4. In $T = \mathrm{Th}(\mathscr{Z})$, a 1-type has either rank 0 or 1.

Proof. To this end, we recall that T has quantifier elimination. As a result, two types q, q' can be distinguished by atomic and negation of atomic formulas.

Let $\mathscr{G}$ be a monster model of T and $A \subset G$. Let $p = \mathrm{tp}\,(a/A)$ be a non-algebraic type. Thus $U(p) \geq 1$. Let us show that we cannot have $U(p) \geq 2$. We will show that for all $B \supset A$, p has at most $2^{\aleph_0}$ non-algebraic extensions in $S_1(B)$. In fact we will show that there are $2^{\aleph_0}$ non-algebraic types over B.

Recall that any term $t(x, \bar{y})$ is of the form $nx + t'(\bar{y})$, where $n \in \mathbf{Z}$ and $t'(\bar{y})$ is a term. Let $B \supset A$. Let $q \in S_1(B)$ be non-algebraic. Let $n \in \mathbf{Z}$ and $t(\bar{y})$ be a term. The formula $nx + t(\bar{b}) = 0$, $\bar{b} \in B^{|\bar{y}|}$, has at most one solution in $\mathscr{G}$, unless $n = 0$ and $t(\bar{b}) = 0$. So for all term $t(\bar{y})$ and $n \in \mathbf{Z}$, if $n \neq 0$, $nx + t(\bar{b}) \neq 0 \in q$ for all $\bar{b} \in B^{|\bar{y}|}$. Also, the formula $t(\bar{b}) = 0 \in q$ whenever $\mathscr{G} \models t(\bar{b}) = 0$, where $t(\bar{y})$ is a term and $\bar{b} \in B^{|\bar{y}|}$. As a result any two non-algebraic $q, q' \in S_1(B)$ extensions of p have the same equations and negation of equations with parameters in B. So the atomic formulas that may distinguish q from p are of the form $D_k(nx + t(\bar{b}))$.

Let $q \in S_1(B)$ be non-algebraic. Let $k \in \mathbf{N}^{>1}$ and consider the formula in q of the form $D_k(nx + t(\bar{b}))$, where $n \in \mathbf{Z}$, $t(\bar{y})$ is a term and $\bar{b} \in B^{|\bar{y}|}$. We may assume that $n \neq 0$. The set $D_k(nG + t(\bar{b}))$ is a coset of the subgroup $D_k(nG)$. Thus, for all $k \in \mathbf{N}^{>1}$ and $n \in \mathbf{Z} \setminus \{0\}$, the set of formulas

$$\{D_k(nx + t(\bar{b})) \mid D_k(nx + t(\bar{b})) \in q, t(\bar{y}) \text{ is a term and } \bar{b} \in B^{|\bar{y}|}\}$$

is implied by one of its formulas, since the intersection of two cosets of a subgroup is either empty or equal to the cosets. Thus the set

$$\{D_k(nx + t(\bar{b})) \mid D_k(nx + t(\bar{b})) \in q, n \in \mathbf{Z} \setminus \{0\}, t(\bar{y}) \text{ is a term and } \bar{b} \in B^{|\bar{y}|}\}$$

is determined by a countable set of formulas. Hence there are only $2^{\aleph_0}$ non-algebraic types over B. Finally, we get that $U(p) = 1$ if p is non-algebraic. □

The U-rank allows to divide the extensions of a given type in two categories.

Definition 1.1.5. Let $A, B \subset G$. Let $p \in S_1(A)$ and $q \in S_1(B)$ an extension of p. We say that q is a *forking extension* (resp. a *non-forking extension*) of p if $U(q) < U(p)$ (resp. $U(q) = U(p)$).

This definition says that a forking extension of a type contains much more information that the type itself.

In $\mathrm{Th}(\mathscr{Z})$, if $p \in S(A)$ and $U(p) = 1$, then the forking extensions of p are exactly the algebraic ones.

We define the U-rank of T as follows.

Definition 1.1.6. The U-rank of T, noted $U(T)$, is defined as $\sup\{U(p) \mid p \in S_1(\emptyset)\}$.

In view of Example 1.1.4, we have $U(\mathrm{Th}(\mathscr{Z})) = 1$.

In our context, one can show that the U-rank of T is the U-rank of certain types whose set of realizations in $\mathscr{G}$ is large. More precisely, the types in question contain only formulas that define sets X that can cover G with only finitely many translates (if X is a subgroup, this means that X has finite index in G).

Definition 1.1.7 ([43, Section 5.a])**.** Let $X \subset G$ be a definable set, possibly with parameters. Then X is said to be *generic* if there are $k \in \mathbf{N}$ and $\bar{g} \in G^k$ such that $G = (g_1 + X) \cup \cdots \cup (g_k + X)$. A formula is called *generic* if it defines a generic set. Likewise, a type over A is called *generic* if it contains only generic formulas with parameters in A.

Observe that the notions of generic set, formulas and types over A are invariant under automorphisms. Generic sets enjoy also the following properties:

1. if X is generic and $Y \supset X$ is definable, then Y is generic;
2. either X is generic or $G \setminus X$ is generic (here, the stability of T is needed, see [43, Lemme 5.1]);
3. if $X \cup Y$ is generic, then either X or Y is generic. This is a consequence of the previous item. Indeed, assume that there are $g_1, \ldots, g_k \in G$ such that
$$G = \bigcup_{i \in [k]} g_i + (X \cup Y).$$
Then we have
$$G = \bigcup_{i \in [k]} g_i + X \cup \bigcup_{i \in [k]} g_i + Y.$$
Therefore, by item 2 either $X' = \bigcup_{i \in [k]} g_i + X$ is generic or $G \setminus X'$ is generic. As $G \setminus X' \subset \bigcup_{i \in [k]} g_i + Y$, we have that X' or $Y' = \bigcup_{i \in [k]} g_i + Y$ is generic. Let us

show that if X' is generic then X is generic, the case of Y' being identical. If X' is generic, there are $h_1, \dots, h_n \in G$ such that

$$G = \bigcup_{i \in [n]} h_i + X'.$$

But by definition of X' we get

$$G = \bigcup_{i \in [n]} \bigcup_{j \in k} (h_i + g_j) + X,$$

which shows that X is generic.

However, the set of generics is not stable under intersection. For instance, in $\mathscr{G}$, $3G$ is a definable generic set since it is a subgroup of finite index and so is any coset of $3G$, who are disjoint from $3G$. On the other hand, one can show that the set of complements of non-generic sets is a filter. This yields the following existence result for generic types.

Proposition 1.1.8 ([43, Corollaire 5.2]). *Let $A \subset G$. Then there is a generic type over A.*

We again illustrate the notion of generic type in the theory of $\mathscr{Z}$.

Example 1.1.9. The generic types in $T = \mathrm{Th}(\mathscr{Z})$ are those of U-rank 1.

Proof. Indeed let $p \in S_1(A)$. If p is generic, we must have $U(p) = 1$, since an algebraic type must contain an algebraic formula. On the other hand, in $U(p) = 1$, then all formulas in p are non-algebraic. So we only need to check non-algebraic formulas are generic. By quantifier elimination in T, a definable set in G is a boolean combination of finite sets and cosets of subgroups of the form nG. Thus an infinite definable set X in G is a finite union of cosets of subgroups of finite index minus a finite set. Hence X is generic. □

The only extensions of a generic type are the non-forking ones.

Proposition 1.1.10 ([43, Lemme 5.5]). *Let $A, B \subset G$. Let $p \in S(A)$ and $q \in S(B)$ such that q is an extension of p. Assume that p is generic. Then q is generic if and only if it is a non-forking extension of p.*

For a given set of parameters, we cannot in general say that there exists a unique generic type. However, one can identify a subgroup of $\mathscr{G}$, called the generic component of $\mathscr{G}$, in which there is only one generic type over models.

Definition 1.1.11 ([43, Section 1.d]). Let $\varphi(x, \bar{y})$ be a $\mathcal{L}$-formula. The *connected component* of $\mathscr{G}$, denoted $\mathscr{G}^0$, is the intersection of all finite index subgroups of G definable by an $\mathcal{L}$-formula.

Remark 1.1.12. We could have defined the connected component allowing parameters in any given set A, giving a *connected component over* A denoted $\mathscr{G}_A^0$. However, as explained in [43, Section 1.d], since we work in a stable theory, we have $\mathscr{G}_A^0 = \mathscr{G}^0$. Thus the notion of a connected component over A is not necessary.

Example 1.1.13. In $T = \mathrm{Th}(\mathscr{Z})$, $\mathscr{G}^0$ is the subgroup $\bigcap_{n \in \mathbf{N}} nG$.

Proof. This again follows from quantifier elimination in T, which implies that the only groups definable in $\mathscr{G}$ are of the form nG, which have finite index in $\mathscr{G}$. Hence $\mathscr{G}^0 = \bigcap_{n \in \mathbf{N}} nG$. □

Proposition 1.1.14 ([43, Proposition 5.9 and Section 5.c]). *Let $\mathscr{H} \prec \mathscr{G}$. There exists a unique generic type p over H such that $p(H) \subset \mathscr{G}^0$ and it is called* the principal generic type *over $\mathscr{H}$. Furthermore, $U(T) = U(p)$.*

For $T = \mathrm{Th}(\mathscr{Z})$, the principal generic over $\mathbf{Z}$ is the type determined by $\{D_n(x) | n \in \mathbf{N}^{>1}\}$.

By stationarity of types over models, we get that the generic type over $\mathscr{H}$ has a unique generic extension to any parameter set.

Proposition 1.1.15. *Let $\mathscr{H} \prec \mathscr{G}$. Let $H \subset A \subset G$. Then the principal generic type over H has a unique generic extension in $S_1(A)$.*

1.2 Dependent theories

In this section, we present the material needed in Chapter 5 on dependent theories. We fix T a complete $\mathcal{L}$-theory with infinite models and we let $\mathscr{M}$ be a monster model of T. The main reference on dependent theories we used is [50].

Definition 1.2.1 ([50, Definition 2.1]).

1. Let $\varphi(\bar{x}, \bar{y})$ be an $\mathcal{L}$-formula. We say that $\varphi(\bar{x}, \bar{y})$ is *dependent* if there do not exist $A \subset M^{|\bar{x}|}$ infinite countable and $(\bar{b}_I \mid I \subset A)$ a sequence in $M^{|\bar{y}|}$ such that

$$\mathscr{M} \models \varphi(\bar{a}, \bar{b}_I) \text{ if and only if } a \in I, \text{ for all } \bar{a} \in A.$$

2. T is said *dependent* if all $\mathcal{L}$-formulas $\varphi(\bar{x}, \bar{y})$ are dependent.

A standard fact about dependent theories is that it is enough to check the dependency of formulas $\varphi(\bar{x}, \bar{y})$ with $|\bar{x}| = 1$ (see [50, Proposition 2.11]). Even more, we have the following result when T has quantifier elimination.

Lemma 1.2.2 ([50, Lemma 2.9]). *Assume that T has quantifier elimination. Then T is dependent if and only if all atomic formulas $\varphi(x, \bar{y})$ are dependent.*

The following characterization of independence will be useful in Chapter 5.

Proposition 1.2.3 ([50, Lemma 2.7]). *Let $\varphi(\bar{x},\bar{y})$ be an $\mathcal{L}$-formula. Then $\varphi(\bar{x},\bar{y})$ is independent if and only if there exist an indiscernible sequence $(\bar{a}_i \mid i \in \omega)$ in $M^{|\bar{x}|}$ and $\bar{b} \in M^{|\bar{y}|}$ such that*

$$\mathscr{M} \models \varphi(\bar{a}_i, \bar{b}) \text{ if and only if } i \text{ is even.}$$

As a result, it is possible to associate to each dependent formula $\varphi(\bar{x},\bar{y})$ a minimal natural number, called *alternation number*, $n = \mathrm{alt}(\varphi)$ such that for all indiscernible sequences $(\bar{a}_i \mid i \in \omega)$ in $M^{|\bar{x}|}$ and $\bar{b} \in M^{|\bar{y}|}$, we cannot find $i_1 < \cdots < i_n \in \omega$ with $\neg(\varphi(\bar{a}_{i_j}, \bar{b}) \leftrightarrow \varphi(\bar{a}_{i_{j+1}}, \bar{b}))$ for all $j \in [n]$. (See [50, Page 9] for more details.)

A similar characterization of dependence in terms of indiscernible sequence holds.

Proposition 1.2.4 ([50, Proposition 2.8]). *Let λ be ω or ω_1. Let $\varphi(\bar{x},\bar{y})$ be an $\mathcal{L}$-formula. Then $\varphi(\bar{x},\bar{y})$ is dependent if and only if for all indiscernible sequence $(\bar{a}_i \mid i \in \lambda)$ in $M^{|\bar{x}|}$ and all $\bar{b} \in M^{|\bar{y}|}$, the truth value of $\varphi(\bar{a}_i, \bar{b})$ is eventually constant.*

Remark 1.2.5. The statement of [50, Proposition 2.8] actually requires that Proposition 1.2.4 holds for arbitrary linear orders instead of just $(\lambda, <)$. But as stated after the proof of [50, Proposition 2.8], it is enough to restrict ourselves to a fixed $(\lambda, <)$, by compactness and Ramsey's Theorem.

As we almost always use this last characterization of dependency, we shall omit to explicitly appeal to Proposition 1.2.4 in the rest of this text.

Let us give examples of dependent theories.

Example 1.2.6. Any stable theory is dependent.

Proof. One way of showing this is as follows. Recall that in a stable theory an indiscernible sequence $(a_i \mid i \in \omega)$ over A is totally indiscernible: for all $\bar{i}, \bar{j} \in \omega^n$, $\mathrm{tp}\left(a_{i_1}, \ldots, a_{i_n} / A\right) = \mathrm{tp}\left(a_{j_1}, \ldots, a_{j_n} / A\right)$ (see [53, Lemma 9.1.1]). We will show that given $\varphi(x,\bar{y})$ and $(a_i \mid i \in \omega)$ totally indiscernible, there exists $n \in \mathbf{N}$ such that for all $\bar{b} \in M^{|\bar{y}|}$, either $|\{i \in \omega \mid \mathscr{M} \models \varphi(a_i, \bar{b})\}| \leq n$ or $|\{i \in \omega \mid \mathscr{M} \models \neg\varphi(a_i, \bar{b})\}| \leq n$. This of course implies that $\varphi(x,\bar{y})$ is dependent, and so does T by [50, Proposition 2.11].

Assume towards a contradiction that for all $n \in \mathbf{N}$, there exists $\bar{b}_n \in M^{|\bar{y}|}$ such that $|\{i \in \omega \mid \mathscr{M} \models \varphi(a_i, \bar{b}_n)\}| > n$ and $|\{i \in \omega \mid \mathscr{M} \models \neg\varphi(a_i, \bar{b}_n)\}| > n$. Let $\kappa \geq \aleph_0$. We want to show that there exists $A \subset M$ such that $|S_{|\bar{y}|}(A)| = 2^\kappa$, contradicting stability. By [53, Lemma 5.1.3], there exists $(a'_i \mid i \in \kappa)$ indiscernible such that for all $\bar{i} \in \kappa^n$, $\mathrm{tp}\left(a'_{i_1}, \ldots, a'_{i_n}\right) = \mathrm{tp}\left(a_1, \ldots, a_n\right)$. Let $A = \{a'_i \mid i \in \kappa\}$ and let us show that $|S_{|\bar{y}|}(A)| = 2^\kappa$. Given disjoint subsets I_1, I_2 of κ, the formula

$$\bigwedge_{i \in I_1} \varphi(a'_i, \bar{y}) \wedge \bigwedge_{i \in I_2} \neg\varphi(a'_i, \bar{y})$$

is consistent, because of total indiscernibility of $(a'_i \mid i \in \kappa)$. Hence $|S_{|\bar{y}|}(A)| = 2^\kappa$. □

Example 1.2.7. Presburger arithmetic, that is $\mathrm{Th}(\mathbf{Z}, +, 0, <)$, is dependent.

Proof. This is done using quantifier elimination of $\mathrm{Th}(\mathbf{Z}, +, 0, <)$ in the language $\{+, -, 0, <, D_n \mid n \in \omega\}$ (see [33, Corollary 3.1.21]) and Lemma 1.2.2. By quantifier elimination, an atomic formula $\varphi(x, \bar{y})$ is one of the following formulas:

1. $nx + t(\bar{y}) > 0$ for some $n \in \mathbf{Z}$ and $t(\bar{y})$ a term;
2. $nx + t(\bar{y}) = 0$ for some $n \in \mathbf{Z}$ and $t(\bar{y})$ a term;
3. $D_n(kx)$, for some $n \in \mathbf{N}^{>1}$ and $k \in \mathbf{Z}$;
4. $D_n(t(\bar{y}))$, for some $n \in \mathbf{N}^{>1}$ and $t(\bar{y})$ a term.

The last two formulas are dependent, because their truth value depend either only on x or on $\bar{y}$. The same is true if $n = 0$ in the first two cases. For the first two cases, with $n \neq 0$, we take advantage of the separation of variables (and this is a property of terms that we will try to obtain later in Chapter 5).

Let $\mathscr{M}$ be a monster model. Let $(a_i \mid i \in \omega)$ be indiscernible and $\bar{b} \in M^{|\bar{y}|}$. In particular $(a_i \mid i \in \omega)$ is either constant or strictly monotone. If it is constant, then $\varphi(a_i, \bar{b})$ has constant truth value. If $(a_i \mid i \in \omega)$ is strictly monotone, then there exists $i_0 \in \omega$ such that

1. either $na_i > -t(\bar{b})$ for all $i \geq i_0$. In that case $na_i + t(\bar{b}) > 0$ and $\neg(na_i + t(\bar{b}) = 0)$ for all $i \geq i_0$;
2. or $na_i < -t(\bar{b})$ for all $i \geq i_0$. In that case $\neg(na_i + t(\bar{b}) > 0)$ and $\neg(na_i + t(\bar{b}) = 0)$ for all $i \geq i_0$.

This concludes the proof. □

Our last example will be useful in Chapter 5, as one of our tasks will be to reduce dependency of expansions of $(\mathbf{Z}, +, 0, <)$ by a sparse set R to the dependency of the theory of R with the successor function, the order and predicates for congruence relations.

Example 1.2.8. Let $\mathcal{P}$ be a subset of $\mathfrak{P}(\mathbf{N})$. Let $\mathcal{L}$ be the language $\{S, S^{-1}, 0, <, P_X \mid X \in \mathcal{P}\}$, where S and S^{-1} are unary functions and P_X is a unary predicate for all $X \subset R$. Let $\mathscr{N}_{\mathcal{P}}$ be the $\mathcal{L}$-structure with domain $\mathbf{N}$, $S(n) = n + 1$, $S^{-1}(n + 1) = n$, $S^{-1}(0) = 0$, $P_X(\mathbf{N}) = X$ and $<$ is the usual order on $\mathbf{N}$. Then $T = \mathrm{Th}(\mathscr{N}_{\mathcal{P}})$ is dependent.

Proof. This is done as in [50, Section A.1.1], after noticing that S and S^{-1} are definable in $(\mathbf{N}, 0, <)$. □

This last example will be used as follows. Consider a sequence (r_n) in $\mathbf{N}$ that is strictly increasing. Let $R = \{r_n \mid n \in \mathbf{N}\}$. We equip R with the following structure in the language $\{S, S^{-1}, c, <, D_{n,k} \mid k < n \in \mathbf{N}\}$: $S(r_n) = r_{n+1}$, $S^{-1}(r_{n+1}) = r_n$, $c = r_0$, $<$ is the order induced by the one on $\mathbf{N}$ and $D_{n,k}$ is the set of elements in R that are equal to k modulo n. Let $\mathscr{R}$ denote this structure. Then $\mathscr{R}$ is definable in $\mathscr{N}_{\mathcal{P}}$ where $\mathcal{P}$ is obtained from the sets $D_{n,k}$. Thus $\mathrm{Th}(\mathscr{R})$ is dependent.

1.3 Linear recurrence sequences

In the course of Chapter 3, we will need several facts on linear recurrence sequences. We collect in this section these facts, without proofs, unless we could not find one in the literature.

We first define the Kepler limit of a general sequence of complex numbers. This limit, if it exists, indicates how fast the sequence grows.

Definition 1.3.1. Let $(r_n) \subset \mathbf{C}$. Then the *Kepler limit* of (r_n) is the following limit, if it exists in $\mathbf{C}_\infty$,

$$\lim_{n\to\infty} \frac{r_{n+1}}{r_n}.$$

We now define what is a linear recurrence sequence.

Definition 1.3.2. Let $(r_n) \subset \mathbf{C}$. Then (r_n) is a *linear recurrence sequence* if there exist $k \in \mathbf{N}^{>0}$ and $a_0, \ldots, a_{k-1} \in \mathbf{C}$ such that for all $n \in \mathbf{N}$,

$$a_0 r_n + a_1 r_{n+1} + \cdots + a_{k-1} r_{n+k-1} = r_{n+k}. \tag{1.1}$$

The polynomial $P(X) = X^k - a_{k-1}X^{k-1} - \cdots - a_0$ is called the *companion polynomial* of (r_n) associated to (1.1). The numbers $r_0, \ldots, r_{k-1}$ are called *the initial conditions* of (r_n) associated to (1.1).

Definition 1.3.3. Let (r_n) be a linear recurrence sequence. Let P be the companion polynomial of (r_n) of smallest degree. Then P is called the *minimal polynomial* of (r_n), and $\deg(P)$ is called the *order* of (r_n).

In the following result, given $P \in \mathbf{C}[X]$ of degree $k \in \mathbf{N}$, $P^-(X)$ is the polynomial

$$\sum_{i=0}^{k} a_i X^{k-i},$$

where $P(X) = a_0 + a_1 X + \cdots + a_k X^k$.

Theorem 1.3.4 ([51, Theorem 4.1.1]). *Let (r_n) be a sequence in $\mathbf{C}$. Let $P \in \mathbf{C}[X]$, $P(X) = X^k - a_{k-1}X^{k-1} - \cdots - a_0$. Let $\theta_1, \dots, \theta_\ell$ be the distinct roots of P and k_i their multiplicity. Then the following are equivalent:*

1. *(r_n) satisfy a linear recurrence relation with companion polynomial $P(X)$;*
2. *there exists $Q \in \mathbf{C}[X]$ such that $\deg(Q) < \deg(P)$ and*
$$\sum_{n \in \mathbf{N}} r_n X^n = \frac{Q(X)}{P^-(X)};$$
3. *(Binet's formula) there exists $P_1, \dots, P_\ell \in \mathbf{C}[X]$ such that $\deg P_i < k_i$ for all $i \in [\ell]$ and for all $n \in \mathbf{N}$*
$$r_n = \sum_{i=1}^{\ell} P_i(n)\theta_i^n.$$

Corollary 1.3.5 ([51, Corollary 4.2.1]). *Let (r_n) be a linear recurrence sequence and P a polynomial. In the context of Theorem 1.3.4, the following are equivalent:*

1. *P is the minimal polynomial of (r_n);*
2. *Q and P^- are relatively prime;*
3. *$\deg P_i = k_i - 1$ for all $i \in [\ell]$.*

Example 1.3.6.

1. The Fibonacci sequence Fib is the following recurrence sequence: $r_0 = 0$, $r_1 = 1$ and $r_{n+1} = r_{n+1} + r_n$. Its minimal polynomial is $X^2 - X - 1$. Fib has a Kepler limit, which is the golden ratio $(1 + \sqrt{5})/2$. We also have
$$r_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right).$$
2. for any $q \in \mathbf{C}$, the sequence (q^n) is a linear recurrence sequence, with minimal polynomial $P(X) = X - q$: $r_0 = 1$ and $r_{n+1} = qr_n$.

We define an action of $\mathbf{C}[X]$ on the set of sequences of integers. We let X act as the shift σ: for all $n \in \mathbf{N}$ and all sequence (s_n), $\sigma(s_n) = s_{n+1}$. Likewise, X^i acts as σ^i. We extend this by linearity: if $Q(X) = \sum_{i=0}^{d} a_i X^i$, then Q acts as $\sum_{i=0}^{d} a_i \sigma^i$. This action has the following property: for all $Q, Q' \in \mathbf{C}[X]$, QQ' acts as the action of Q' followed by the action of Q. Let $Q\cdot$ denote the action of Q. Note that if P is the minimal polynomial of the linear recurrence sequence (r_n), then $P_R \cdot (r_n) = (0)$.

The next proposition states that the polynomials in $\mathbf{C}[X]$ that are companion polynomials of a given linear recurrence sequence (r_n) form an ideal and this ideal is generated by the minimal polynomial of (r_n).

Proposition 1.3.7. *Let $R = (r_n)$ be a linear recurrence relation and let $Q \in \mathbf{C}[X]$, $Q(X) = \sum_{i=0}^{d} a_i X^i$. Let P_R be the minimal polynomial of R. The following are equivalent*

1. *P_R divides Q (in $\mathbf{C}[X]$);*
2. *$Q \cdot (r_n) = (0)$, that is for all $n \in \mathbf{N}$, $a_0 r_n + a_1 r_{n+1} + \cdots + a_d r_{n+d} = 0$.*

Proof. By the euclidean algorithm (in $\mathbf{C}[X]$, see [26, Theorem 2.14]), we have $Q = Q_1 P_R + Q_2$, for some $Q_1, Q_2 \in \mathbf{C}[X]$ with $\deg(Q_2) < \deg(P_R)$.

First assume that P_R divides Q. Then $Q = Q_1 P_R$. Thus, $Q \cdot (r_n) = (Q_1 P_R) \cdot (r_n) = Q_1 \cdot (P_R \cdot (r_n)) = Q_1 \cdot (0) = (0)$. This implies that $Q \cdot (r_n) = (0)$.

Second assume that $Q \cdot (r_n) = (0)$. Since $P_R \cdot (r_n) = (0)$, we get that $Q_2 \cdot (r_n) = (0)$, which contradicts the minimality of P_R, unless $Q_2 = 0$. So P_R must divide Q. □

The existence of a Kepler limit for a given linear recurrence sequence have been investigated in full generality by A. Fiorenza and G. Vincenzi in [22]. We reformulate here [22, Theorem 2.3].

From now on when (r_n) a linear recurrence sequence is considered with companion polynomial P, we assume the following. Let $\theta_1, \ldots, \theta_\ell$ be roots of P such that for all $n \in \mathbf{N}$

$$r_n = \sum_{i=1}^{\ell} P_i(n)\theta_i^n,$$

where for all $i \in [\ell]$ $\deg(P_i) < k_i$, k_i is the multiplicity of θ_i as a root of P, and $P_i \neq 0$. Assume furthermore that $|\theta_i| \geq |\theta_{i+1}|$ and $|\theta_i| = |\theta_{i+1}| \Rightarrow \deg(P_i) \geq \deg(P_{i+1})$ for all $i \in [\ell]$. In summary, we order the roots of P decreasingly according to their modulus first and then according to the degree of their coefficient in Binet's formula.

Theorem 1.3.8 ([22, Theorem 2.3])**.** *Assume that that $\{n \in \mathbf{N} \mid r_n = 0\}$ is finite. Then (r_n) has a Kepler limit if and only if $|\theta_1| > |\theta_2|$ or $\deg(p_1) > \deg(p_2)$. Furthermore, if (r_n) has a Kepler limit, then it is equal to θ_1.*

A remarkable result on linear recurrence sequence is the so called Skolem-Mahler-Lech Theorem, which states that the set of indices at which a linear recurrence sequence takes the value 0 is an ultimately periodic subset of $\mathbf{N}$.

Theorem 1.3.9 ([21, Theorem 2.1])**.** *Let (r_n) be a linear recurrence sequence in $\mathbf{C}$. Then there exist a finite set $F \subset \mathbf{N}$ and $(a_1, b_1), \ldots, (a_\ell, b_\ell) \in \mathbf{N} \times \mathbf{N}^{>0}$ such that*

$$\{n \in \mathbf{N} \mid r_n = 0\} = F \cup \bigcup_{i=1}^{\ell} (a_i + b_i \mathbf{N}).$$

An important class of linear recurrence sequences are the so called non-degenerate ones.

Definition 1.3.10. Let (r_n) be a linear recurrence sequence with minimal polynomial P. Then (r_n) is called *non-degenerate* if for all $\theta_1 \neq \theta_2$, two roots of P, θ_1/θ_2 is not a root of unity.

One useful tool in the study of linear recurrence sequence is the following result.

Theorem 1.3.11 ([21, Theorem 1.2]). *Let (r_n) be a linear recurrence sequence. Then there exists $m \in \mathbf{N}$ such that for all $i \in \{0, \dots, m-1\}$, the sequence (r_{i+nm}) is either identically 0 or non-degenerate.*

This theorem often allows to study general linear recurrence sequence via their non-degenerate subsequences.

We now concentrate on linear recurrence sequences in $\mathbf{Z}$.

Proposition 1.3.12. *Let (r_n) be a linear recurrence sequence in $\mathbf{Z}$. Then for all $k \in \mathbf{N}$, the sequence $(r_n \mod k)$ is ultimately periodic.*

Proof. Let d be the order of (r_n) and $P(X) = X^d - a_{d-1}X^{d-1} + \cdots + a_0$ be its minimal polynomial. Consider the tuples $x_n = (r_{n+d-1} \mod k, \dots, r_n \mod k) \in (\mathbf{Z}/k\mathbf{Z})^d$. Since $(\mathbf{Z}/k\mathbf{Z})^d$ is finite, there exists $n_0, \ell \in \mathbf{N}$ such that $x_{n_0} = x_{n_0+\ell}$. We show by induction that $x_n = x_{n+\ell}$ for all $n \geq n_0$. Let $n > n_0$. Let us show that $x_n = x_{n+\ell}$, that is

$$(r_{n+d-1} \mod k, \dots, r_n \mod k) = (r_{n+\ell+d-1} \mod k, \dots, r_{n+\ell} \mod k).$$

We know by induction hypothesis that for all $i \in \{n_0, \dots, n-1\}$, $x_i = x_{i+\ell}$. This implies in particular that for all $i \in \{n_0, \dots, n-1\}$, $r_{i+d} \mod k = r_{i+\ell+d} \mod k$, so that $r_{n+\ell+d-1} \mod k = r_{n+d-1} \mod k$. And since $x_{n-1} = x_{n-1+\ell}$, we have that $r_{n+i} \mod k = r_{n+i+\ell} \mod k$ for all $i \in \{0, \dots, d-2\}$. Thus $x_n = x_{n+\ell}$. □

We end the preliminaries on linear recurrence sequences with an answer to the following question: under what conditions on $Q \in \mathbf{Q}[X]$ of degree $d > 0$ and $r_0, \dots, r_{d-1} \in \mathbf{Z}$, do we have that the linear recurrence with companion polynomial Q and initial conditions $\bar{r}$ is always in $\mathbf{Z}$?

Proposition 1.3.13. *Let $P, Q \in \mathbf{Z}[X]$, with $\deg(P) < \deg(Q)$. Assume that P^- and Q are relatively prime in $\mathbf{Z}[X]$ and that there exists $(r_n) \subset \mathbf{C}$ such that*

$$\sum_{i \in \mathbf{N}} r_n X^n = \frac{Q(X)}{P^-(X)}.$$

Then $r_n \in \mathbf{Z}$ for all $n \in \mathbf{N}$ if and only if $P^-(0) = \pm 1$.

Proof. The only if part of the proposition is done in [47, Lemmas I and II].

For the if part, assume that $P^-(0) = \pm 1$. So, there exists $R \in \mathbf{Z}[X]$ such that $P^-(X) = \pm(1 - XR(X))$. Therefore

$$\frac{Q(X)}{P^-(X)} = \pm Q(X) \sum_{i \in \mathbf{N}} (XR(X))^i.$$

This shows that r_n is in $\mathbf{Z}$ for all $\in \mathbf{N}$. □

CHAPTER

2

Tame pairs

In this chapter, we make precise the general theme of this dissertation.

In the first section, we discuss results that concern tame *pairs*, with a focus on stability and dependency. Here, a pair is just an $\mathcal{L}$-structure $\mathscr{M}$ expanded by a subset A of M, denoted $\mathscr{M}_A$. Among other things, preservation of stability and dependency are discussed for pairs and these results show that tameness of the induced structure on A (that is the trace of definable subsets of $\mathscr{M}$ on the cartesian powers of A) is important, as well as a syntactical condition of the pair $\mathscr{M}_A$ called *boundedness*. Because this notion can be difficult to check, we list several more manageable properties that imply boundedness, most notably, in the stable context, the lack of a relative notion of the finite cover property for $\mathscr{M}$. Most results of this section are due to E. Casanovas and M. Ziegler for the stable case and A. Chernikov and P. Simon for the dependent case.

In the second section, we give an overview of known results on pairs $\mathscr{Z}_A$, where $A \subset \mathbf{Z}$. We review the work of D. Palacín and R. Sklinos, B. Poizat, G. Conant and C. Laskowski on expansions of $\mathscr{Z}$ by a unary predicate. Apart from various examples, we promote the work of G. Conant and C. Laskowski who managed to show that the stability of a pair $\mathscr{Z}_A$ depends only on the stability of the induced structure on A. We also take some time to discuss expansions by a set enumerated by a linear recurrence sequence, where interesting phenomena occur. We end this section with a review of pairs $\mathscr{Z}_A$, where A is the set of integers whose absolute value is a prime number.

2.1 Stable and dependent pairs

The main theme of this thesis is the study of expansions of $(\mathbf{Z}, +, 0)$ and $(\mathbf{Z}, +, 0, <)$ by a unary predicate. Of particular interest to us is the preservation of properties such as superstability in case of $(\mathbf{Z}, +, 0)$ and dependency in case of $(\mathbf{Z}, +, 0, <)$. These question arise at a very general level in abstract model theory, in the study of *pairs*. Let us give the setting in which the study of pairs is formulated.

We start with a complete $\mathcal{L}$-theory T with infinite models, a model $\mathscr{M}$ of T and $A \subset M$ infinite. The pair associated to $\mathscr{M}$ and A is the following structure: setting

$\mathcal{L}_R = \mathcal{L} \cup \{R\}$, where R is a new unary predicate symbol, we consider the natural $\mathcal{L}_R$-expansion on $\mathscr{M}$ with $R(M) = A$, which we note $\mathscr{M}_A$. In this setting, the main problem in the study of pairs is the following: under what conditions of $\mathscr{M}$ and A do nice properties of $\mathscr{M}$ (such as stability and dependency) transfer to the pair $\mathscr{M}_A$? In what follows, we concentrate mostly on the transfer of stability and dependency of $\mathscr{M}$ to a pair $\mathscr{M}_A$. Therefore, we assume now that $\mathscr{M}$ is either stable on dependent and use the adjective *tame* to refer to these properties.

The first relevant observation in this context is that the trace of definable subsets on A must be tame. This means that for all $\mathcal{L}$-formula $\varphi(\bar{x})$, $|\bar{x}| = n$, the formula $\varphi(\bar{x}) \wedge R(x_1) \wedge \cdots \wedge R(x_n)$ must be tame, that is either stable or dependent, depending on which tameness notion is considered. Therefore, the following definition is useful.

Definition 2.1.1. To each $\mathcal{L}$-formula $\varphi(x_1, \ldots, x_n)$, we associate a new n-ary predicate $R_{\varphi,n}$ and we denote by $\mathcal{L}_{\text{ind}}$ the language

$$\{R_{\varphi,n} \mid \varphi(x_1, \ldots, x_n) \text{ is an } \mathcal{L}\text{-formula}\}.$$

The *induced structure* on A (by $\mathscr{M}$), denoted A_{ind}, is the $\mathcal{L}_{\text{ind}}$-structure whose domain is A and $R_{\varphi,n}(A) = \varphi(M^n) \cap A^n$.

With this terminology, the observation we made before is that if $\mathscr{M}_A$ is tame, then A_{ind} is tame. However, it is not enough in general that A_{ind} be tame in order to deduce that $\mathscr{M}_A$ remains tame, although we shall see in the next section of this chapter that it is sometimes the case. One recurrent sufficient condition is the *boundedness* of the pair $\mathscr{M}_A$, which we recall now.

Definition 2.1.2. We say that the pair $\mathscr{M}_A$ is *bounded* if for all $\mathcal{L}_R$-formulas $\varphi(\bar{x})$ there is an $\mathcal{L}$-formula $\psi(\bar{x}, \bar{y})$ such that φ is equivalent (in $\text{Th}(\mathscr{M}_A)$) to the formula

$$Q_1 y_1 \in R \ldots Q_n y_n \in R\, \psi(\bar{x}, \bar{y}),$$

where $Q_i \in \{\exists, \forall\}$.

Now we have the following preservation results. The first is due to E. Casanovas and M. Ziegler and the second is due to A. Chernikov and P. Simon.

Theorem 2.1.3 ([11, Proposition 3.1])**.** *Let $\mathscr{M}$ be an $\mathcal{L}$-structure and let $A \subset M$. Suppose that the pair $\mathscr{M}_A$ is bounded. Then for all $\lambda \geq |\mathcal{L}|$, if $\mathscr{M}$ and A_{ind} are λ-stable, then $\mathscr{M}_A$ is λ-stable.*

Theorem 2.1.4 ([12, Corollary 2.6])**.** *Let $\mathscr{M}$ be an $\mathcal{L}$-structure and let $A \subset M$. Suppose that the pair $\mathscr{M}_A$ is bounded. Then $\mathscr{M}_A$ is dependent if both $\mathscr{M}$ and A_{ind} are dependent.*

Checking that a pair is bounded can be difficult, unless some desirable property such as quantifier elimination is known. Hence the need for criteria to get boundedness. On the stable side, the following definition is relevant.

Definition 2.1.5.
1. We say that $\mathscr{M}$ *does not have the finite cover property* (in short: $\mathscr{M}$ is nfcp) if for all formulas $\varphi(x, \bar{y})$, there exists $k \in \mathbf{N}$ such that for all $X \subset M^{|\bar{y}|}$ if the set
$$\{\varphi(x, \bar{m}) \mid m \in X\}$$
is k-consistent, then it is consistent.
2. We say that $\mathscr{M}$ *has* nfcp *over* A if for all formulas $\varphi(x, \bar{y}, \bar{z})$, there exists $k \in \mathbf{N}$ such that for all $\bar{m} \in M^{|\bar{z}|}$, $X \subset A^{|\bar{y}|}$, if the set
$$\{\varphi(x, \bar{a}, \bar{m}) \mid \bar{a} \in X\}$$
is k-consistent, then it is consistent.

The relevance of nfcp is captured in the following result, called the *f.c.p theorem*, due to S. Shelah.

Theorem 2.1.6 ([17, Fact 3.9.(a)])**.** *(We do not assume that $\mathscr{M}$ is tame here) $\mathscr{M}$ is nfcp if and only if $\mathscr{M}$ is stable and* $\mathrm{Th}(\mathscr{M})$ *eliminates $\exists^\infty$ in all imaginary sorts.*

One example of nfcp structure is $(\mathbf{Z}, +, 0)$ and this can be seen using quantifier elimination in the expanded language $\{+, -, 0, D_n \mid n \in \mathbf{N}^{>1}\}$, which implies that a definable subset is a boolean combination of a finite sets and cosets of subgroups.

Definition 2.1.7. We say that A is *small* if there is an $\mathcal{L}_R$-structure $\mathscr{N}_{R(N)}$ elementary equivalent to $\mathscr{M}_A$ such that for all finite subsets B of N, any type in $\mathcal{L}$ over $B \cup R(N)$ is realized in $\mathscr{N}$.

It turns out that the smallness of A and nfcp over A implies the boundedness of $\mathscr{M}_A$.

Theorem 2.1.8 ([11, Proposition 2.1])**.** *Assume that A is small. If $\mathscr{M}$ is nfcp over A then $\mathscr{M}_A$ is bounded.*

Now the cost of using nfcp is the need to show that A is small. A way of avoiding this is given in the following result, extracted from the proof of [11, Proposition 2.1] by Palacín and Sklinos (see also [37, Lemma 3.5]).

Theorem 2.1.9. *Let $\mathscr{M}$ be a stable $\mathcal{L}$-structure and $A \subset M$. Assume that for all $\mathcal{L}$-formulas $\varphi(\bar{x}, y, \bar{z})$, there exists $k \in \mathbf{N}$ such that*

$$\mathscr{M}_A \models \forall \bar{x} \left(\left(\forall \bar{z}_1 \in R \ldots \forall \bar{z}_k \in R \exists y \bigwedge_{j \in [k]} \varphi(\bar{x}, y, \bar{z}_j) \right) \to \exists y \forall \bar{z} \in R \varphi(\bar{x}, y, \bar{z}) \right).$$

Then $\mathscr{M}_A$ is bounded.

Proof. This is done by induction on the number of quantifiers of a formula. For the case where there are no quantifiers, we first point out that $R(x)$ is equivalent to $\exists y \in R(x = y)$. This shows that every quantifier-free formula is bounded. Now assume that $\varphi(\bar{x}, y)$ is bounded and let us show that $\exists y \varphi(\bar{x}, y)$ is equivalent to a bounded formula.

As $\varphi(\bar{x}, y)$ is bounded, there exists an $\mathcal{L}$-formula $\psi(\bar{x}, y, \bar{z})$ such that $\varphi(\bar{x}, y)$ is of the form

$$Q_1 z_1 \in R \ldots Q_n z_n \in R \psi(\bar{x}, y, \bar{z}),$$

where $n = |\bar{z}|$ and $Q_i \in \{\exists, \forall\}$ for all $i \in [n]$. Since $\mathscr{M}$ is stable, there exists an $\mathcal{L}$-formula $\theta(\bar{z}, \bar{w})$ such that for all $\mathscr{N}_B \models \mathrm{Th}(\mathscr{M}_A)$, for all $\bar{m} \in N^{|\bar{x}|}$ and all $m' \in N$, there exists $\bar{b} \in B^{|\bar{w}|}$ such that for all $\bar{a} \in B^n$

$$\psi(\bar{x}, y, \bar{a}) \in \mathrm{tp}^{\mathcal{L}}\left(\bar{m}, m'/A\right) \text{ if and only if } \mathscr{N} \models \theta(\bar{a}, \bar{b}).$$

As a result, $\varphi(\bar{x}, y)$ is equivalent to the bounded formula

$$\exists \bar{w} \in R \big(\forall \bar{z} \in R (\psi(\bar{x}, y, \bar{z}) \leftrightarrow \theta(\bar{z}, \bar{w})) \land Q_1 z_1 \in R \ldots Q_n z_n \in R \theta(\bar{z}, \bar{w}) \big).$$

Let $\tau(\bar{x}, y, \bar{z}, \bar{w})$ be the formula $\psi(\bar{x}, y, \bar{z}) \leftrightarrow \theta(\bar{z}, \bar{w})$. By assumption, there exists $k \in \mathbf{N}$ such that

$$\mathscr{M}_A \models \forall \bar{x} \forall \bar{w} \left(\left(\forall \bar{z}_1 \in R \ldots \forall \bar{z}_k \in R \exists y \bigwedge_{j \in [k]} \tau(\bar{x}, y, \bar{z}_j, \bar{w}) \right) \to \exists y \forall \bar{z} \in R \tau(\bar{x}, y, \bar{z}, \bar{w}) \right).$$

Thus, $\exists y \varphi(\bar{x}, y)$ is equivalent to the bounded formula

$$\exists \bar{w} \in R \left(\left(\forall \bar{z}_1 \in R \ldots \forall \bar{z}_k \in R \exists y \bigwedge_{j \in [k]} \tau(\bar{x}, y, \bar{z}_j, \bar{w}) \right) \right.$$
$$\left. \land Q_1 z_1 \in R \ldots Q_n z_n \in R \theta(\bar{z}, \bar{w}) \right).$$ □

Now concerning the dependent case, nfcp is useless to infer boundedness of a pair, in view of Shelah's f.c.p Theorem. However A. Chernikov and P. Simon developed a candidate for a version of nfcp in the dependent context, called *dnfcp* (nfcp for definable sets of parameters).

Definition 2.1.10 ([13, Definition 38])**.** We say that $\mathscr{M}$ *has* dnfcp *over* A if for all formulas $\varphi(x, \bar{y}, \bar{z})$, there exists $k \in \mathbf{N}$ such that for all $\bar{m} \in M^{|\bar{z}|}$, if the set

$$\{\varphi(x, \bar{a}, \bar{m}) \mid \bar{a} \in A^{|\bar{y}|}\}$$

is k-consistent, then it is consistent.

Observe that the content of Theorem 2.1.9 precisely states that in the stable context, if $\mathscr{M}$ is dnfcp over A in a strong form, then $\mathscr{M}_A$ is bounded. Also, we have that nfcp over A implies dnfcp over A.

In the dependent context, dnfcp over A and smallness of A is not quite sufficient to deduce that $\mathscr{M}_A$ is bounded. What is missing, in comparison to stable case, is that A is not necessarily uniformly stably embedded, a property automatically satisfied in the stable context. This property states in particular that the trace on A of a definable set with parameters in M can be defined with parameters in A.

Definition 2.1.11. We say that A is *uniformly stably embedded* if for any $\mathcal{L}$-formula $\varphi(\bar{x}, \bar{y})$ there exists a formula $\psi(\bar{x}, \bar{z})$ such that for all $\bar{m} \in M^{|\bar{y}|}$, there exists $\bar{a} \in A^{|\bar{z}|}$ such that

$$\varphi(A^{|\bar{x}|}, \bar{m}) = \psi(A^{|\bar{x}|}, \bar{a}).$$

Theorem 2.1.12 ([13, Theorem 37])**.** *Assume that A is small and uniformly stably embedded. If $\mathscr{M}$ has dnfcp over A, then $\mathscr{M}_A$ is bounded.*

Even though in general sets in dependent theories are not uniformly stably embedded, a weak form of this property still holds.

Theorem 2.1.13 ([50, Theorem 3.13])**.** *Let $\varphi(\bar{x}, \bar{y})$ be an $\mathcal{L}$-formula and $\bar{b} \in M^{|\bar{y}|}$. Assume that $\varphi(\bar{x}, \bar{y})$ is dependent. Then there exists an elementary extension $\mathscr{M}'_{A'}$ of $\mathscr{M}_A$, an $\mathcal{L}$-formula $\psi(\bar{x}, \bar{z})$ and $\bar{a} \in A'^{|\bar{z}|}$ such that*

$$\varphi(A^{|\bar{x}|}, \bar{b}) \subset \psi(A'^{|\bar{x}|}, \bar{a}) \subset \varphi(A'^{|\bar{x}|}, \bar{b}).$$

The formula $\psi(\bar{x}, \bar{a})$ is called an *honest definition* for $\varphi(\bar{x}, \bar{b})$. For the proof of [12, Corollary 2.6], A. Chernikov and P. Simon introduced a relative notion of honest definition, which we will need in the course of Chapter 5. We therefore end this section with a presentation of this notion and the relevant properties.

In what follows, we assume that T is dependent and $\mathscr{N}$ is a monster model of T containing $\mathscr{M}$.

Definition 2.1.14. Let $\mathscr{M} \prec \mathscr{N}$, $\varphi(\bar{x}, \bar{y})$ an $\mathcal{L}_R$-formula and $\bar{a} \in M^{|\bar{y}|}$.

1. We say that $\varphi(\bar{x},\bar{y})$ has an *honest definition over R* if there exists an $\mathcal{L}_R$-formula $\theta(\bar{x},\bar{z})$ and $\bar{c} \in R(N)^{|\bar{z}|}$ such that $\theta(R(M),\bar{c}) = \varphi(R(M),\bar{a})$ and

$$\mathscr{N}_{R(N)} \models \forall \bar{x} \in R(\theta(\bar{x},\bar{c}) \to \varphi(\bar{x},\bar{a})).$$

2. We say that $\varphi(\bar{x},\bar{y})$ is *dependent over R* if there do not exist $(\bar{a}_i \mid i \in \omega)$ in $R(N)^{|\bar{x}|}$ $\mathcal{L}_R$-indiscernible and $\bar{b} \in N^{|\bar{y}|}$ such that $\mathscr{N} \models \varphi(\bar{a}_i,\bar{b})$ if and only if i is even.

Honest definition behave well under existential quantifications, conjunctions and disjunctions.

Lemma 2.1.15 ([12, Lemma 2.1]). *Let $\mathscr{M} \prec \mathscr{N}$. Let $\varphi(\bar{x}_1,\bar{x}_2,\bar{y})$ be an $\mathcal{L}_R$ formula and let $\bar{a} \in M^{\bar{y}}$. Assume that $\theta(\bar{x}_1,\bar{x}_2,\bar{c})$ is an honest definition over R for $\varphi(\bar{x}_1,\bar{x}_2,\bar{a})$. Then $\exists \bar{x}_1 \in R\theta(\bar{x}_1,\bar{x}_2,\bar{c})$ is an honest definition over R for $\exists \bar{x}_1 \in R\varphi(\bar{x}_1,\bar{x}_2,\bar{a})$.*

Lemma 2.1.16. *Let $\mathscr{M} \prec \mathscr{N}$, $\varphi_1(\bar{x},\bar{y})$, $\varphi_2(\bar{x},\bar{w})$ $\mathcal{L}_R$-formulas, $\bar{a}_1 \in N^{|\bar{y}|}$ and $\bar{a}_2 \in N^{|\bar{z}|}$. Assume that $\theta_1(\bar{x},\bar{c}_1)$ is an honest definition over R for $\varphi_1(\bar{x},\bar{a}_1)$ and $\theta_2(\bar{x},\bar{c}_2)$ is an honest definition over R for $\varphi_2(\bar{x},\bar{a}_2)$. Then $\theta_1(\bar{x},\bar{c}_1) \wedge \theta_2(\bar{x},\bar{c}_2)$ (resp. $\theta_1(\bar{x},\bar{c}_1) \vee \theta_2(\bar{x},\bar{c}_2)$) is an honest definition over R for $\varphi_1(\bar{x},\bar{a}_1) \wedge \varphi_2(\bar{x},\bar{a}_2)$ (resp. $\varphi_1(\bar{x},\bar{a}_1) \vee \varphi_2(\bar{x},\bar{a}_2)$).*

The next lemma gives a link between the dependency of $\mathrm{Th}(\mathscr{N}_{R(N)})$ and the dependency of $R(N)_{\mathrm{ind}}$.

Lemma 2.1.17 ([12, Lemma 2.3]). *Let $(\bar{a}_i \mid i \in \omega)$ be $\mathcal{L}_R$-indiscernible, $(\bar{b}_{2i} \mid i \in \omega)$ a sequence of elements in $R(N)^m$ and $\delta(\bar{x}_1,\ldots,\bar{x}_n,\bar{y}_1,\ldots,\bar{y}_n)$ an $\mathcal{L}_R$-formula with $|\bar{x}_i| = |\bar{b}_0|$ and $|\bar{y}_i| = |\bar{a}_0|$ for all $i \in [n]$ and n is even. Assume the following:*

1. *$\mathscr{N} \models \delta(\bar{b}_{2i_1},\ldots,\bar{b}_{2i_n},\bar{a}_{2i_1},\ldots,\bar{a}_{2i_n})$ for all $i_1,\ldots,i_n \in \omega$;*
2. *$\delta(\bar{x}_1,\ldots,\bar{x}_n,\bar{a}_0,\ldots,\bar{a}_{n-1})$ has an honest definition over A $\theta(\bar{x}_1,\ldots,\bar{x}_n,\bar{c})$ such that $\exists \bar{x}_1\bar{x}_3\ldots\bar{x}_{n-1} \in R\theta(\bar{x}_1,\ldots,\bar{x}_n,\bar{z})$ is dependent over R.*

Then there exists $i_1,\ldots,i_n \in \omega$ with $i_j \equiv_2 j$ and $(\bar{b}_{i_j} \mid j \equiv_2 1, j \leq n)$ a tuple in $R(N)^m$ such that $\mathscr{N} \models \delta(\bar{b}_{i_1},\ldots,\bar{b}_{i_n},\bar{a}_{i_0},\ldots,\bar{a}_{i_n})$.

Remark 2.1.18. The statement of [12, Lemma 2.3] actually requires A_{ind} to be dependent. However the proof of [12, Lemma 2.3] shows that the dependency over R of $\exists \bar{x}_1\bar{x}_3\ldots\bar{x}_{n-1} \in R\theta(\bar{x}_1,\ldots,\bar{x}_n,\bar{z})$ is sufficient.

2.2 Tame expansions of the group of integers

In this section, we give a partial account of known results of tame expansions of $(\mathbf{Z},+,0)$ and $(\mathbf{Z},+,0,<)$ by a unary predicate. We concentrate on three notions of tameness: stability, dependency and simplicity. We retain the setting of the previous sections: we consider expansions of $(\mathbf{Z}, +, 0)$ (resp. $(\mathbf{Z}, +, 0, <)$) in the language $\mathcal{L} = \{+, 0, R\}$ (resp. $\mathcal{L}_< = \{+, 0, <, R\}$) where R is a unary predicate.

We start with stability. The following question, attributed to J. Goodrick by D. Palacín and R. Sklinos in [37], is still open.

Question 2.2.1 ([37, Question 4.7]). *Characterize the subsets A of $\mathbf{Z}$ for which $(\mathbf{Z}, +, 0, A)$ is (super-)stable.*

The first explicit examples of superstable expansions of $(\mathbf{Z}, +, 0)$ are by the sets $A_q = \{q^n \mid n \in \mathbf{N}\}$, $q \in \mathbf{N}^{>1}$, and $B = \{n! \mid n \in \mathbf{N}\}$. Both expansions are the subject of [37] and expansions by A_q have been independently considered by B. Poizat in [44]. The superstability of the pair $(\mathbf{Z}, +, 0, A_q)$ is also a consequence of the work of R. Moosa and T. Scanlon [36].

Theorem 2.2.2. *Let $q \in \mathbf{N}^{>1}$. Then*

1. ([36, Theorem 6.11], [37, Theorem 2] and [44, Théorème 25]) $\mathrm{Th}(\mathbf{Z}, +, 0, A_q)$ *is superstable of Lascar rank ω;*
2. ([37, Proposition 4.2]) $\mathrm{Th}(\mathbf{Z}, +, 0, B)$ *is superstable of Lascar rank ω.*

The approach in [37] rely on Theorem 2.1.3 while the approach in [44] rely on a back-and-forth argument to characterize ω-saturated elementary extensions of $(\mathbf{Z}, +, 0, A_q)$.

Let us briefly outline the proof of [44, Théorème 25]. The main step is to show that a ω-saturated model $\mathscr{M}$ of $\mathrm{Th}(\mathbf{Z}, +, 0, 1, A_q)$ is of the form

$$\hat{\mathbf{Z}} \times D_1 \times D_2,$$

where $\hat{\mathbf{Z}}$ is the profinite completion of $\mathbf{Z}$, D_1 and D_2 are divisible groups of infinite dimension when considered as $\mathbf{Q}$-vector spaces and $\hat{\mathbf{Z}} \times D_1$ contains the group generated by $R(M)$. This is done using a back-and-forth argument observing the following crucial property: any element in the group G generated by $R(M)$ can be uniquely written in the form $n_0 + n_1 a_1 + \cdots + n_k a_k$ where $n_0 \in \mathbf{Z}$, $n_1 \ldots, n_k \in \mathbf{Z}$ are prime to q and $a_1, \ldots, a_k \in R(M) \setminus A_q$ are in pairwise different orbits, that is $a_i \notin \{q^n a_j \mid n \in \mathbf{Z}\}$ for all $i \neq j$. Then using the characterization of ω-saturated, $\mathrm{Th}(\mathbf{Z}, +, 0, 1, A_q)$ is shown to be superstable by calculating $U(a/\emptyset)$ for all a in all ω-saturated models. For the rank calculation, it is observed that the rank over $\emptyset$ of an element a in G is equal to the number of elements in $R(M) \setminus A_q$ in the decomposition of a. This shows that the rank is at least ω. For the other inequality, it is observed that the principal generic type is the type of an element that is infinitely divisible, which has rank $\leq \omega$.

As we said earlier, the approach of D. Palacín and R. Sklinos in [37] relies on Theorem 2.1.3 and is quite different from the one of B. Poizat. As we shall give the details of the calculation of the rank in Chapter 3 (see Theorem 3.5.1), we only sketch the arguments needed to show superstability, in the case of A_q. Let us however point out that [37, Theorem 1] states that $(\mathbf{Z}, +, 0)$ has no proper expansion of finite Lascar rank.

The proof of [37, Theorem 2] is done in two main steps

1. first they show that $\mathrm{Th}(\mathbf{Z}, +, 0, 1, A_q)$ is bounded. This is mainly due to the fact that $(\mathbf{Z}, +, 0, 1)$ has dnfcp over A_q ([37, Lemmas 3.3 and 3.4]) and a version of this argument is done in Chapter 3, see Corollary 3.4.6. The main point here is that a finite union of sets of the form

$$\{n_0 + n_1 a_1 + \cdots + n_k a_k \mid \bar{a} \in A_q^k\},$$

where $\bar{n} \in \mathbf{Z}^{k+1}$, cannot cover a non-trivial subgroup of $\mathbf{Z}$;

2. second they analyze traces of equations and congruence relations on A_q. For $k < n$, let $X_{q,k,n}$ be the elements of A_q that are equal to k modulo n. The relevant case is when k and n are coprime and n and q are coprime. Then we have (see [37, Lemma 3.9 and Remark 3.10])

$$X_{q,k,n} = \{q^{m_0 + \varphi(n)m} \mid m \in \mathbf{N}\},$$

where m_0 is minimal such that $q^{m_0} \equiv_n k$ and φ is Euler's phi function. Now given an equation $n_1 x_1 + \cdots + n_k x_k = \ell$, $\bar{n} \in \mathbf{Z}^k$ and $\ell \in \mathbf{Z}$, setting S to be the set of non-degenerate solutions (that is solutions for which no proper sub-sum vanishes) in $\mathbf{Z}^k$ of this equation, [37, Lemma 3.11] states that $S \cap A_q^k$ contains only tuples $(q^{m_1}, \ldots, q^{m_k})$ such that $\max\{|m_i - m_j| \mid i, j \in [k]\}$ is bounded by some constant only depending on $\bar{n}$, ℓ and k. As a result of this analysis, they show that $A_{q,\mathrm{ind}}$ is superstable by interpreting it in the superstable structure $(\mathbf{N}, s, D_{n,k} \mid k < n \in \mathbf{N})$, where s is the successor function and $D_{n,k}$ is the set of natural numbers equal to k modulo n.

After the publications of [37, 44], the examples treated there were generalized independently in two directions by G. Conant ([16]) and by F. Point and the author ([29]). These generalizations try to capture the exponential growth of the examples in [37, 44], via the notion of *geometrically sparse set* (see below) by G. Conant and by the existence of a Kepler limit by F. Point and the author. G. Conant published two other papers on superstable expansions of $(\mathbf{Z}, +, 0)$ [15, 17], the last one in collaboration with C. Laskowski, which also uses techniques from [29]. In the papers [16, 15, 17], the results of E. Casanovas and M. Ziegler are also used to analyze superstable expansions of $(\mathbf{Z}, +, 0)$. What's more, Theorem 2.1.3 is improved in two ways. Let $A \subset \mathbf{Z}$ and consider $(\mathbf{Z}, +, 0, A)$. We first need two definitions.

Definition 2.2.3. We let A^0_{ind} be the reduct of A_{ind} in the language

$$\mathcal{L}^0_{\text{ind}} = \{R_{\varphi,r} \mid \varphi(\bar{x}) \text{ is of the form } a_1x_1 + \cdots + a_nx_n = 0, \bar{a} \in \mathbf{Z}^n\}.$$

Therefore A^0_{ind} is the induced structure on A by equations.

Definition 2.2.4. Let $\mathscr{M}$ be an $\tilde{\mathcal{L}}$-structure. Let $\tilde{\mathcal{L}}^1$ be the expansion of $\tilde{\mathcal{L}}$ by new unary predicates R_X, for all $X \subset M$. We let $\mathscr{M}^1$ be the natural $\tilde{\mathcal{L}}^1$-expansion of $\mathscr{M}$.

The first improvement is the reduction of the stability of A_{ind} to the stability of A^0_{ind}.

Theorem 2.2.5 ([16, Corollary 5.7])**.** *If A^0_{ind} is definably interpretable in a structure $\mathscr{M}$ such that $\mathscr{M}^1$ is λ-stable, then A_{ind} is λ-stable.*

The second improvement is that boundedness is automatic for pairs $(\mathbf{Z}, +, 0, A)$, $A \subset \mathbf{Z}$.

Theorem 2.2.6 ([17, Theorem 2.14])**.** *For all $A \subset \mathbf{Z}$, the pair $(\mathbf{Z}, +, 0, A)$ is bounded.*

Remark 2.2.7. More generally, [17, Theorem 2.8] states that for any *weakly minimal* complete theory T, $\mathscr{M}_0 \prec \mathscr{M} \models T$ and $A \subset M$, $(\mathscr{M}, A, c_m \mid m \in M_0)$ is bounded. (Recall that a theory is weakly minimal if it is superstable of Lascar rank 1, see also [5, §5]).

We now summarize the examples of superstable expansions of $(\mathbf{Z}, +, 0)$ given in [16, 15, 17]. Let $A \subset \mathbf{Z}$. Then $(\mathbf{Z}, +, 0, A)$ is superstable in the following cases:

1. ([16, Theorem A]) A is geometrically sparse, that is there exists $f : A \to \mathbf{R}^{>0}$ such that $\sup\{|a - f(a)| \mid a \in A\}$ is finite and the set $\{s/t \mid s, t \in f(A), t \leq s\}$ is closed and discrete. Instances of geometrically sparse sets are A_q, $\{n! \mid n \in \mathbf{N}\}$ and Fib (see Example 1) or sets enumerated by a sequence (a_n) such that

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \infty.$$

   We shall compare the notion of a geometrically sparse set with our notion of regular set later on in Section 3.6;

2. ([15, Theorem 3.1]) A is an infinite subset of a finitely generated multiplicative submonoid of $\mathbf{N}$;

These examples highlight the fact that superstable expansions of $(\mathbf{Z}, +, 0)$ are often by sparse sets, that is sets that can be enumerated by fast growing sequences. However, this is not sufficient nor necessary:

1. ([15, Corollary 3.14]) there exists $A \subset \mathbf{Z}$ such that $(\mathbf{Z}, +, 0, A)$ is superstable and A is enumerated by a sequence (a_n) such that

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = 1.$$

   This set can be chosen as a finitely generated multiplicative submonoid generated by $n, m \in \mathbf{N}$, as long as n and m are multiplicatively independent;

2. ([15, Theorem 4.8]) the expansion of $(\mathbf{Z}, +, 0)$ by the set $\{2^n + n \mid n \in \mathbf{N}\}$ is unstable. So having an enumeration by a sequence (a_n) such that

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} \in \mathbf{R}$$

   does not guarantee superstability.

We end this short review of superstable expansions of $(\mathbf{Z}, +, 0)$ with the case where A can be enumerated by a linear recurrence sequence (a_n). The examples given in [16, 15] and Chapter 3 are characterized by one of the following properties: (a_n) has a Kepler limit θ and either

1. in the papers [16, 15] the minimal polynomial of (a_n) has exactly one root of modulus > 1, namely θ;
2. in Chapter 3, the minimal polynomial of (a_n) is the minimal polynomial of θ.

These examples were later on generalized in the following result.

Theorem 2.2.8 ([17, Theorem 4.9])**.** *Let $A \subset \mathbf{Z}$ be enumerated by a linear recurrence sequence (r_n) with minimal polynomial P. Assume that no repeated root of P is a root of unity. Then* $\mathrm{Th}(\mathbf{Z}, +, 0, A)$ *is superstable of rank ω.*

While the converse of the previous theorem does not hold (for instance $\mathbf{Z}$ is enumerated by a linear recurrence sequence whose minimal polynomial is $(X-1)^2(X+1)^2)$, the following conjecture is formulated in [17].

Question 2.2.9 (see [17, Remark 4.15])**.** *Let A be enumerated by a linear recurrence sequence (r_n) with minimal polynomial P. Assume that (r_n) is non-degenerate. Is it true that if ± 1 is a repeated root of P, then* $\mathrm{Th}(\mathbf{Z}, +, 0, A)$ *unstable?*

It is plausible that this question has a positive answer, as the following examples support it:

1. we have already mentioned that $(2^n + n)$ gives an unstable expansion of $(\mathbf{Z}, +, 0)$ and it is apparent that $(2^n + n)$ is non-degenerate, since its minimal polynomial is $(X-2)(X-1)^2$;
2. let $P \in \mathbf{Z}[X]$ be non-constant and consider $A = P(\mathbf{N})$. A is then enumerated by the linear recurrence sequence $(P(n))$ which is non-degenerate, with minimal polynomial $(X-1)^{\deg(P)+1}$. Then by [16, Corollary 8.17 and Fact 8.18], $\mathrm{Th}(\mathbf{Z}, +, 0, A)$ is unstable.
3. let $P \in \mathbf{Z}[X]$ be non-constant, say $P(X) = a_0 + a_1 X + \cdots + a_d X^d$, and consider $A_{P,\pm} = \{P(n)(-1)^n \mid n \in \mathbf{N}\}$. As in the previous example, A is then enumerated by the sequence $(P(n)(-1)^n)$ which is non-degenerate, with minimal polynomial $(X+1)^{\deg(P)+1}$. Assume that $a_1 + a_2 + \cdots + a_d$ is odd. Then $\mathrm{Th}(\mathbf{Z}, +, 0, A)$ is also unstable. To see this, it is enough to show that $P(\mathbf{N})$ is definable in $(\mathbf{Z}, +, 0, A)$. We may assume that $A \subset \mathbf{N}$ and work in $(\mathbf{Z}, +, -, 0, 1, A, D_n \mid 1 < n \in \mathbf{N})$. By assumption, we have $P(2n) \equiv_2 a_0$ and $P(2n+1) \equiv_2 a_0 + 1$ for all $n \in \mathbf{N}$. Therefore, $P(\mathbf{N})$ is defined by

$$(x \in R \wedge D_2(x - a_0)) \vee (-x \in R \wedge D_2(x - a_0 - 1)).$$

We do not know whether we can remove the assumption that $a_1 + a_2 + \cdots + a_d$ is odd.

Question 2.2.10. *Is $(\mathbf{Z}, +, 0, A_{P,\pm})$ unstable when $a_1 + a_2 + \cdots + a_d$ is even, where $P(X) = a_0 + a_1 X + \cdots + a_d X^d$?*

Let us mention that Question 2.2.9 can be handled easily if instead of working in $(\mathbf{Z}, +, 0, A)$, we would work in $(\mathbf{Z}, +, 0, S, S^{-1}, A)$, where S is the successor function on A and S^{-1} its inverse. In that case, one can reduce Question 2.2.9 to the case of expansions by sets of the form $A_{P,\pm}$, by observing that the set defined by $\exists x \in R(y = S(x) - x)$ is enumerable by a linear recurrence relation with minimal polynomial of degree strictly less than the one of A. However, it is unclear to us if working in $(\mathbf{Z}, +, 0, S, S^{-1}, A)$ is harmless. More precisely, we ask the following question.

Question 2.2.11. *Is there a subset A of $\mathbf{Z}$ such that $(\mathbf{Z}, +, 0, S, S^{-1}, A)$ is unstable while $(\mathbf{Z}, +, 0, A)$ is (super-)stable?*

A natural generalization of the discussion on expansions by sets of the form $P(\mathbf{N})$, $P \in \mathbf{Z}[X]$, is this: what about the sets $P(\mathbf{Z})$, $P \in \mathbf{Z}[X]$? It is easy to see that if $P \in \mathbf{Z}[X]$ has degree 1, then $(\mathbf{Z}, +, 0, P(\mathbf{Z}))$ is superstable, since $P(\mathbf{Z})$ is definable in $(\mathbf{Z}, +, 0)$. For higher degree, the problem appears to be difficult and the best result we know of is due to H. Pasten and X. Vidaux under a strong algebro-geometric conjecture, namely the *uniform boundedness conjecture for rational points* (see [38, Conjecture 1.3]).

Theorem 2.2.12 ([38, Theorem 1.4])**.** *Assume uniform boundedness conjecture for rational points holds. Let $P \in \mathbf{Z}[X]$ of degree at least 2. Then the graph of multiplication is positive-existentially definable en $(\mathbf{Z}, +, 0, P(\mathbf{Z}))$.*

This result implies in particular that $(\mathbf{Z}, +, 0, P(\mathbf{Z}))$ is unstable and in fact as wild as possible in the sense of model theory.

We should point out, as a conclusion, that there are no known example of a strictly stable expansion of $(\mathbf{Z}, +, 0)$ by a unary predicate, that is a stable expansion that is not superstable.

We now move on to dependent expansions of $(\mathbf{Z}, +, 0)$ and $(\mathbf{Z}, +, 0, <)$. Apart from the stable ones, we do not know of an example of a dependent expansion of $(\mathbf{Z}, +, 0)$. However, to find such an expansion, G. Conant and C. Laskowski provide a useful tool: $(\mathbf{Z}, +, 0, A)$ is dependent if and only if A_{ind} is dependent (see [17, Theorem 2.9]). For $(\mathbf{Z}, +, 0, <)$, it was announced in [2, 3] that $(\mathbf{Z}, +, 0, <, A_q)$ and $(\mathbf{Z}, +, 0, <, \text{Fib})$ and the first complete proof of this fact appeared in [29]. This was based on a quantifier elimination result due to F. Point (see [41, Proposition 9]). This last paper focused on the decidability of expansions of $(\mathbf{Z}, +, 0, <)$ by a *sparse* set in the sense of Semenov [49]. Particular instances of sparse sets are A_q and Fib. In Chapter 5, we revisit the papers [41, 49] and provide a quantifier elimination result for expansions by sparse sets and reduce the dependency of those expansion to $\text{Th}(\mathscr{R})$ defined after Example 1.2.8. The ordered group of integers $(\mathbf{Z}, +, 0, <)$ satisfy a property that is an analogue of superstability but in the dependent setting, namely it is *strongly dependent* (see [50, Definition 4.23]). This property is defined using a notion of rank called the dp-rank (see [50, Definition 4.12]) and $(\mathbf{Z}, +, 0, <)$ has the minimal rank possible: its dp-rank is 1. In this case, we say that $(\mathbf{Z}, +, 0, <)$ is *dp-minimal*. It is interesting to see that being dp-minimal and strongly dependent is a strong property on $(\mathbf{Z}, +, 0, <)$. Indeed, it is shown in [18, Corollary 2.20] that $(\mathbf{Z}, +, 0, <)$ has no proper strongly dependent expansion. In particular, it has no proper dp-minimal expansion, a result that was proved in [2, Proposition 6.6]. This is similar to [37, Theorem 1], which states that $(\mathbf{Z}, +, 0)$ has no proper expansion of finite Lascar rank. As a consequence of [18, Corollary 2.20], the examples treated Chapter 5 are not strongly dependent.

Let us end with a short account of expansions of $(\mathbf{Z}, +, 0)$ by prime numbers. Let $\mathbf{P}$ be the set of prime numbers. The only result we know in this case are conditional to Dickson's conjecture.

Conjecture 1 (Dickson's conjecture)**.** *Let $k \geq 1$, a_i, b_i be integers such that $a_i \geq 1$ and $b_i \geq 0$ for all $i < k$. Let $f_i(x)$ be the polynomial $a_i x + b_i$. Assume that the following condition holds:*

$(\star_{\bar{f}})$ there does not exist any integer $n > 1$ dividing $\prod_{i<k} f_i(s)$ for all $s \in \mathbf{N}$.
Then, there exist infinitely many $m \in \mathbf{N}$ such that $f_i(m)$ is prime for all $i < k$.

The main result concerning the tameness of expansions by prime numbers is due to I. Kaplan and S. Shelah.

Theorem 2.2.13 ([28, Theorem 1.2]). *Let T be the theory of $(\mathbf{Z}, +, 0, \mathbf{P} \cup -\mathbf{P})$. Then T is independent and, if Dickson's conjecture is true, T is supersimple.*

For an introduction to simple theories, we refer to [53, Chapter 7].

Notice that the structure $(\mathbf{Z}, +, 0, \mathbf{P})$ has the order property[1] (and is in particular unstable), which is why the expansion $(\mathbf{Z}, +, 0, \mathbf{P} \cup -\mathbf{P})$ is considered: simple theories lack the order property.

In relation to [28, Theorem 1.2], under the assumption that Dickson's conjecture is true, P. T. Bateman, C. G. Jockusch and A. R. Woods showed that the theory of $\mathscr{Z}_{<,\mathbf{P}} = (\mathbf{Z}, +, 0, <, \mathbf{P})$ is undecidable and, in fact, that the multiplication is definable (see [4, Theorem 1]). This result was slightly improved in [7] by M. Boffa, who obtained the same result for $\mathscr{Z}_{<,\mathbf{P}_{m,r}}$, where, for coprime natural numbers $r < m$, $\mathbf{P}_{m,r}$ is the set $\{p \mid p \equiv_m r \text{ and } p \in \mathbf{P}\}$. In particular $(\mathbf{Z}, +, 0, <, \mathbf{P})$ and $(\mathbf{Z}, +, 0, <, \mathbf{P}_{m,r})$ have independent theories.

In the spirit of M. Boffa's improvement of [4, Theorem 1], it would be interesting to know if the statement of [28, Theorem 1.2] holds for $(\mathbf{Z}, +, 0, <, \mathbf{P}_{m,r} \cup -\mathbf{P}_{m,r})$, k coprime with n.

Question 2.2.14. *Let $r < m$ be coprime. Is* $\mathrm{Th}(\mathbf{Z}, +, 0, <, \mathbf{P}_{m,r} \cup -\mathbf{P}_{m,r})$ *superstable and independent?*

[1] For instance, a Theorem of Tao (see [52]) states that every natural number greater than 1 is the sum of at most five prime numbers.

CHAPTER

3

Expansion of $(\mathbf{Z}, +, 0)$ by a regular set: superstability

In this chapter we identify a class of sets of natural numbers, which we call *regular*, that provide superstable expansions of $\mathscr{Z}$. These sets R are enumerated by a sequence (r_n) that grows fast, in the sense that they have a Kepler limit in $\mathbf{R}_{\infty}^{>1}$. An extra condition is required when the Kepler limit θ is algebraic: we impose that the enumeration is a linear recurrence sequence whose minimal polynomial is the minimal polynomial of θ. We also show that those expansions have Lascar rank ω, see Theorem 3.5.1.

The proof of Theorem 3.5.1 follows the same strategy used by D. Palacín and R. Sklinos in [37], that is we use the results of E. Casanovas and M. Ziegler on stable pairs.

This chapter is organized as follows. In Section 3.1, we define precisely what a regular set is and we provide examples and counter-examples of such sets.

In Section 3.2, given a regular set R enumerated by (r_n), we introduce a family of functions, called *operators*, that are intended to detect the recurrence relations satisfied by (r_n). More precisely, an operator is a function $\mathsf{f} : \mathbf{N} \to \mathbf{Z} : n \mapsto a_0 r_n + a_1 r_{n+1} + \cdots + a_d r_{n+d}$, where $\bar{a} \in \mathbf{Z}^{d+1}$. The main property of these functions is that sets of the form $\{n \mid \mathsf{f}(n) = z\}$, $z \in \mathbf{Z}$, are always finite if $z \neq 0$ and in case $z = 0$, then they are either finite or $\mathbf{N}$. Furthermore, this is detected by the Kepler limit of (r_n).

Then in Section 3.3, we begin the analysis of the trace of equations on R. There, given operators $\mathsf{f}_1, \ldots, \mathsf{f}_k$ and $z \in \mathbf{Z}$, we consider the set

$$\{\bar{n} \in \mathbf{N}^k \mid \mathsf{f}_1(n_1) + \cdots + \mathsf{f}_k(n_k) = z\}.$$

The analysis of these sets is reduced to the set X of *non-degenerate solutions* of the equation $\mathsf{f}_1(n_1) + \cdots + \mathsf{f}_k(n_k) = z$. These solutions satisfy the condition

$$\sum_{i \in I} \mathsf{f}_i(n_i) \neq 0 \text{ for all } J \subsetneq [k].$$

Proposition 3.3.1 imply that there exists a constant m depending only on $\mathsf{f}_1, \ldots, \mathsf{f}_k$ and z such that if $\bar{n} \in X$ then $\max\{|n_i - n_j| \mid i, j \in [k]\} \leq m$. All this work allows us to to show that R^0_{ind} is superstable by interpreting it in $\mathscr{N} = (\mathbf{N}, S, S^{-1}, 0)$, for which $\mathscr{N}^1$ is superstable.

In Section 3.4 we show that for R regular, the pair $\mathscr{Z}_R$ is bounded. This is done by showing that given operators $\mathsf{f}_1, \ldots, \mathsf{f}_k$ the set

$$\{\mathsf{f}_1(n_1) + \cdots + \mathsf{f}_k(n_k) \mid \bar{n} \in \mathbf{N}^k\} \cap \mathbf{N}$$

is not *piecewise syndetic*: it does not contain arbitrarily long sequences with bounded gaps. This allows us to appeal to Theorem 2.1.9, showing that $\mathscr{Z}$ dnfcp over R. The main theorem of this chapter is then proved in Section 3.5.

We end this chapter with three other sections. In Section 3.6, we compare the notions of regular and geometrically sparse sets. In Section 3.7, we point out the analogy between Proposition 3.3.1 and the Mann property in fields and give a quick proof of a special case of a result of G. Conant [15, Theorem 3.1] that the pair $\mathscr{Z}_A$ is superstable when A is an infinite subset of a finitely generated multiplicative submonoid of $\mathbf{N}$. Finally, in Section 3.8, we prove that $(\mathbf{Q}, +, 0, R)$ and $(\mathbf{R}, +, 0, R)$ are ω-stable for any regular set $R \subset \mathbf{N}$, using the work done on the trace of equations on R.

3.1 Regular sets

In this section, we define the main objects studied in this chapter. These objects, called *regular sets* and *regular sequences*, highlight the common behavior in the superstable expansions of $\mathscr{Z}$ studied by D. Palacín and R. Sklinos in [37]. Typical examples of the expansions studied in [37] are $(\mathbf{Z}, +, 0, A_q)$ and $(\mathbf{Z}, +, 0, \{n! \mid n \in \mathbf{N}\})$. One common crucial property of these expansions is that the sequence (2^n) and $(n!)$ have a Kepler limit in $\mathbf{R}^{\geq 0}_{\infty}$. This common property is the core of our definition of regular sets: these sets can be enumerated by an increasing sequence that have a Kepler limit in $\mathbf{R}^{\geq 0}_{\infty}$. In other words regular sets must have at least exponential growth. In our definition of regular set, another condition is added when the Kepler limit of an increasing enumeration is algebraic over $\mathbf{Q}$: we require that the enumeration satisfy a recurrence relation whose minimal polynomial is the minimal polynomial of its Kepler limit (note that (2^n) satisfy this condition). This definition is designed so as to allow us to control the set of solutions of equations satisfied by elements of regular sets by looking where those solutions appear in an increasing enumeration: the exponential growth does not allow solutions where the components are too far away from each other in an increasing enumeration, unless some of those components satisfy another equation with fewer variables. Furthermore, we show that the exponential growth implies that the set of solutions of an equation satisfied by elements of a regular set are determined by operators. These functions behave nicely: they are either constantly

0 or ultimately injective. This explains why we needed an extra condition when the Kepler limit is algebraic over $\mathbf{Q}$: in this case the constant operators on a regular set correspond to the polynomials in the ideal generated by the minimal polynomial of its Kepler limit. In the case where a regular set has a Kepler limit not algebraic over $\mathbf{Q}$ or infinite, there is only one constant operator.

Definition 3.1.1. Let (r_n) be a sequence of natural numbers and $R \subset \mathbf{N}$.

1. We say that (r_n) is *regular* if $\lim_{n\to\infty} r_{n+1}/r_n = \theta \in \mathbf{R}_\infty^{\geq 1}$ and, if θ is algebraic over $\mathbf{Q}$, (r_n) satisfies a linear recurrence relation whose minimal polynomial is the minimal polynomial of θ.
2. We say that R is *regular* if it can be enumerated by a regular sequence.

We observe that a regular sequence must be ultimately strictly increasing. Indeed, let (r_n) be a regular sequence, as witnessed by $\theta \in \mathbf{R}_\infty^{\geq 1}$. With the convention that $\infty^{-1} = 0$, we have that $r_n/r_{n+1} \to \theta^{-1}$. As a consequence, for all $n \in \mathbf{N}$ sufficiently large, we have $|r_n - \theta^{-1} r_{n+1}| < (1-\theta^{-1})r_{n+1}$. Thus, $r_n < (1-\theta^{-1})r_{n+1} + \theta^{-1}r_{n+1} = r_{n+1}$. In particular, a regular set is automatically infinite. Also, if (r_n) is regular, as witnessed by $\theta \in \mathbf{R}_\infty^{\geq 1}$, then for all $k \in \mathbf{N}^{>1}$,

$$\lim_{n\to\infty} \frac{r_{n+k}}{r_n} = \begin{cases} \infty & \text{if } \theta = \infty \\ \theta^k & \text{otherwise.} \end{cases}$$

These observations will be used freely in the rest of this text.

Let us now give examples and counter-examples of regular sequences. We first consider the non-algebraic case.

Example 3.1.2.

1. The sequence $(n!)$ is regular with $\theta = \infty$.
2. Let $\theta \in \mathbf{R}^{>1}$ and consider the sequences $(\lfloor \theta^n \rfloor)$ and $(\lfloor \theta^n \rfloor + n)$. We have

$$\lim_{n\to\infty} \frac{\lfloor \theta^n \rfloor}{\theta^n} = \lim_{n\to\infty} \frac{\lfloor \theta^n \rfloor + n}{\theta^n} = 1.$$

Thus

$$\lim_{n\to\infty} \frac{\lfloor \theta^{n+1} \rfloor}{\lfloor \theta^n \rfloor} = \lim_{n\to\infty} \frac{\lfloor \theta^{n+1} \rfloor + n + 1}{\lfloor \theta^n \rfloor + n} = \theta.$$

Thus, when θ is transcendental, $(\lfloor \theta^n \rfloor)$ and $(\lfloor \theta^n \rfloor + n)$ are regular sequences.

We note that when (r_n) is a regular sequence with non-algebraic Kepler limit, then $(r_n + o(r_n))$ is also a regular sequence.

We now look at the algebraic case.

Example 3.1.3. Let $k \in \mathbf{N}^{>1}$. Then sequence (k^n) is regular. Indeed, it satisfies the recurrence relation $r_{n+1} = kr_n$, whose minimal polynomial $X-k$ is the minimal polynomial of k. Likewise, the Fibonacci sequence is regular, with minimal polynomial X^2-X-1 and $\theta = (1+\sqrt{5})/2$. Finally, the sequence $r_{n+2} = 5r_{n+1} + 7r_n$ with $r_1 = 1$ and $r_0 = 0$ is also regular with minimal polynomial X^2-5X-7 and $\theta = (5+\sqrt{53})/2$ and will be discussed further in Section 3.6.

In this case, we cannot say in general that if (r_n) is a regular sequence with algebraic Kepler limit, then $(r_n + o(r_n))$ is regular. For instance, the sequence (3^n+2^n) has Kepler limit 3, but its minimal polynomial is $(X-3)(X-2)$. Hence (3^n+2^n) is not regular. However, G. Conant and C. Laskowsky recently showed in [17, Theorem 4.9] that $(\mathbf{Z},+,0,\{3^n+2^n \mid n \in \mathbf{N}\})$ is superstable.

Many examples of unstable expansions of $\mathscr{Z}$ are by sets with Kepler limit 1.

Counter-example 3.1.4.

1. The sequence (n) is not regular as its Kepler limit is 1. The same holds for arithmetic sequences, *i.e.* sequences of the form $(a+nb)$, where $a,b \in \mathbf{N}$ and $b \neq 0$. Note that expansions of $\mathscr{Z}$ by such sequences are unstable[1] ($\mathbf{N}$ is defined by the formula $\exists y \in R(a+bx=y)$) and satisfy the linear recurrence $r_{n+2} = 2r_{n+1} - r_n$.
2. The sequence (p_n) of prime numbers is not regular. This is a consequence of the Prime Number Theorem (see [27, Theorem 3.4.3 and Proposition 3.5.3]): $p_{n+1}/p_n \to 1$.

3.2 Operators

In this section, we fix a regular sequence (r_n) and we let $\theta = \lim_{n\to\infty} r_{n+1}/r_n$. Our goal here is to determine the recurrence relations satisfied by (r_n). To this end, we will need the following definition.

Definition 3.2.1. Let $Q \in \mathbf{Z}[X]$. Assume $Q(X) = a_0 + a_1X + \cdots + a_dX^d$, where $\bar{a} \in \mathbf{Z}^{d+1}$. The *operator* associated to Q, denoted f_Q or simply f, is the function $\mathsf{f}: \mathbf{N} \to \mathbf{Z}: n \mapsto a_0r_n + a_1r_{n+1} + \cdots + a_dr_{n+d}$.

Given an operator f and $z \in \mathbf{Z}$, we let $\mathsf{S}_{\mathsf{f},z}$ be the set

$$\mathsf{f}^{-1}(z) = \{n \in \mathbf{N} \mid \mathsf{f}(n) = z\}.$$

Let us show that the sets $\mathsf{S}_{\mathsf{f},z}$ are quite simple: they are either finite or $\mathbf{N}$.

[1] This is also true for any sequence (r_n) such that there exists $k \in \mathbf{N}$ such that for all $n \in \mathbf{N}$, $|r_{n+1} - r_n| \leq k$.

Proposition 3.2.2. *Let $Q \in \mathbf{Z}[X] \setminus \{0\}$ and $z \in \mathbf{Z}$. Let $\mathsf{f} = \mathsf{f}_Q$. Then*

1. *if $z = 0$, then*
   a) *if $\theta = \infty$ or $Q(\theta) \neq 0$, then $\mathsf{S}_{\mathsf{f},z}$ is finite;*
   b) *$\mathsf{S}_{\mathsf{f},z} = \mathbf{N}$ if and only if $Q(\theta) = 0$;*
2. *if $z \neq 0$ then $\mathsf{S}_{\mathsf{f},z}$ is finite.*

Proof. Assume $Q(X) = a_0 + a_1 X + \cdots + a_d X^d$, $\bar{a} \in \mathbf{Z}^{d+1}$, $a_d \neq 0$. Then

$$\lim_{n\to\infty} \frac{\mathsf{f}(n)}{r_{n+d}} = \begin{cases} a_d & \text{if } \theta = \infty, \\ \theta^{-d} Q(\theta) & \text{otherwise.} \end{cases}$$

Also $\lim\limits_{n\to\infty} z/r_{n+d} = 0$. So, if $\theta = \infty$ or $Q(\theta) \neq 0$, we get that $\mathsf{S}_{\mathsf{f},z}$ is finite. Now, if $Q(\theta) = 0$, then (r_n) follows a linear recurrence relation whose minimal polynomial is the minimal polynomial P_θ of θ. Thus, Q is divisible by P_θ. By Proposition 1.3.7, this implies that $\mathsf{S}_{\mathsf{f},0} = \mathbf{N}$. □

In the remainder of this section, we indicate how the data from Proposition 3.2.2 can be estimated from θ. This will be needed when we address the decidability of expansions of $\mathscr{Z}$ by regular sets.

Given $z \in \mathbf{Z}$ and $Q \in \mathbf{Z}[X] \setminus \{0\}$, we want to estimate the size of $\mathsf{S}_{\mathsf{f}_Q,z}$ when it is finite. Let $\delta : \mathbf{N} \to \mathbf{N}$ be a modulus of convergence for (r_n/r_{n+1}), that is δ satisfies $\forall n \in \mathbf{N}\, \forall m \geq \delta(n)\ |r_m/r_{m+1} - \theta^{-1}| \leq 1/2^n$, where $\theta^{-1} = 0$ if $\theta = \infty$, and let $n_0 \in \mathbf{N}$ be such that $\theta^{-1} + 1/2^{n_0} < 1$. For each $k \in \mathbf{N}^{>0}$, we recursively define a function $\delta_k : \mathbf{N} \to \mathbf{N}$ such that $\forall n \in \mathbf{N}\, \forall m \geq \delta_k(n)\ |r_m/r_{m+k} - \theta^{-k}| \leq 1/2^n$. We let $\delta_1 = \delta$ and assuming δ_{k-1} is constructed, we define δ_k by $\delta_k(n) = \max\{n_0, \delta(n+1), \delta_{k-1}(n+1)\}$. Let us check that δ_k is in fact a modulus of convergence for (r_n/r_{n+k}). Let $m \geq \delta_k(n)$. Then

$$\begin{aligned} |r_m/r_{m+k} - \theta^{-k}| &= |(r_m/r_{m+1} - \theta^{-1}) r_{m+1}/r_{m+k} + \theta^{-1}(r_{m+1}/r_{m+k} - \theta^{-k+1})| \\ &\leq |r_m/r_{m+1} - \theta^{-1}| + |r_{m+1}/r_{m+k} - \theta^{-k+1}| \\ &\leq 1/2^{n+1} + 1/2^{n+1} = 1/2^n. \end{aligned}$$

In a similar way, we can define, for all $Q \in \mathbf{Z}[X] \setminus \{0\}$, a modulus of convergence for $(\mathsf{f}_Q(n)/r_{n+d})$, where $d = \deg(Q)$. Indeed, if $Q(X) = a_0 + \cdots + a_d X^d$, then the function δ_Q defined by $\delta_Q(n) = \max\{\delta_d(n+\ell), \delta_{d-1}(n+\ell), \ldots, \delta_1(n+\ell)\}$, where $\ell \in \mathbf{N}$ is such that $|a_0| + |a_1| + \cdots + |a_d| \leq 2^\ell$, is a modulus of convergence for $(\mathsf{f}_Q(n)/r_{n+d})$. Let us check this when $\theta \in \mathbf{R}^{>1}$. Let $m \geq \delta_Q(n)$. Then

$$\begin{aligned}\left|\frac{\mathsf{f}_Q(m)}{r_{m+d}} - \theta^{-d}Q(\theta)\right| &= \left|\sum_{i=0}^{d} a_i \frac{r_{m+i}}{r_{m+d}} - \sum_{i=0}^{d} a_i\theta^{i-d}\right| \\ &\leq \sum_{i=0}^{d} |a_i| \left|\frac{r_{m+i}}{r_{m+d}} - \theta^{i-d}\right| \\ &\leq \sum_{i=0}^{d} |a_i| 2^{-(n+\ell)} \\ &= 2^{-(n+\ell)} \sum_{i=0}^{d} |a_i| \\ &\leq 2^{-n}.\end{aligned}$$

Now let $z \in \mathbf{Z}$ and $Q \in \mathbf{Z}[X]\backslash\{0\}$. Set $\mathsf{f} = \mathsf{f}_Q$. We want to estimate $|\mathsf{S}_{\mathsf{f},z}|$ using the moduli defined above, assuming either $\theta = \infty$, $Q(\theta) \neq 0$ or $z \neq 0$. We concentrate on the case $\theta \in \mathbf{R}^{>1}$. Let $u = Q(\theta)$ and $\|u\| = |a_0| + \cdots + |a_d|$. Let $n_1 \in \mathbf{N}$ be such that $\|u\|/2^{n_1} < |u|$. Then for all $m \geq \delta_Q(n_1)$, we have, since $|a_i r_{m+i} - a_i\theta^i r_m| < |a_i| r_m/2^{n_1}$ whenever $a_i \neq 0$,

$$0 < r_m(|u| - \|u\|/2^{n_1}) < |\mathsf{f}_Q(m)|.$$

Let $n_2 \in \mathbf{N}$ such that $r_{n+1} > r_n$ for all $n \geq n_2$. Then if $m \geq \max\{\delta_Q(n_1), n_1, n_2 + \lfloor |z|/(|u| - \|u\|/2^{n_1}) \rfloor\}$, we have $|z| < |\mathsf{f}_Q(m)|$, so that $|\mathsf{S}_{\mathsf{f},z}| < \max\{\delta_Q(n_1), n_1, n_2 + \lfloor |z|/(|u| - \|u\|/2^{n_1}) \rfloor\}$.

3.3 Equations

Let (r_n) be a regular sequence and θ be its Kepler limit. Let $R = \{r_n \mid n \in \mathbf{N}\}$. In order to show that R^0_{ind} is superstable, we have to understand the trace on R of equations with coefficients in $\mathbf{Z}$. In what follows, we show that the trace of an equation is either finite or determined by a finite number of operators on R.

Let $Q_1, \ldots, Q_s \in \mathbf{Z}[X]$ be operators and let $\mathsf{f}_i = \mathsf{f}_{Q_i}$ for all $i \in [s]$. Let $z \in \mathbf{Z}$. We consider the equation $\mathsf{f}_1(x_1) + \cdots + \mathsf{f}_s(x_s) = z$. We call a tuple $\bar{n} \in \mathbf{N}^s$ a *non-degenerate solution* of $\mathsf{f}_1(x_1) + \cdots + \mathsf{f}_s(x_s) = z$ when the following conditions hold:

1. $\mathsf{f}_1(n_1) + \cdots + \mathsf{f}_s(n_s) = z$;
2. for all $I \subsetneq [s]$ $\sum_{i \in I} \mathsf{f}_i(n_i) \neq 0$.

We now explain how to decompose $\mathsf{S}_{\bar{\mathsf{f}},z} = \{\bar{n} \in \mathbf{N}^s \mid \mathsf{f}_1(n_1) + \cdots + \mathsf{f}_s(n_s) = z\}$ into sets of non-degenerate solutions.

Let $\bar{I} = (I_1, \ldots, I_k) \in \mathsf{Part}([n])$. To this partition we associate the following system of equations:

$$\begin{cases} \sum_{i\in I_1} \mathsf{f}_i(n_i) = z, \\ \sum_{i\in I_2} \mathsf{f}_i(n_i) = 0, \\ \vdots \\ \sum_{i\in I_k} \mathsf{f}_i(n_i) = 0. \end{cases} \tag{3.1}$$

Let

$$\mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}},z,\bar{I}} = \Big\{ \bar{n} \in \mathbf{N}^s \Big| \bar{n}_{I_1} \text{ is a non-degenerate solution of } \sum_{i\in I_1} \mathsf{f}_i(n_i) = z \text{ and}$$
$$\text{for all } j \in [k]^{>1}\ \bar{n}_{I_j} \text{ is a non-degenerate solution of } \sum_{i\in I_j} \mathsf{f}_i(n_i) = 0 \Big\}.$$

When $\bar{I} = ([s])$, we use $\mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}},z}$ instead of $\mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}},z,\bar{I}}$. In this setting, we decompose $\mathsf{S}_{\bar{\mathsf{f}},z}$ as

$$\mathsf{S}_{\bar{\mathsf{f}},z} = \bigcup_{\bar{I}\in\mathsf{Part}([s])} \mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}},z,\bar{I}}. \tag{3.2}$$

This decomposition will prove to be quite useful as the set of non-degenerate solutions of $\mathsf{f}_1(x_1) + \cdots + \mathsf{f}_s(x_s) = z$ is easily understood. For instance, Proposition 3.3.1 implies that for some constant m depending only on $\bar{\mathsf{f}}$ and z, if $\bar{n}$ is a non-degenerate solution, then $\max\{|n_i - n_j| \mid i, j \in [s]\} \leq m$.

Proposition 3.3.1. *Let $\mathsf{f}_1, \ldots, \mathsf{f}_s$ be operators and $z \in \mathbf{Z}$. Then, there exist $k \in \mathbf{N}$ and $\bar{m}_1, \ldots, \bar{m}_k \in \mathbf{Z}^s$ such that for all $\bar{\ell} \in \mathbf{N}^s$, if $\bar{\ell} \in \mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}},z}$ then for some $i \in [k]$, $\ell_j = \ell_1 + m_{ij}$ for all $j \in [s]$.*

Proof. Assume that $\mathsf{f}_j = \mathsf{f}_{Q_j}$ where $Q_j(X) = \sum_{i=0}^{d_j} a_{ji}X^i$ and $a_{jd_j} \neq 0$. Suppose, towards a contradiction, that the proposition is false: for all $k \in \mathbf{N}$ and $\bar{m}_1, \ldots, \bar{m}_k \in \mathbf{Z}^s$, there exists $\bar{\ell} \in \mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}},z}$ such that for all $i \in [k]$, $\ell_j \neq \ell_1 + m_{ij}$ for some $j \in [s]$. From this, we construct two sequences $(\bar{\ell}_i) \subset \mathbf{N}^s$ and $(\bar{m}_i) \subset \mathbf{Z}^s$ that will help us reach a contradiction.

Start with any $\bar{\ell}_1 \in \mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}},z}$ and define $\bar{m}_1$ as $m_{1j} = \ell_{1j} - \ell_{11}$ for all $j \in [s]$. Assuming $\bar{\ell}_i$ and $\bar{m}_i$ are constructed, we let $\bar{\ell}_{i+1}$ be a non-degenerate solution obtained from our assumption that the proposition is false with $k = i$ and $\bar{m}_1, \ldots, \bar{m}_i$. We define $\bar{m}_{i+1}$ as $m_{(i+1)j} = \ell_{(i+1)j} - \ell_{(i+1)1}$ for all $j \in [s]$. We want to perform calculations on these sequences using the fact that (r_n) has a Kepler limit as in the proof of Proposition 3.2.2. However, the sequences $(\bar{\ell}_i)$ and $(\bar{m}_i)$ are not nice enough to perform the kind of calculations done in the proof of Proposition 3.2.2, so that we first need to rearrange them slightly and pass to subsequences to obtain nice properties that will allow us to estimate $\sum_{j\in[s]} \mathsf{f}_j(\ell_{i1} + m_{ij}) = \sum_{j\in[s]} \mathsf{f}_j(\ell_{ij})$.

We first reorder the tuples $\bar{m}_i$. We may assume, up to a permutation of $\bar{\mathsf{f}}$ and passing to a subsequence using the pigeonhole principle, that

1. for all $i \in \mathbf{N}$, $m_{ij} \leq m_{i(j+1)}$ for all $j < s$.

We know by construction that each tuple $\bar{m}_i$ has a coordinate whose value is 0. By the pigeonhole principle, there exists $j^* \in [s]$ such that $m_{ij^*} = 0$ for infinitely many $i \in \mathbf{N}$. Thus, up to passing to a subsequence, we may assume that

2. there is $j^* \in [s]$ such that for all $i \in \mathbf{N}$, $m_{ij^*} = 0$;

We want the sequence (m_{is}) to be strictly increasing. By passing to a subsequence, me may assume that

3. for all $i \in \mathbf{N}$, $m_{is} < m_{(i+1)s}$. In particular, $m_{is} \to \infty$.

We now decompose the tuples $\bar{m}_i$, $i \in \mathbf{N}$, in two parts according to whether the differences $m_{is} - m_{ij} = \ell_{is} - \ell_{ij}$ are bounded. Let $J \subset [s]$ be of maximal size such that for all $j \in J$, $\max\{m_{is} - m_{ij} \mid i \in \mathbf{N}\} < \infty$. Notice that $s \in J$. Also, by 2 and 3, $j^* \notin J$.

For all $j \in J$, we have that the sequence $(m_{is} - m_{ij})$ is bounded, hence take finitely many values. So applying successively the pigeonhole principle for each $j \in J$, we may assume that

4. for all $j \in J$, there exists $k_j \in \mathbf{N}$ such that $m_{is} - m_{ij} = k_j$ for all $i \in \mathbf{N}$.

For each $j \notin J$, we have that the sequence $(m_{is} - m_{ij})$ is unbounded, hence has a subsequence that converges to ∞. Thus applying successively the pigeonhole principle for each $j \notin J$, we may assume that

5. for all $j \notin J$, $m_{is} - m_{ij} \to \infty$.

We are now ready to estimate $\sum_{j\in[s]} \mathsf{f}_j(\ell_{i1} + m_{ij}) = \sum_{j\in[s]} \mathsf{f}_j(\ell_{ij})$. More precisely, our goal is to calculate the limit

$$\lim_{i\to\infty} \sum_{j=1}^{s} \frac{\mathsf{f}_j(\ell_{i1} + m_{ij})}{r_{\ell_{i1}+m_{ij_0}+d}},$$

where $j_0 = \min J$.

For all $i \in \mathbf{N}$ and $j \in J$, rewrite ℓ_{ij} as $\ell_{ij_0} + (m_{ij} - m_{ij_0}) = \ell_{ij_0} + (k_{j_0} - k_j)$ (note that by 1, $k_{j_0} - k_j \geq 0$). Set $Q'_j(X) = \sum_{n=1}^{d_j} a_{jn} X^{k_{j_0}-k_j+n}$ and $\mathsf{f}'_j = \mathsf{f}_{Q'_j}$. We have, for all $j \in J$,

$$\begin{aligned}\mathsf{f}_j(\ell_{ij}) &= \mathsf{f}_j(\ell_{ij_0} + (m_{ij} - m_{ij_0}))\\ &= \mathsf{f}_j(\ell_{ij_0} + (k_{j_0} - k_j))\\ &= \mathsf{f}'_j(\ell_{ij_0}).\end{aligned}$$

Define $Q(X) = \sum_{j \in J} Q'_j(X)$, let d be the degree of Q and a_d be the coefficient of X^d in Q. For all $i \in \mathbf{N}$,

$$\begin{aligned}\sum_{j \in J} \mathsf{f}_j(\ell_{ij}) &= \sum_{j \in J} \mathsf{f}'_j(\ell_{ij_0}) \\ &= \mathsf{f}_Q(\ell_{ij_0}).\end{aligned}$$

Before we move on, let us check that $Q \neq 0$. If it were not the case, then, we would have

$$\sum_{j \in J} \mathsf{f}_j(\ell_{ij}) = 0 \text{ for all } i \in \mathbf{N}.$$

But, since J is a non-empty proper subset of $[s]$ (recall that $s \in J$ and $j^* \notin J$), this implies that $\bar{\ell}_i$ is a non-degenerate solution of $\mathsf{f}_1(x_1) + \cdots + \mathsf{f}_s(x_s) = z$, in contradiction with our assumption that $\bar{\ell}_i$ is non-degenerate for all $i \in \mathbf{N}$. Hence $Q \neq 0$ and in particular $a_d \neq 0$.

Since, for all $n \in \mathbf{N}$, $\lim\limits_{k \to \infty} r_n / r_{n+k} = 0$, for all $j \notin J$,

$$u_j = \lim_{i \to \infty} \frac{\mathsf{f}_j(\ell_{i1} + m_{ij})}{r_{\ell_{i1} + m_{ij_0} + d}} = \lim_{i \to \infty} \sum_{n=0}^{d_j} \frac{a_{jn} r_{\ell_{i1} + m_{ij} + n}}{r_{\ell_{i1} + m_{ij_0} + d}} = 0.$$

The last equality comes from the fact that by 4 and 5, for all $j \notin J$, $m_{ij_0} - m_{ij} = m_{is} - m_{ij} - k_{j_0} \to \infty$.

We perform a similar calculation for all $j \in J$. Recall that for all $k \in \mathbf{N}^{>0}$,

$$\lim_{n \to \infty} r_n / r_{n+k} = \begin{cases} \theta^{-k} & \text{if } \theta \in \mathbf{R}^{>1} \\ 0 & \text{if } \theta = \infty. \end{cases}$$

So we have that

$$\begin{aligned} u_J &= \lim_{i \to \infty} \frac{\sum_{j \in J} \mathsf{f}_j(\ell_{i1} + m_{ij})}{r_{\ell_{i1} + m_{ij_0} + d}} \\ &= \lim_{i \to \infty} \frac{\mathsf{f}_Q(\ell_{ij_0})}{r_{\ell_{ij_0} + d}} \\ &= \begin{cases} \theta^{-d} Q(\theta) & \text{if } \theta \in \mathbf{R}^{>1} \\ a_d & \text{if } \theta = \infty. \end{cases} \end{aligned}$$

Thus,

$$0 = \lim_{i \to \infty} \frac{z}{r_{\ell_{i1} + m_{ij_0} + d}} = \lim_{i \to \infty} \sum_{j=1}^{s} \frac{\mathsf{f}_j(\ell_{i1} + m_{ij})}{r_{\ell_{i1} + m_{ij_0} + d}} = u_J + \sum_{j \notin J} u_j = u_J,$$

where that last equality comes from the fact that, by 3 and 4, the sequence $(r_{\ell_{i1}+m_{ij_0}+d})$ is not bounded.

Since $u_J = 0$ and $a_d \neq 0$, we must have $\theta \in \mathbf{R}^{>1}$. In particular, $u_J = \theta^d Q(\theta)$. So $Q(\theta) = 0$. Since R is regular and $Q \neq 0$, R satisfies a linear recurrence relation and P_R divides Q. So by Proposition 3.2.2 $\mathsf{S}_{\mathsf{f}_Q,0} = \mathbf{N}$, in contradiction with the assumption that $\bar{\ell}_i$ is non-degenerate for all $i \in \mathbf{N}$ and the fact that J is a proper non-empty subset of $[s]$. □

In preparation of Corollary 3.3.6, where we establish the superstability of R^0_{ind}, the induced structure on R of equations with coefficients in $\mathbf{Z}$, we make a few comments on Proposition 3.3.1. Let us first state precisely what we meant when we wrote that the set of solutions of equations is either finite or determined by operators. Per equation 3.2, we may focus on the set of non-degenerate solutions.

For an operator f_Q, $Q(X) = \sum_{i=0}^d a_i X^i$, let

$$\mathsf{S}^\circ_{\mathsf{f}_Q,z} = \{n \in \mathsf{S}_{\mathsf{f}_Q,z} \mid \text{for all } I \subsetneq [d] \sum_{i\in I} a_i r_{n+i} \neq 0\}.$$

For an s-tuple $\bar{\mathsf{f}}$ of operators and $\bar{n} \in \mathbf{N}^s$, we let $\mathsf{f}_{\bar{n}}(\ell) = \sum_{j=1}^s \mathsf{f}_j(\ell + n_j)$.

Remark 3.3.2. Let $\mathsf{f}_1, \ldots, \mathsf{f}_s$ be operators and $z \in \mathbf{Z}$. By Proposition 3.3.1, there exist k and $\bar{m}_1, \ldots, \bar{m}_k \in \mathbf{Z}^s$ such that, letting $m_i = \min\{m_{i1}, \ldots, m_{is}\}$ and $\bar{n}_i = \bar{m}_i - m_i$:

$\bar{\ell} \in \mathsf{S}^{\text{nd}}_{\bar{\mathsf{f}},z}$ if and only if for some $i \in [k]$, $\ell_1 + m_i \in \mathsf{S}^\circ_{\mathsf{f}_{\bar{n}_i},z}$ and $\ell_j = \ell_1 + m_{ij}$ for all $j \in [s]$.

The operators $\mathsf{f}_{\bar{n}_i}$ of Remark 3.3.2 are the ones that determine the set of non-degenerate solutions.

Corollary 3.3.3. *Let $\mathsf{f}_1, \ldots, \mathsf{f}_s$ be operators and $z \in \mathbf{Z}$. Let $k \in \mathbf{N}$ and $\bar{m}_1, \ldots, \bar{m}_k \in \mathbf{Z}^s$ be given by Proposition 3.3.1. Then $\mathsf{S}^{\text{nd}}_{\bar{\mathsf{f}},z}$ is infinite if and only if $z = 0$ and $\mathsf{S}^\circ_{\mathsf{f}_{\bar{m}_i},0}$ is infinite for some $i \in [k]$.*

Proof. This follows from Proposition 3.2.2 and Proposition 3.3.1. □

The following corollary states that operators are ultimately injective functions, unless $\mathsf{S}_{\mathsf{f},0}$ is infinite.

Corollary 3.3.4. *Let $Q \in \mathbf{Z}[X]$. Then exactly one of the following holds:*

- $\mathsf{S}_{\mathsf{f}_Q,0} = \mathbf{N}$;
- $\mathsf{S}_{(\mathsf{f}_Q,-\mathsf{f}_Q),0} \backslash \{(n,n) \mid n \in \mathbf{N}\}$ *is finite.*

Proof. Assume $\mathsf{S}_{\mathsf{f}_Q,0}$ is finite. Let us then show that $\mathsf{S}_{(\mathsf{f}_Q,-\mathsf{f}_Q),0} \backslash \{(n,n) \mid n \in \mathbf{N}\}$ is finite. Since $\mathsf{S}_{\mathsf{f}_Q,0}$ is finite, $\mathsf{S}_{(\mathsf{f}_Q,-\mathsf{f}_Q),0} \backslash \mathsf{S}^{\text{nd}}_{(\mathsf{f}_Q,-\mathsf{f}_Q),0}$ is finite, so we only need to show that

$\mathsf{S}^{\mathrm{nd}}_{(\mathsf{f}_Q,-\mathsf{f}_Q),0} \setminus \{(n,n) \mid n \in \mathbf{N}\}$ is finite. By Proposition 3.3.1, this amounts to show that for all $k \in \mathbf{N}^{>0}$, the operator $\mathsf{f}_k(n) = \mathsf{f}_Q(n) - \mathsf{f}_Q(n+k)$ is such that $\mathsf{S}_{\mathsf{f}_k,0}$ is finite. Notice that the polynomial associated to f_k is $Q(X)(1-X^k)$. Therefore, since $\theta > 1$ and $Q(\theta) \neq 0$ (or $\theta = \infty$) by assumption, we get by Proposition 3.2.2 that $\mathsf{S}_{\mathsf{f}_k,0}$ is finite. □

As a corollary of Proposition 3.3.1 and the following result, we obtain the superstability of R^0_{ind}, the induced structure on R by equations with coefficients in $\mathbf{Z}$ (see Definition 2.2.3).

Proposition 3.3.5 ([16, Proposition 5.9]). *Let $\mathcal{N}$ be the structure $(\mathbf{N}, S, S^{-1}, 0)$, where $S(n) = n+1$, $S^{-1}(n+1) = n$ and $S^{-1}(0) = 0$. Then $\mathcal{N}^1$ is superstable of U-rank 1.*

Recall that $\mathcal{N}^1$ is the expansion of $\mathcal{N}$ by unary predicates for all subsets of $\mathcal{N}$.

Corollary 3.3.6. *Let R be a regular set. Then R^0_{ind} is definably interpreted in $\mathcal{N}$.*

Proof. We interpret the domain of R^0_{ind} as $\mathbf{N}$. Let $a_1, \dots, a_s \in \mathbf{Z}\setminus\{0\}$. We need to interpret in $\mathcal{N}$ the set of s-tuples of elements in R that satisfy the equation $a_1x_1 + \dots + a_sx_s = 0$. For all $i \in [s]$, let f_i be the operator $n \mapsto a_i r_n$. We interpret $\{\bar{x} \in R^s \mid a_1x_1 + \dots + a_sx_s = 0\}$ as $\mathsf{S}_{\bar{\mathsf{f}},0}$ in $\mathcal{N}$. Let us show that $\mathsf{S}_{\bar{\mathsf{f}},0}$ is definable in $\mathcal{N}$. As explained at the beginning of Section 3.3, the set

$$\mathsf{S}_{\bar{\mathsf{f}},0} = \bigcup_{\bar{I} \in \mathsf{Part}([s])} \mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}},0,\bar{I}}.$$

So we need only to show that $\mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}},0,\bar{I}}$ is definable in $\mathcal{N}$ for all $\bar{I} \in \mathsf{Part}([s])$. Let $\bar{I} \in \mathsf{Part}([s])$. Recall that $\mathsf{S}_{\bar{\mathsf{f}},0,\bar{I}}$ is defined as

$$\mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}},0,\bar{I}} = \left\{ \bar{n} \in \mathbf{N}^s \,\middle|\, \text{for all } j \in [k],\ \bar{n}_{I_j} \text{ is a non-degenerate solution of } \sum_{i \in I_j} \mathsf{f}_i(n_i) = 0 \right\}.$$

Therefore, it is enough to show that for all $j \in [k]$, the set

$$\mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}}_{I_j},0,I_j} = \left\{ \bar{n} \in \mathbf{N}^{|I_j|} \,\middle|\, \bar{n} \text{ is a non-degenerate solution of } \sum_{i \in I_j} \mathsf{f}_i(n_i) = 0 \right\}$$

is definable.

Applying Remark 3.3.2 to $\mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}}_{I_j},0,I_j}$, we only need to show that $\mathsf{S}^{\circ}_{\mathsf{f}_{\bar{n}_i},0}$ is definable for all $i \in [k]$. But as $\mathsf{S}^{\circ}_{\mathsf{f}_{\bar{n}_i},0}$ is either empty or cofinite in $\mathsf{S}_{\mathsf{f}_{\bar{n}_i},0}$, we only need to show that the latter is definable. But we know that, by Proposition 3.2.2, the set $\mathsf{S}_{\mathsf{f}_{\bar{n}_i},0}$ is either finite or $\mathbf{N}$, hence definable. □

3.4 Sums of operators

Throughout this section, we will use the following notations. Let $\bar{\mathsf{f}}$ be a tuple of k operators. Define $\mathrm{Im}(\bar{\mathsf{f}})$ as

$$\{a \in \mathbf{Z} \mid a = \mathsf{f}_1(n_1) + \cdots + \mathsf{f}_k(n_k) \text{ for some } \bar{n} \in \mathbf{N}^k\}.$$

Notice that $\mathrm{Im}(\bar{\mathsf{f}}) = \{a \in \mathbf{Z} \mid \mathsf{S}_{\bar{\mathsf{f}},a} \neq \emptyset\}$. Similarly define $\mathrm{Im}^+(\bar{\mathsf{f}})$ as $\mathrm{Im}(\bar{\mathsf{f}}) \cap \mathbf{N}$. We will prove that sets of the form $\mathrm{Im}^+(\bar{\mathsf{f}})$ are not too dense in $\mathbf{N}$, in the sense that $\mathrm{Im}^+(\bar{\mathsf{f}})$ does not contain arbitrarily long sequences with bounded gaps (such sets are called *piecewise syndetic*). As a result, we show that a set of the form $a + b\mathbf{N}$, where $a, b \in \mathbf{N}$ and $b > 0$, cannot be covered by finitely many sets of the form $z + \mathrm{Im}^+(\bar{\mathsf{f}})$, with $z \in \mathbf{Z}$. This property will be used later on to show that regular sets are bounded, in the sense of Definition 2.1.2.

Definition 3.4.1. Let $A \subset \mathbf{N}$. A is called *piecewise syndetic* if there exists $d \in \mathbf{N}^{>0}$ such that for all $k \in \mathbf{N}$, there exists $a_1 < \cdots < a_k \in A$ such that $a_{i+1} - a_i \leq d$ for all $i \in [k-1]$.

A key property of piecewise syndetic sets is that they are *partition regular*: any partition of a piecewise syndetic set must contain a piecewise syndetic set.

Theorem 3.4.2 (Brown's Lemma [30, Theorem 10.37])**.** *Let $A \subset \mathbf{N}$ be piecewise syndetic. If $A = A_1 \cup \cdots \cup A_n$, then there exists $i \in [n]$ such that A_i is piecewise syndetic.*

The main result of this section is the following.

Theorem 3.4.3. *Let $a, b \in \mathbf{N}$, $b > 0$. Then, the set $a + b\mathbf{N}$ cannot be covered by finitely many sets of the form $z + \mathrm{Im}^+(\bar{\mathsf{f}})$, where $\bar{\mathsf{f}}$ is a tuple of k operators, $k \in \mathbf{N}$ and $z \in \mathbf{Z}$.*

In the next proposition, we show that the image of arbitrary linear combinations of operators is not piecewise syndetic.

Proposition 3.4.4. *Let $\bar{\mathsf{f}}$ be a tuple of k operators. Then $\mathrm{Im}^+(\bar{\mathsf{f}})$ is not piecewise syndetic.*

Before giving a proof of Proposition 3.4.4, let us show how it is used to prove Theorem 3.4.3.

Proof of Theorem 3.4.3. Since $a + b\mathbf{N}$ is piecewise syndetic, if it were covered by sets of the form $z + \mathrm{Im}^+(\bar{\mathsf{f}})$, then one of them would also be piecewise syndetic, by Brown's Lemma. But this would imply that a set of the form $\mathrm{Im}^+(\bar{\mathsf{f}})$ is piecewise syndetic since any translate of a piecewise syndetic set is again piecewise syndetic. This contradicts Proposition 3.4.4. □

The proof of Proposition 3.4.4 is done by induction on the length of the tuple $\bar{\mathsf{f}}$. Our first step is the following lemma, which explains how to apply the induction hypothesis. In this lemma, we show that the elements of $\mathrm{Im}^+(\bar{\mathsf{f}})$ that have a fixed gap with another element of $\mathrm{Im}^+(\bar{\mathsf{f}})$ lie in a finite union of translates of sets of the form $\mathrm{Im}^+(\bar{\mathsf{f}}_I)$, where $I \subsetneq [s]$.

Lemma 3.4.5. *Let $\mathsf{f}_1, \dots, \mathsf{f}_k$ be operators and $e \in \mathbf{N}^{>0}$. Let X_e be the set $\{a \in \mathrm{Im}^+(\bar{\mathsf{f}}) \mid \exists a' \in \mathrm{Im}^+(\bar{\mathsf{f}}), |a - a'| = e\}$. Then there exists a finite set Z of integers such that*

$$X_e \subset \bigcup_{z \in Z} \bigcup_{I \subsetneq [k]} (z + \mathrm{Im}^+(\bar{\mathsf{f}}_I)).$$

Proof. We first identify Z. Let $a \in X_e$. By definition, there is $a' \in \mathrm{Im}^+(\bar{\mathsf{f}})$ such that $e = a - a'$ or $e = a' - a$, that we shorten by $e = \pm(a - a')$. Since both a and a' are in $\mathrm{Im}^+(\bar{\mathsf{f}})$, we can find $\bar{n}, \bar{n}' \in \mathbf{N}^k$ such that

$$a = \sum_{i=1}^{k} \mathsf{f}_i(n_i) \text{ and } a' = \sum_{i=1}^{k} \mathsf{f}_i(n'_i).$$

Since $e = \pm(a - a')$ we can find $I, I' \subset [k]$ such that

$$e = \pm \left(\sum_{i \in I} \mathsf{f}_i(n_i) - \sum_{i \in I'} \mathsf{f}_i(n'_i) \right),$$

$(\bar{n}_I, \bar{n}'_{I'}) \in \mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}}_I \cup -\bar{\mathsf{f}}_{I'}, e} \cup \mathsf{S}^{\mathrm{nd}}_{-\bar{\mathsf{f}}_I \cup \bar{\mathsf{f}}_{I'}, e}$ and

$$0 = \sum_{i \notin I} \mathsf{f}_i(n_i) - \sum_{i \notin I'} \mathsf{f}_i(n'_i).$$

We thus let Z be the set

$$\left\{ \sum_{i \in I} \mathsf{f}_i(n_i), \sum_{i \in I'} \mathsf{f}_i(n'_i) \,\middle|\, I, I' \subset [k], (\bar{n}_I, \bar{n}'_{I'}) \in \mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}}_I \cup -\bar{\mathsf{f}}_{I'}, e} \cup \mathsf{S}^{\mathrm{nd}}_{-\bar{\mathsf{f}}_I \cup \bar{\mathsf{f}}_{I'}, e} \right\} \cup \{0\}.$$

By Corollary 3.3.3, we have that the sets $\mathsf{S}^{\mathrm{nd}}_{-\bar{\mathsf{f}}_I \cup \bar{\mathsf{f}}_{I'}, e}$ and $\mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}}_I \cup -\bar{\mathsf{f}}_{I'}, e}$ are finite. Hence Z is finite.

Let us show that

$$X_e \subset \bigcup_{z \in Z} \bigcup_{I \subsetneq [k]} (z + \mathrm{Im}^+(\bar{\mathsf{f}}_I)).$$

Let $a \in X_e$. As in the first part of the proof there is $a' \in \mathrm{Im}^+(\bar{\mathsf{f}})$ such that $|a - a'| = e$ and there are $I, I' \subset [k]$ such that

$$e = \pm \left(\sum_{i \in I} \mathsf{f}_i(n_i) - \sum_{i \in I'} \mathsf{f}_i(n'_i) \right),$$

$(\bar{n}_I, \bar{n}'_{I'}) \in \mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}}_I \cup -\bar{\mathsf{f}}_{I'}, e} \cup \mathsf{S}^{\mathrm{nd}}_{-\bar{\mathsf{f}}_I \cup \bar{\mathsf{f}}_{I'}, e}$ and

$$0 = \sum_{i \notin I} \mathsf{f}_i(n_i) - \sum_{i \notin I'} \mathsf{f}_i(n'_i).$$

In order to show that

$$a \in \bigcup_{z \in Z} \bigcup_{I \subsetneq [k]} (z + \mathrm{Im}^+(\bar{\mathsf{f}}_I)),$$

we distinguish three cases.

1. $I = [k]$. In that case, $a \in Z$.
2. $\emptyset \neq I \subsetneq [k]$. In that case, $[k] \setminus I$ is a proper subset of $[k]$ and

$$a = z + \sum_{i \in [k] \setminus I} \mathsf{f}_i(n_i),\ z = \sum_{i \in I} \mathsf{f}_i(n_i) \in Z.$$

3. $I = \emptyset$. Since $e > 0$, we have that $I' \neq \emptyset$. Now if $I' = [k]$, we have $a = 0 \in Z$. So let us assume that $I' \subsetneq [k]$. In that case, $a = \sum_{i \in [k] \setminus I'} \mathsf{f}_i(n'_i)$. Since $[k] \setminus I'$ is a proper subset of $[k]$, a has the required form. □

We now prove Proposition 3.4.4 by induction on the length of the tuple $\bar{\mathsf{f}}$.

Proof of Proposition 3.4.4. Let f be an operator. By Lemma 3.4.5, we have that X_e is finite for all $e \in \mathbf{N}^{>0}$. This implies that $\mathrm{Im}^+(\mathsf{f})$ cannot be piecewise syndetic.

Let $k > 1$ and assume that the proposition holds for all tuple $\bar{\mathsf{f}}$ of length $\leq k$. Let $\mathsf{f}_1, \ldots, \mathsf{f}_{k+1}$ be operators such that $\mathrm{Im}^+(\bar{\mathsf{f}})$ is infinite. Suppose, towards a contradiction that $\mathrm{Im}^+(\bar{\mathsf{f}})$ is piecewise syndetic. Assume $d \in \mathbf{N}^{>0}$ witnesses the fact that $\mathrm{Im}^+(\bar{\mathsf{f}})$ is piecewise syndetic. Recall that for $e \in \mathbf{N}^{>0}$ we defined X_e as $\{a \in \mathrm{Im}^+(\bar{\mathsf{f}}) \mid \exists a' \in \mathrm{Im}^+(\bar{\mathsf{f}}), |a - a'| = e\}$. Even though $X_1 \cup \cdots \cup X_d$ may not equal $\mathrm{Im}^+(\bar{\mathsf{f}})$, this subset will play a key role in the rest of the proof, as it is the "piecewise syndetic part of $\mathrm{Im}^+(\bar{\mathsf{f}})$ with respect to d". Indeed, the set $X_1 \cup \cdots \cup X_d$ is itself piecewise syndetic so that by Brown's Lemma, there exists $i \in [d]$ such that X_i is also piecewise syndetic. But by Lemma 3.4.5 we know that X_i is contained in a finite union of sets of the form $z + \mathrm{Im}^+(\bar{\mathsf{f}}')$, where $\bar{\mathsf{f}}'$ is of length $\leq k$. But this is implies, by Brown's Lemma and the fact that a set containing a piecewise syndetic set is itself piecewise syndetic, the existence of a piecewise syndetic set of the form $z + \mathrm{Im}^+(\bar{\mathsf{f}}')$, where $\bar{\mathsf{f}}'$ is of length $\leq k$. This contradicts our induction hypothesis. So $\mathrm{Im}^+(\bar{\mathsf{f}})$ is not piecewise syndetic, which is what we wanted. □

We are now ready to prove that regular sets are bounded.

Corollary 3.4.6. *Let R be a regular set. Then R is bounded.*

Proof. The proof follows [37, Lemma 3.4 and Lemma 3.5] and is an application of Theorem 2.1.9. First let us show that, for any $\mathcal{L}_g$-formula $\varphi(\bar{x}, y, \bar{z})$, any consistent set the form

$$\Gamma(y) = \{\varphi(\bar{b}, y, \bar{\alpha}) \mid \bar{\alpha} \in R^n\},$$

where $\bar{b} \in \mathbf{Z}$ and n is the length of the tuple $\bar{z}$, is realized by some $s \in \mathbf{Z}$.

Using quantifier elimination in $\mathcal{L}_g$, we can assume that φ is a disjunction of conjunctions of atomic formulas and negations of atomic formulas. By the consistency of $\Gamma(y)$, we may select for each $\bar{\alpha} \in R^n$, one of the disjunctive clauses. So we may assume, using the fact that $\forall x(\neg D_n(x) \leftrightarrow \bigvee_{k=1}^{n-1} D_n(x+k))$, that $\varphi(\bar{b}, y, \bar{\alpha})$ is of the form

$$\bigwedge_{i \in I_{1,\bar{\alpha}}} t_i(\bar{b}, y, \bar{\alpha}) = 0 \wedge \bigwedge_{i \in I_{2,\bar{\alpha}}} t_i(\bar{b}, y, \bar{\alpha}) \neq 0 \wedge \bigwedge_{i \in I_{3,\bar{\alpha}}} D_{n_i}(t_i(\bar{b}, y, \bar{\alpha})),$$

where $t_i(\bar{x}, y, \bar{z})$ is a term for all $i \in I_{1,\bar{\alpha}} \cup I_{2,\bar{\alpha}}$ and $j \in I_{3,\bar{\alpha}}$. We may further assume that $I_{1,\bar{\alpha}} = \emptyset$ (otherwise $y \in \mathbf{Z}$ since $\bar{b} \in \mathbf{Z}$ and $R \subset \mathbf{Z}$). Given $i \in I_{2,\bar{\alpha}} \cup I_{3,\bar{\alpha}}$, the term $t_i(\bar{b}, y, \bar{z})$ is equal to $m_i y + d_i + a_1 z_1 + \cdots + a_{\ell_i} z_{\ell_i}$, where $m_i, d_i \in \mathbf{Z}$, $\bar{a}_i \in \mathbf{Z}^{\ell_i}$ and $\ell_i \in \mathbf{N}$. Notice that we may assume that $m_i = m_{\bar{\alpha}}$ for all $i \in I_{2,\bar{\alpha}} \cup I_{3,\bar{\alpha}}$ (otherwise, we multiply the inequation $t_i(\bar{x}, y, \bar{z}) \neq 0$ by $k_i = \prod_{j \neq i} m_j$ for all $i \in I_2 \cup I_3$, and replace $D_{n_i}(t_i(\bar{x}, y, \bar{z}))$ by the equivalent formula $D_{k_i n_i}(k_i t_i(\bar{x}, y, \bar{z})))$. Likewise, since the set $\{m_{\bar{\alpha}} \mid \bar{\alpha} \in R^n\}$ is finite, we may assume that $m_{\bar{\alpha}} = m$ for all $\bar{\alpha} \in R^n$. Also, note that the set

$$C = \bigcup_{\bar{\alpha} \in R} \bigcup_{i \in I_{2,\bar{\alpha}}} \{(d_i, \bar{a}_i)\}$$

is finite.

Thus $\Gamma(y)$ expresses the fact that y is in a coset of a subgroup of $\mathbf{Z}$, say $c + d\mathbf{Z}$ for some $c, d \in \mathbf{N}$, and $-my$ is not in the set

$$X = \bigcup_{(d,\bar{a}) \in C} \{d + a_1 z_1 + \cdots + a_\ell z_\ell \mid \bar{z} \in R^\ell, \ell = |\bar{a}|\}.$$

But, by Theorem 3.4.3, X does not cover $mc + md\mathbf{N}$ (use operators of the form $n \mapsto ar_n$ to apply the theorem). So there is $s \in \mathbf{N}$ such that $m(c + ds)$ is not in X, which is what we wanted.

Let $\varphi(\bar{x}, y, \bar{z})$ be an $\mathcal{L}_g$-formula. Let $\ell = |\bar{x}|$ and $n = |\bar{z}|$. Since $\mathscr{Z}$ has nfcp, there exists $k \in \mathbf{N}$ such that for all $\bar{b} \in \mathbf{Z}^\ell$ if $\Gamma(y) = \{\varphi(\bar{b}, y, \bar{\alpha}) | \bar{\alpha} \in R^n\}$ is k-consistent, then it is consistent. Thus, we have that

$$\mathscr{Z}_R \models \forall \bar{x} \left(\left(\forall \bar{z}_1 \in R \ldots \forall \bar{z}_k \in R \exists y \bigwedge_{j \in [k]} \varphi(\bar{x}, y, \bar{z}_j) \right) \rightarrow \exists y \forall \bar{z} \in R \varphi(\bar{x}, y, \bar{z}) \right).$$

Thus by Theorem 2.1.9, R is bounded. □

Remark 3.4.7. As we explained in Chapter 2, it is unnecessary to establish that the pair $\mathscr{Z}_R$ is bounded, in view of Theorem 2.2.6. However, the material developed in this section will be needed in Chapter 4 for our quantifier elimination result, namely in Proposition 4.4.2.

Using Corollary 3.4.6, we can show that two tuples have the same $\mathcal{L}_R$-type over A if they have the same type in $\mathcal{L}_g$ over $R \cup A$.

Proposition 3.4.8 ([37, Corollary 3.7])**.** *Let R be a regular set. Let $\mathscr{G}$ be a monster model of* $\mathrm{Th}(\mathscr{Z}_R)$. *Let $A \subset G$ and $\bar{a}, \bar{b} \in A^n$. Then* $\mathrm{tp}^{\mathcal{L}_R}(\bar{a}/A) = \mathrm{tp}^{\mathcal{L}_R}(\bar{b}/A)$ *if* $\mathrm{tp}^{\mathcal{L}_g}(\bar{a}/R(G), A) = \mathrm{tp}^{\mathcal{L}_g}(\bar{b}/R(G), A)$.

Proof. Assume that $\mathrm{tp}^{\mathcal{L}_g}(\bar{a}/R(G), A) = \mathrm{tp}^{\mathcal{L}_g}(\bar{b}/R(G), A)$. By Corollary 3.4.6, we have to show that $\bar{a}$ and $\bar{b}$ satisfy the same bounded formulas. We proceed by induction on n, the number of bounded quantifiers in $\varphi(\bar{x})$. Let $\varphi(\bar{x})$ be a bounded formula with parameters in $R(G) \cup A$, that is $\varphi(\bar{x})$ is of the form

$$Q_1 y_1 \in R, \ldots, Q_n y_n \in R, \psi(\bar{x}, \bar{y}),$$

where $n \in \mathbf{N}$, $Q_i \in \{\forall, \exists\}$ for all $i \in [n]$ and $\psi(\bar{x}, \bar{y})$ is a $\mathcal{L}_g$-formula with parameters in $R(G) \cup A$. If $n = 0$, then there is nothing to do since we assumed that $\mathrm{tp}^{\mathcal{L}_g}(\bar{a}/R(G), A) = \mathrm{tp}^{\mathcal{L}_g}(\bar{b}/R(G), A)$. Assume that $n > 0$ and that $\bar{a}$ and $\bar{b}$ satisfy the same bounded formulas with $k < n$ bounded quantifiers. Let us show that $\mathscr{G} \models \varphi(\bar{a}) \leftrightarrow \varphi(\bar{b})$. We may assume, without loss of generality, that $Q_1 = \exists$. Assume that $\mathscr{G} \models \varphi(\bar{a})$. Then there exists $c \in R(G)$ such that

$$\mathscr{G} \models Q_2 y_2 \in R, \ldots, Q_n y_n \in R, \psi(\bar{a}, c, y_2, \ldots, y_n).$$

But the formula $Q_2 y_2 \in R, \ldots, Q_n y_n \in R, \psi(\bar{x}, c, y_2, \ldots, y_n)$ is a bounded formula with parameters in $R(G) \cup A$ and $n - 1$ bounded quantifiers. So by the inductive assumption

$$\mathscr{G} \models Q_2 y_2 \in R, \ldots, Q_n y_n \in R, \psi(\bar{b}, c, y_2, \ldots, y_n).$$

Likewise, we show that if $\mathscr{G} \models \varphi(\bar{b})$ then $\mathscr{G} \models \varphi(\bar{a})$.

In conclusion, $\mathrm{tp}^{\mathcal{L}_R}(\bar{a}/A) = \mathrm{tp}^{\mathcal{L}_R}(\bar{b}/A)$. □

3.5 Superstability

We are now able to prove the main theorem of this chapter, which states that expansions of $\mathscr{Z}$ by regular sets are superstable of Lascar rank ω.

Theorem 3.5.1. *Let R be a regular set. Then* $\mathrm{Th}(\mathscr{Z}_R)$ *is superstable of Lascar rank ω.*

Proof. By Proposition 3.3.5 and Corollary 3.3.6, we get that R^0_{ind} is superstable. Furthermore, by Corollary 3.4.6 R is bounded. So, we deduce from Theorem 2.1.3 that $\mathscr{Z}_R$ is superstable. Since $\mathscr{Z}_R$ is a proper expansion of $\mathscr{Z}$, it must have Lascar rank $\geq \omega$ by [37, Theorem 1]. So what remains to be shown is that the rank is $\leq \omega$. We follow the proof of [37, Theorem 2], which covers the case of $R = \{q^n \mid n \in \mathbf{N}\}$, $q \in \mathbf{N}^{>1}$. First, let us note that the U-rank of R is 1, since by Proposition 3.3.5, $\mathscr{N}$ has U-rank 1 and by Corollary 3.3.6, R_{ind} is definably interpreted in $\mathscr{N}$. Now we prove that $U(\text{Th}(\mathscr{Z}_R)) \leq \omega$. Let $\mathscr{G} \succ \mathscr{Z}_R$ be a monster model. Since $\text{Th}(\mathscr{Z}_R)$ is superstable, by Proposition 1.1.14, it is enough to show that $U(p) \leq \omega$ where p is the principal generic over $\mathbf{Z}$. By definition, this amounts to show that any forking extension of p has finite U-rank.

First let us show that if $h \in \text{acl}^{\mathcal{L}_R}(R(G), B)$, then $U(h/B) < \omega$. Let $\bar{c} \in R(G)^n$ be such that $h \in \text{acl}^{\mathcal{L}_R}(\bar{c}, B)$. In particular $U(h/\bar{c}, B) = 0$. By Proposition 3.3.5 and Corollary 3.3.6 and Lascar's inequality, $U(\bar{c}) \leq n$. Then by Proposition 1.1.3, $U(h, \bar{c}/B) = U(h/\bar{c}, B) + U(\bar{c}/B) = U(\bar{c}/B)$ and $U(h, \bar{c}/B) = U(\bar{c}/h, B) + U(h/B)$. So, $U(h/B) \leq U(\bar{c}/B) \leq n < \omega$.

Now consider a forking extension of p, say $\text{tp}^{\mathcal{L}_R}(b/B)$, $\mathbf{Z} \subset B$, and $q = \text{tp}(a/B)$ a non-forking extension of p. In particular, $U(q) \geq \omega$. We want to show that that $b \in \text{acl}^{\mathcal{L}_R}(R(G), B)$, which would imply, by the calculations above, that $U(q)$ is finite.

Assume on the contrary that $b \notin \text{acl}^{\mathcal{L}_R}(R(G), B)$. Since $U(p) = U(a/B) \geq \omega$, we have that $a \notin \text{acl}^{\mathcal{L}_R}(R(G), B)$. Hence $a \notin \text{acl}^{\mathcal{L}_g}(R(G), B)$. So, in the $\mathcal{L}_g$-theory of $\mathscr{Z}$, $U(a/R(G), B) = 1$. Thus, $\text{tp}^{\mathcal{L}_g}(a/R(G), B)$ is a generic type. By the same reasoning, this is also true for $\text{tp}^{\mathcal{L}_g}(b/R(G), B)$. So by Proposition 1.1.15 $\text{tp}^{\mathcal{L}_g}(b/R(G), B) = \text{tp}^{\mathcal{L}_g}(a/R(G), B)$. This implies, by Proposition 3.4.8, $\text{tp}^{\mathcal{L}_R}(b/B) = \text{tp}^{\mathcal{L}_R}(a/B)$. Therefore $\text{tp}^{\mathcal{L}_R}(b/B)$ is also generic. By Proposition 1.1.10, we have that q is a non-forking extension of p, a contradiction. □

3.6 Comparison with geometrically sparse sets

When we were developing the material of this chapter, G. Conant independently developed the notion of geometrically sparse set in [16]. Because there is an overlap between Theorem 3.5.1 and [16, Theorem 7.1], we believe that a comparison between the two notions involved is in order.

Definition 3.6.1 ([16, Definition 6.2])**.** Let $A \subset \mathbf{Z}$, $(r_n) \subset \mathbf{Z}$ and $B \subset \mathbf{R}^{>0}$.

1. B is said to be *geometric* if $\{a/b \mid a \geq b, a, b \in B\}$ is closed and discrete.
2. A is said to be *geometrically sparse* if there is a function $f : A \to \mathbf{R}^{>0}$ such that $f(A)$ is geometric and $\sup\{|a - f(a)| \mid a \in A\} < \infty$. A sequence (r_n) is said to be *geometrically sparse* if $\{r_n \mid n \in \mathbf{N}\}$ is geometrically sparse.

In case (r_n) is strictly increasing, then if (r_n) is geometrically sparse, there exists an increasing sequence $(\lambda_n) \subset \mathbf{R}^{\geq 1}$ such that $X = \{\lambda_m/\lambda_n \mid n < m\}$ is closed and discrete and $\sup_{n\in\mathbf{N}} |r_n - \lambda_n| < \infty$, see [16, Proposition 7.2]. The proof of [16, Proposition 7.2] requires a bit of work since a witness $f : \{r_n | n \in \mathbf{N}\} \to \mathbf{R}^{>0}$ of the fact that (r_n) is geometrically sparse is not necessarily strictly increasing.

We recall the main result of [16].

Theorem 3.6.2 ([16, Theorem 7.1]). *Let $A \subset \mathbf{Z}$. If A is geometrically sparse, then* $\mathrm{Th}(\mathscr{Z}_A)$ *is superstable of U-rank ω.*

In our comparison between our theorem and [16, Theorem 7.1], we will use the following lemma about geometrically sparse sequences (r_n) with a Kepler limit in $\mathbf{R}^{>1}$.

Lemma 3.6.3. *Let (r_n) be a geometrically sparse sequence such that $r_{n+1}/r_n \to \theta \in \mathbf{R}^{>1}$. Then there exists $\tau \in \mathbf{R}^{\geq 1}$ such that $r_n/\theta^n \to \tau$.*

Proof. Let A and $f : \{r_n \mid n \in \mathbf{N}\} \to A$ witness the fact that (r_n) is geometrically sparse. Put $\lambda_n = f(r_n)$. As $\sup_{n\in\mathbf{N}} |r_n - \lambda_n| < k$, we have that

$$\lim_{n\to\infty} \frac{\lambda_n}{r_n} = 1.$$

Since $r_{n+1}/r_n \to \theta$, we get

$$\begin{aligned}\lim_{n\to\infty} \frac{\lambda_{n+1}}{\lambda_n} &= \lim_{n\to\infty} \frac{\lambda_{n+1}}{r_{n+1}} \frac{r_n}{\lambda_n} \frac{r_{n+1}}{r_n} \\ &= \theta.\end{aligned}$$

But, as X is closed and discrete, if the sequence $(\lambda_{n+1}/\lambda_n)$ converges, then it is ultimately constant and so ultimately equal to θ. As a result, there exists $\tau \in \mathbf{R}^{\geq 1}$ such that $\lambda_n = \tau\theta^n$ for all $n \in \mathbf{N}$ sufficiently large. Hence (λ_n/θ^n) converges to τ. □

Now, let us discuss the overlap between Theorem 3.5.1 and [16, Theorem 7.1]:

– the case where $r_{n+1}/r_n \to \infty$ is completely covered by [16, Theorem 7.1] (as a consequence of [16, Proposition 6.3]);

– the case where $r_{n+1}/r_n \to \theta$ and θ is algebraic, there are examples of regular sequences that are not geometrically sparse. For instance, we can show by direct calculations that the sequence defined by $r_{n+2} = 5r_{n+1} + 7r_n$, $r_1 = 1$ and $r_0 = 0$, is regular but not geometrically sparse. Indeed, first notice that for all $n \in \mathbf{N}$, $r_n = \alpha(\lambda_+^n - \lambda_-^n)$, where $\alpha = 1/\sqrt{53}$, $\lambda_\pm = (5 \pm \sqrt{53})/2$. Then assume that there is a sequence (λ_n) such that $\sup\{|r_n - \lambda_n| \mid n \in \mathbf{N}\} = k \in \mathbf{R}$. Let $\kappa_n = \lambda_n/\lambda_+^n$

and note that $\kappa_n \to \alpha$. Now we have $|r_n - \lambda_n| = |(\alpha - \kappa_n)\lambda_+^n - \alpha\lambda_-^n|$. We want to show that $\{\lambda_m/\lambda_n \mid n \leq m\}$ cannot be both closed and discrete. Assume towards a contradiction that $\{\lambda_m/\lambda_n \mid n \leq m\}$ is closed and discrete. In that case, λ_{n+1}/λ_n is ultimately equal to λ_+. Thus κ_{n+1}/κ_n is ultimately equal to 1. This in turn implies that $\kappa_n = \alpha$ for all sufficiently large $n \in \mathbf{N}$. But in this case, $|r_n - \lambda_n| = |\alpha\lambda_-^n|$ for all sufficiently large $n \in \mathbf{N}$, in contradiction with the boundedness of $(|r_n - \lambda_n|)$;

– for the case where $r_{n+1}/r_n \to \theta$ and θ is transcendental, we can construct non-geometrically sparse sequences that are regular from geometrically sparse sequences. More precisely, if (r_n) is geometrically sparse, that is $\sup_{n\in\mathbf{N}} |r_n - \lambda_n|$ is finite for some sequence $(\lambda_n) \subset \mathbf{R}^{\geq 1}$ such that $X = \{\lambda_m/\lambda_n \mid n < m\}$ is closed and discrete, then the sequence $(r_n + n)$ is not geometrically sparse but satisfies Theorem 3.5.1. Since there exists $\tau \in \mathbf{R}^{\geq 1}$ such that $r_n/\theta^n \to \tau$, we may assume, by Lemma 3.6.3, $\lambda_n = \tau\theta^n$.
Assume, towards a contradiction, that there is (λ'_n) such that $\sup_{n\in\mathbf{N}} |r_n + n - \lambda'_n| < \infty$ and $X' = \{\lambda'_m/\lambda'_n \mid n \leq m\}$ is closed in discrete. Now let $\kappa_n = \tau\theta^n + n - \lambda'_n$. Notice that (κ_n) is bounded since we assumed (r_n) geometrically sparse. So we have that $\lambda'_{n+1}/\lambda'_n \to \theta$. Since X' is closed in discrete, this last sequence is ultimately constant: for all $n \in \mathbf{N}$ sufficiently large,

$$\theta = \frac{\tau\theta^{n+1} + n + 1 - \kappa_{n+1}}{\tau\theta^n + n - \kappa_n}.$$

So, for all sufficiently large $n \in \mathbf{N}$, $n(\theta - 1) = 1 - \kappa_{n+1} + \theta\kappa_n$, a contradiction.

For the remainder of this section we investigate further the class of geometrically sparse sequence that satisfy a linear recurrence relation. We characterize geometrically sparse sequences among linear recurrence sequences that have a Kepler limit in $\mathbf{R}^{>1}$. In particular, we show that those sequences must have a Kepler limit that is either a Pisot number or a Salem number[2].

Theorem 3.6.4. *Let $(r_n) \subset \mathbf{Z}$ be a linear recurrence sequence, with minimal polynomial P. Assume that (r_n) has a Kepler limit $\theta \in \mathbf{R}^{>1}$. Then (r_n) is geometrically sparse if and only if θ is the only root of P such that $|\theta| > 1$, has multiplicity 1 and if θ' is a root of P of modulus 1 then it is a root of multiplicity 1. In particular, θ is either a Pisot or Salem number.*

Proof. Using Binet's formula (cf. Theorem 1.3.4), we know that for all $n \in \mathbf{N}$,

$$r_n = \sum_{i=1}^{k} P_i(n)\theta_i^n,$$

[2] Let $\theta \in \mathbf{R}^{>1}$ be algebraic over $\mathbf{Q}$ and let P_θ be its minimal polynomial. Then θ is a *Pisot number* (resp. *Salem number*) if it is the only root of P_θ of modulus ≥ 1 (resp. > 1 and P_θ has at least one root of modulus 1).

where $\deg(P_i) + 1$ is the multiplicity of θ_i as a root of P. We may assume that $|\theta_i| \geq |\theta_{i+1}|$ and $|\theta_i| = |\theta_{i+1}| \Rightarrow \deg(P_i) \geq \deg(P_{i+1})$.

Assume first θ has multiplicity 1 and is the only root of P of modulus > 1 and if θ' is a root of P of modulus 1 then it is a root of multiplicity 1. In that case, by Theorem 1.3.8, $\theta = \theta_1$ and $|\theta_1| > |\theta_2|$. Also, $P_1(n) = c_1$ since θ has multiplicity 1 as a root of P. Let $A = \{c_1\theta^n \mid n \in \mathbf{N}\}$ and let $f : \{r_n \mid n \in \mathbf{N}\} \to A : r_n \mapsto c_1\theta^n$. Let us show that (r_n) is geometrically sparse using A and f. We have $X = \{a/b \mid a \geq b, a, b \in A\} = \{\theta^n | n \in \mathbf{N}\}$. Also $X \subset \mathbf{R}^{>0}$, since $\theta > 0$. As $\theta > 1$, any convergent sequence in X is either ultimately constant or tends towards ∞, hence X is closed and discrete. This shows that A is geometric. To conclude, let us show that $\sup\{|r_n - f(r_n)| \mid n \in \mathbf{N}\}$ is finite. We have, for all $n \in \mathbf{N}$,

$$\begin{aligned}
|r_n - f(r_n)| &= \left|\sum_{i=1}^{k} P_i(n)\theta_i^n - c_1\theta_1^n\right| \\
&= \left|\sum_{i=2}^{k} P_i(n)\theta_i^n\right| \\
&\leq \sum_{\substack{i\in[k] \\ |\theta_i|=1}} |P_i(n)\theta_i^n| + \sum_{\substack{i\in[k] \\ |\theta_i|<1}} |P_i(n)\theta_i^n|.
\end{aligned}$$

By assumption, we have that $P_i(n) = c_i$ for all $i \in [k]$ such that $|\theta_i| = 1$. Also,

$$\sum_{\substack{i\in[k] \\ |\theta_i|<1}} |P_i(n)\theta_i^n| \to 0.$$

So, for all $n \in \mathbf{N}$ sufficiently large,

$$|r_n - f(r_n)| \leq \sum_{\substack{i\in[k] \\ |\theta_i|=1}} |c_i| + 1.$$

Hence $\sup\{|r_n - f(r_n)| \mid n \in \mathbf{N}\}$ is finite.

Let us now assume that (r_n) is geometrically sparse, as witnessed by the geometric set A and the map $f : \{r_n \mid n \in \mathbf{N}\} \to A$. Let $\lambda_n = f(r_n)$. Since $\sup\{|r_n - \lambda_n| \mid n \in \mathbf{N}\} < \infty$, we have that (λ_n) has Kepler limit θ. Therefore, since $\{\lambda_m/\lambda_n \mid \lambda_n \leq \lambda_n\}$ is closed and discrete, there exists $c \in \mathbf{R}$ such that $\lambda_n = c\theta^n$ for all n sufficiently large. Hence we may assume that $\lambda_n = c\theta^n$ for all $n \in \mathbf{N}$. As a result, for all $n \in \mathbf{N}$,

$$\begin{aligned}
r_n - \lambda_n &= (P_1(n) - c)\theta_1^n + \sum_{i=1}^{k} P_i(n)\theta_i^n \\
&= (P_1(n) - c)\theta_1^n + o(|P_1(n)\theta_1^n|).
\end{aligned}$$

Thus, since $\sup\{|r_n - \lambda_n| \mid n \in \mathbf{N}\} < \infty$, we must have $P_1(n) = c$ for all $n \in \mathbf{N}$ and $|\theta_2| \leq 1$. To finish the proof, we need to show that $\deg(P_2) = 0$. But this is a consequence of [1, Lemma 3.2], which states that a sequence (s_n) of the form $s_n = \sum_{i=1}^{\ell} Q_i(n)\eta_i^n$, where $Q_i \in \mathbf{C}[X]$ and $\eta_1, \ldots, \eta_\ell$ are pairwise distinct complex numbers of modulus 1, is bounded only when $\deg(Q_i) = 0$ for all $i \in [\ell]$. □

We point out that the minimal polynomial of a linear recurrence sequence with a Kepler limit that is geometrically sparse is not necessarily the minimal polynomial of its Kepler limit. For instance, consider $P(X) = (X^2 - 2X - 1)(X - 1)$ and the recurrence $r_n = (1+\sqrt{2})^n + (1-\sqrt{2})^n + 1$. By Theorem 3.6.4, (r_n) is geometrically sparse, with Kepler limit $1+\sqrt{2}$, whose minimal polynomial, $X^2 - 2X - 1$, is not the minimal polynomial of (r_n). Notice that (r_n) is not a regular sequence.

The situation when we remove the Kepler limit assumption is less clear. While we do not believe that the statement of Theorem 3.6.4 is true without this assumption, we think that it is plausible if we require (r_n) to be non-degenerate. However, we do not have a proof of such a result.

3.7 Expansions of $(\mathbf{Z},+,0)$ and the Mann property

Let $\mathscr{K}$ be a field of characteristic zero and consider equations of the form $\sum_{i=1}^{n} q_i x_i = 1$, where $q_i \in \mathbf{Q} \setminus \{0\}$. Let $A \subset K$. A solution $\bar{a}$ in A^n is non-degenerate if $\sum_{j \in J} q_j a_j \neq 0$, for any *proper* subset J of $[n]$. The set A has the Mann property if any such equation has only finitely many non-degenerate solutions. This terminology comes from the work of L. van den Dries and A. Günaydın on expansions of algebraically closed fields or real-closed fields K by a *small* (in the sense of [20, Section 2], and which differs from Definition 2.1.7) subgroup G of the multiplicative group of the field. In [20] it is mentioned that the expression *Mann property* comes from a result of H. Mann that states that the multiplicative group of roots of unity in $\mathbf{C}$ has the Mann property (see [32]). The paper [20] is concerned with the model theory of pairs $\mathscr{K}_G$, where $\mathscr{K}$ is either algebraically closed of characteristic 0 or real closed, and where G has the Mann property. One of their results is a characterization of elementary equivalence between those structures (see [20, Theorems 1.2 and 1.3]).

In this section, we consider expansions of the form $(\mathbf{Z},+,0,M)$ where $(M,\cdot,1)$ is a submonoid of $(\mathbf{Z},\cdot,1)$ with the Mann property. Let G be the subgroup of $(\mathbf{Q}\setminus\{0\},\cdot,1)$ generated by M, then it has the Mann property. An examples of such monoid is $\langle 2^{\mathbf{Z}}, 3^{\mathbf{Z}}\rangle \cap \mathbf{N} = \langle P_2, P_3\rangle$. More generally, any finitely generated submonoid $(M,\cdot,1)$ of $(\mathbf{Z},\cdot,1)$ has the Mann property, since the corresponding group $(G,\cdot,1)$ has finite rank (as an abelian group, that is $\dim_{\mathbf{Q}} G \otimes_{\mathbf{Z}} \mathbf{Q}$ is finite), see [31, Théorème 1]. As mentioned in the introduction of this chapter, the result of this section is a special case of [15, Theorem 3.1], but with a short proof.

Let $(M, \cdot, 1)$ be a submonoid of $(\mathbf{Z}, \cdot, 1)$. Let $\mathcal{L}_M$ be the language $\{1, s | s \in M\}$, where s is a unary function interpreted as $s(m) = s \cdot m$. Let $\mathscr{M}$ be the $\mathcal{L}_M$-structure $(M, 1, s | s \in M)$. Finally let $\mathscr{M}^1$ be the expansion of $\mathscr{M}$ by unary predicates R_X for all subsets X of M. As before, we let $\mathcal{L}^1$ be the resulting language.

Lemma 3.7.1. *$T = \mathrm{Th}(\mathscr{M}^1)$ has quantifier elimination and is superstable.*

Proof. Let us show that T has quantifier elimination. Let $\mathscr{N}_1, \mathscr{N}_2$ be models of T and let $\mathscr{A}$ be a common substructure. Let $\varphi(x, \bar{y})$ be a quantifier-free formula. Let $\bar{a} \in A^{|\bar{y}|}$ and assume that there exists $b \in N_1 \setminus A$ such that $\mathscr{N}_1 \models \varphi(b, \bar{a})$. Let us show that these exists $b' \in N_2$ such that $\mathscr{N}_2 \models \varphi(b', \bar{a})$.

We may assume that $\varphi(x, \bar{y})$ is of the form

$$\bigwedge_{i \in I_1} s_i(x) = s'_i(y_i) \wedge \bigwedge_{i \in I_2} s_i(x) \neq s'_i(y_i) \wedge \psi(x),$$

where $\psi(x)$ is a quantifier-free formula. As for all $s \in M$, $T \models \forall xy\ (x = y \leftrightarrow s(x) = s(y))$, we can assume that for all $i \in I_1 \cup I_2$, $s_i = s$. Thus, $\varphi(x, \bar{y})$ is of the form

$$\bigwedge_{i \in I_1} s(x) = s'_i(y_i) \wedge \bigwedge_{i \in I_2} s(x) \neq s'_i(y_i) \wedge \psi(x).$$

We first assume that $I_1 \neq \emptyset$. Let $i_0 \in I_1$. Consider $X = \{m \in \psi(M) \mid m = sn \text{ for some } n \in M\}$. Then we have that $\mathscr{N} \models P_X(s'_{i_0}(a_{i_0}))$. But then there exists $b' \in \psi(N)$ such that $sb' = s'_{i_0} a_{i_0}$. Thus $\mathscr{N} \models \varphi(b', \bar{a})$.

Let us now assume that $I_1 = \emptyset$. We distinguish two cases, according to whether $\psi(N_1)$ is infinite. We first point out that, for $i \in \{1, 2\}$, the set

$$X_i = \left\{ x \in N_i \,\middle|\, \bigvee_{i \in I_2} s(x) = s'_i(a_i) \right\}$$

is finite in M_i.

1. If $\psi(N_1)$ is infinite, then $\psi(N_2)$ is also infinite (because T is complete). As a result, we can find $b' \in N_2$ such that $\mathscr{N}_2 \models \varphi(b', \bar{a})$, since X_2 is finite.
2. If $\psi(N_1)$ is finite. Let $n = |\psi(N_1)|$. We have that n is independent of $\mathscr{N}_1$, again by completeness. In particular, there are $s''_1, \ldots, s''_n \in \mathcal{L}_M$ such that

$$T \models \forall x \left(\psi(x) \to \bigvee_{i \in [n]} x = s''_i(1) \right).$$

Because $\mathscr{A}$ is a substructure of $\mathscr{N}_1$, we get that $b \in A$, a contradiction with our assumption.

Thus T has quantifier elimination. Then one could show superstability by counting types. We propose another approach. Quantifier elimination in T shows definable sets are boolean combinations of finite sets and $\emptyset$-definable sets. Therefore, T is *quasi strongly minimal* and by [5, Theorem 20] these theories are precisely the superstable with Lascar rank 1 ones. □

Theorem 3.7.2. *Let* $(M, \cdot, 1)$ *be a submonoid of* $(\mathbf{Z}, \cdot, 1)$ *such that* $(G, \cdot, 1)$*, the subgroup of* $(\mathbf{Q} \setminus \{0\}, \cdot, 1)$ *generated by M, has the Mann Property. Then* $\mathrm{Th}(\mathscr{Z}_M)$ *is superstable.*

Proof. We apply the same strategy as in the proof of Theorem 3.5.1. By Theorem 2.2.6, the pair $\mathscr{Z}_M$ is bounded. So to conclude the proof, we only need to show that M_{ind} is superstable. By Theorem 2.2.5, it is enough to show that M^0_{ind} is definably interpreted in $\mathscr{M}^1$.

Let $\varphi(\bar{x})$ be the equation $a_1x_1 + \cdots + a_nx_n = 0$. We have to show that the set $\varphi(M)$ corresponds to a definable subset of $\mathscr{M}^1$. As in the proof of Corollary 3.3.6, we only have to show that the set of non-degenerate solutions of $\varphi(\bar{x})$ corresponds to a definable subset of $\mathscr{M}^1$. This is an adaptation of the work done in [20, Section 5]. Indeed, since G has the Mann property the set of non-degenerate solutions of $\varphi(\bar{x})$ (in G) is

$$\bigcup_{(g_2,\ldots,g_n)\in S} (1, g_2, \ldots, g_n)G,$$

where S is the (finite) set of non-degenerate solutions of the equation

$$a_1 + a_2x_2 + \cdots + a_nx_n = 0.$$

Let $(g_2, \ldots, g_n) \in S$. Let $s \in M$ minimal such that $(sg_2, \ldots, sg_n) \in M^{n-1}$. Let $X_s = sG \cap M$. Consider the $\mathcal{L}^1$-formula $\psi_{s,\bar{g}}(\bar{x})$ defined as, setting $g_1 = 1$,

$$\exists x \Big(R_{X_s}(x) \wedge \bigwedge_{i\in[n]} x_i = sg_i(x) \Big).$$

Then we have that $\bar{x}$ is a non-degenerate solution of $\varphi(\bar{x})$ in M if and only if

$$\bigvee_{(g_2,\ldots,g_n)\in S} \psi_{s,\bar{g}}(\bar{x}).$$

This shows that the set of non-degenerate solutions of $\varphi(\bar{x})$ in M is definable in $\mathscr{M}^1$. □

3.8 Expansions of divisible torsion-free abelian groups by regular set

In this final section, we show that given a regular set $R \subset \mathbf{N}$, the pairs $(\mathbf{Q}, +, 0, R)$ and $(\mathbf{R}, +, 0, R)$ are ω-stable. This will be a consequence of quantifier elimination in

$\mathrm{Th}(\mathbf{Q}, +, 0)$ and the work done in Section 3.3 on the trace of equations with coefficients in $\mathbf{Z}$ on R.

Theorem 3.8.1. *Let R be a regular set. Then $(\mathbf{Q}, +, 0, R)$ and $(\mathbf{R}, +, 0, R)$ are ω-stable.*

Proof. Recall that $T = \mathrm{Th}(\mathbf{Q}, +, -, 0)$ has quantifier elimination (see [33, Theorem 3.1.9]). As a result of this quantifier elimination, we get that T is ω-stable and also strongly minimal. Therefore, we may appeal to Remark 2.2.7: the pairs $(\mathbf{Q}, +, -, 0, R)$ and $(\mathbf{R}, +, -, 0, R)$ are bounded. Hence, what needs to be done is to show that the induced structures on R by $(\mathbf{Q}, +, -, 0)$ and $(\mathbf{R}, +, -, 0)$ are ω-stable. But by quantifier elimination, we only have to look in both cases at the trace of homogeneous equations with coefficients in $\mathbf{Z}$. By Corollary 3.3.6, we know that in both cases the induced structure is definable in $\mathscr{N} = (\mathbf{N}, S, S^{-1}, 0)$, which has an ω-stable theory (this can be seen by quantifier elimination and counting types). Therefore, $(\mathbf{Q}, +, 0, R)$ and $(\mathbf{R}, +, 0, R)$ are ω-stable, by Theorem 2.1.3. $\square$

Remark 3.8.2. Instead of using Theorem 2.2.6, we could have used [11, Corollary 5.4], which states that every pair $\mathscr{M}_A$, where $\mathscr{M}$ is strongly minimal and $A \subset M$, is bounded. Note that [11, Corollary 5.4] was first proved by A. Pillay in [40, Proposition 3.1]

CHAPTER

4

Expansion of $(\mathbf{Z}, +, 0)$ by a regular sequence: quantifier elimination

In this chapter, we axiomatize, in a language $\mathcal{L} \supset \mathcal{L}_g$, the theory T_R of $\mathscr{Z}_R = (\mathbf{Z}, +, -, 0, R)$, where R is a regular set. We show that T_R has quantifier elimination in $\mathcal{L}$ and has a prime model (and hence T_R is complete). Using this quantifier elimination result, we then prove, by means of counting of types, that T_R is superstable. As a consequence, using a variation of elementary amalgamation [25, Corollary 6.6.2], we deduce that the $\mathcal{L}_R$-theory $\mathrm{Th}(\mathscr{Z}_R)$ is superstable. This is necessary because the $\mathcal{L}$-structure on $\mathscr{Z}$ is not a definitional expansion of $\mathscr{Z}_R$. We then close this chapter with a decidability result.

From now on, we fix a regular set $R \subset \mathbf{N}$ that is enumerated by a regular sequence (r_n). We know then from the previous chapter that

1. (Proposition 3.3.1) for all $Q_1, \ldots, Q_s \in \mathbf{Z}[X]$, there is $k = k(\bar{Q}) \in \mathbf{N}$ and a finite set $E = E_{\bar{Q}} \subset \mathbf{Z}^k$ such that for all $\bar{\ell} \in \mathbf{N}^s$, if $\bar{\ell} \in \mathsf{S}^{\mathrm{nd}}_{\bar{\mathsf{f}},0}$ then for some $\bar{m} \in E$, $l_i = l_1 + m_i$ for all $i \in [k]$, where $\mathsf{f}_i = \mathsf{f}_{Q_i}$ for all $i \in [s]$;
2. (Corollary 3.3.4) for all $Q \in \mathbf{Z}[X]$, either $\mathsf{S}_{\mathsf{f}_Q,0} = \mathbf{N}$ or there is $e = e(Q)$ such that for all $n, m > e$, if $\mathsf{f}_Q(n) = \mathsf{f}_Q(m)$ then $n = m$. Let Triv be the set of $Q \in \mathbf{Z}[X]$ such that $\mathsf{S}_{\mathsf{f}_Q,0} = \mathbf{N}$. Note that $\mathrm{Triv} = \{0\}$ unless θ is algebraic, in which case Triv is the ideal of $\mathbf{Z}[X]$ generated by P_R (see Propositions 1.3.7 and 3.2.2).

Our choice of $\mathcal{L}$ will allow us to express the two properties above in a first-order way. It includes in particular the language $\mathcal{L}_S = \{S, S^{-1}, c\}$. Furthermore, $\mathcal{L}$ will include predicates that will allow to handle formulas such as $\exists x_1, x_2 \in R(x_1 + x_2 = y \wedge D_2(x_1) \wedge D_5(x_2 + 2))$ in a quantifier-free way.

This chapter is organized as follows.

Section 4.1 presents in details the theory T_R and gives the strategy for quantifier elimination: we show that T_R has algebraically prime models (that is, given a substructure $\mathscr{A}$ of a model $\mathscr{M}$ of T_R, we can construct a smallest model of T_R containing $\mathscr{A}$ independently of $\mathscr{M}$) and is 1-e.c., that is existentially closed for existential formulas with at most one existential quantifier.

In Section 4.2, we analyze in models of T_R equations, in the spirit of the work done in Section 3.3. Then in Section 4.3, we give a construction of algebraically prime models. Roughly, given a substructure $\mathscr{A}$ of a model $\mathscr{M}$, the algebraically prime model containing $\mathscr{A}$ will be the divisible closure of the substructure of $\mathscr{M}$ generated by A and certain tuples $\bar{b}$ of elements in $R(M)$ (an instance of such tuples is (b_1, b_2) such that $b_1 + b_2 \in A$). We then proceed in Section 4.4 with a proof that T_R is 1-e.c., concluding the proof of quantifier elimination. Superstability of T_R is then showed in Section 4.5. We conclude this chapter with Section 4.6, where we address decidability of T_R.

4.1 The theory T_R

Let us define the language in which we axiomatize $\mathrm{Th}(\mathscr{Z}_R)$. Recall that $\mathcal{L}_g$ is the language $\{+,-,0,D_n \mid n \in \mathbf{N}^{>1}\}$ and $\mathcal{L}_S$ is the language $\{S, S^{-1}, c\}$. These new symbols are interpreted in $\mathscr{Z}$ as follows: c is interpreted as r_0, for all $n \in \mathbf{N}$, $S(r_n) = r_{n+1}$, $S^{-1}(r_{n+1}) = r_n$, $S^{-1}(r_0) = r_0$ and $S(z) = z = S^{-1}(z)$ for all $z \in \mathbf{Z} \setminus R(\mathbf{Z})$. To each $Q \in \mathbf{Z}[X]$, we let $\mathsf{f} = \mathsf{f}_Q$ be the $\mathcal{L}_g \cup \mathcal{L}_S$-term $\sum_{i=0}^{d} n_i S^i(x)$, where $Q(X) = \sum_{i=0}^{d} n_i X^i$ and $S^0(x) = x$. Notice that such terms are similar to the operators of the previous chapter: in fact a term of the form $\sum_{i=0}^{d} n_i S^i(x)$ composed with the function $n \mapsto r_n$ will be an operator in the sense of Definition 3.2.1. This explains why we decided to keep the same notations and we will also call such terms operators. Furthermore, in this section, the symbol f will always denote an operator.

We now work in $\mathcal{L}_g \cup \{1\} \cup \mathcal{L}_S$, where 1 is a constant symbol that is interpreted in $\mathscr{Z}_R$ by the integer 1. For $n, m \in \mathbf{N}$ we let $\mathbf{Z}[X]^{n\times m}$ be the set of $n \times m$ matrices with entries in $\mathbf{Z}[X]$.

Let $[Q] = (Q_{ij}) \in \mathbf{Z}[X]^{n\times m}$ and let $\varphi_{[Q]}(\bar{x}, \bar{y})$ be the formula

$$\bigwedge_{i\in[n]} \sum_{j\in[m]} \mathsf{f}_{Q_{ij}}(x_j) = y_i.$$

Notice that, working in $\mathscr{Z}_R$, the formula $\exists \bar{x} \in R\ \varphi_{[Q]}(\bar{x}, \bar{y})$ expresses the fact that

$$\bigcap_{i\in[n]} \mathsf{S}_{\bar{\mathsf{f}}_i, y_i} \neq \emptyset.$$

Let $D = \{(P_i, \ell_i, k_i) | i \in [m]\}$ be a (finite) subset of $\mathbf{Z}[X] \times \mathbf{N} \times \mathbf{N}$ such that if $(P, \ell, k) \in D$, then $k < \ell$. We call such a set D a *set of divisibility conditions* and define $\varphi_D(\bar{x})$ as the formula

$$\bigwedge_{i\in[m]} D_{\ell_i}(\mathsf{f}_{P_i}(x_i) + k_i).$$

To $[Q]$ and D as above, we associate an n-ary predicate $\mathrm{Im}_{[Q],D}(\bar{y})$. When D is empty, we write $\mathrm{Im}_{[Q]}$ instead of $\mathrm{Im}_{[Q],D}$. In $\mathscr{Z}_R$, the predicate $\mathrm{Im}_{[Q],D}(\bar{y})$ is interpreted as follows: for all $\bar{m} \in \mathbf{Z}^{|\bar{y}|}$, $\mathrm{Im}_{[Q],D}(\bar{m})$ if and only if $\exists \bar{x} \in R(\varphi_{[Q]}(\bar{x}, \bar{m}) \wedge \varphi_D(\bar{x}))$. So this symbol states that $\bar{m}$ is in the image of sums of operators, and a witness of this fact satisfies certain divisibility conditions.

Finally let $\mathcal{L}$ be the language

$$\mathcal{L}_g \cup \{1\} \cup \mathcal{L}_S \cup \{R\} \cup \{\mathrm{Im}_{[Q],D} \mid [Q] \text{ and } D \text{ as above}\},$$

and we let $\mathscr{Z}_{R,\mathcal{L}}$ be the $\mathcal{L}$-expansion of $\mathscr{Z}_R$ described above.

We fix an axiomatization T_1 of $\mathrm{Th}(\mathbf{Z}, +, -, 0, 1, D_n \mid 1 < n \in \mathbf{N})$ (see [46, Chapter 15, Section 15.1]) and we let T_2 be the following universal axiomatization of $\mathrm{Th}(R, S, S^{-1}, c)$:

$$T_2 = \{\forall x(x \neq c \to S(S^{-1}(x)) = x), \forall x(S^{-1}(S(x)) = x), \forall x(S(x) \neq c), S^{-1}(c) = c\}.$$

We will denote by T_2^R the theory obtained by relativizing to the predicate R the quantifiers appearing in each element of T_2. We will frequently use the fact that, modulo T_1, a formula of the form $\neg D_n(x)$ is equivalent to

$$\bigvee_{k=1}^{n-1} D_n(x+k).$$

Let $\mathscr{M}$ be an $\mathcal{L}$-structure. Let $Q_1, \dots, Q_n \in \mathbf{Z}[X]$. We say that $\bar{a} \in M^n$ is a *non-degenerate solution* of $\sum_{i=1}^n \mathsf{f}_{Q_i}(x_i) = y$ if it is a solution of this equation and no proper sub-sum is equal to 0. This can be expressed by the following first-order formula $\mathsf{S}^{\mathrm{nd}}_{\bar{Q}}(\bar{x}, y)$:

$$\sum_{i=1}^n \mathsf{f}_{Q_i}(x_i) = y \wedge \bigwedge_{I \subsetneq [n]} \sum_{i \in I} \mathsf{f}_{Q_i}(x_i) \neq 0.$$

Let T_R be the following set of axiom schemes.

(Ax.1) T_1;

(Ax.2) T_2^R;

(Ax.3) $c = r_0$ (that is c equals the term $\underbrace{1 + \cdots + 1}_{r_0 \text{ times}}$);

(Ax.4) $\forall x(\neg R(x) \to S(x) = x)$;

(Ax.5) for all $\ell_1, \dots, \ell_n$, all $0 \leq k_i \leq \ell_i$ and $\bar{Q} \in \mathbf{Z}[X]^n$, if

$$\left\{ \bar{z} \in R(\mathbf{Z})^n \;\middle|\; \mathscr{Z}_{\mathcal{L},R} \models \bigwedge_{i \in [n]} D_{\ell_i}(\mathsf{f}_{Q_i}(z_i) + k_i) \right\} = \{\bar{w}_1, \dots, \bar{w}_m\}$$

then we add the axiom

$$\forall \bar{x} \in R \left(\bigwedge_{i\in[n]} D_{\ell_i}(\mathsf{f}_{Q_i}(x_i)+k_i) \to \bigvee_{i\in[m]} \bigwedge_{j\in[n]} x_j = w_{ij} \right);$$

(Ax.6) for all $[Q], D$ as above, we add the axiom

$$\forall \bar{y} \left(\mathrm{Im}_{[Q],D}(\bar{y}) \leftrightarrow \exists \bar{x} \in R(\varphi_D(\bar{x}) \wedge \varphi_{[Q]}(\bar{x},\bar{y})) \right);$$

(Ax.7) for all $Q \in \mathrm{Triv}$, we add the axiom

$$\forall x \in R \ \ \mathsf{f}_Q(x) = 0$$

and for all $Q \notin \mathrm{Triv}$ we add

$$\forall x, y \in R\, (x > e \wedge y > e \wedge x \neq y \to \mathsf{f}_Q(x) \neq \mathsf{f}_Q(y)),$$

where $e = e(Q)$ (see 2 on page 53);

(Ax.8) for every $Q_1, \ldots, Q_s \in \mathbf{Z}[X]$, we add the axiom

$$\forall \bar{x} \in R \left(\mathsf{S}^{\mathrm{nd}}_{\bar{Q}}(\bar{x},0) \to \bigvee_{\bar{m}\in E} \bigwedge_{i\in[s]} x_i = S^{m_i}(x_1) \right),$$

where $E = E_{\bar{Q}}$ (see 1 on page 53).

Note that $\mathscr{Z}_{R,\mathcal{L}}$ is a model of T_R. Indeed, (Ax.7) follows from Corollary 3.3.4 and (Ax.8) follows from Proposition 3.3.1. In particular T_R is consistent. Also note that all axiom schemes but the defining axioms for the congruences D_n and the predicates $\mathrm{Im}_{[Q],D}$ are universal.

The main result of this chapter is the following theorem.

Theorem 4.1.1. *The $\mathcal{L}$-theory T_R has quantifier elimination.*

Notice that the set $R = \{2^n + n \mid n \in \mathbf{N}\}$ does not satisfy (Ax.8). Indeed, consider the term $\mathsf{f}(x) = S^2(x) - 3S(x) + 2x$. For $n \in \mathbf{N}$, we have $\mathsf{f}(2^n + n) = 4 \cdot 2^n + n + 2 - 6 \cdot 2^n - 3n - 3 + 2 \cdot 2^n + 2n = -1$. Thus, for all $n, m \in \mathbf{N}$, $(2^n + n, 2^m + m)$ is a non-degenerate solution of the equation $\mathsf{f}(x_1) - \mathsf{f}(x_2) = 0$. In particular, it has infinitely many non-degenerate solutions. In view of Theorem 4.5.1, which states that T_R is superstable whenever R is regular, this is not surprising because the structure $(\mathbf{Z},+,0,1,R,S)$ is known to be unstable: $\mathbf{N}$ is definable by the formula $\exists y \in R\ y \neq 1 \wedge (2y - S(y) = x)$. However, as we already said in Question 2.2.11, we do not know if there exists a sequence R such that $\mathscr{Z}_R$ is (super)stable and

$(\mathbf{Z}, +, 0, 1, R, S)$ unstable. Nevertheless, Theorem 4.1.1 indicates that working with the successor function is sometimes harmless.

To establish quantifier elimination, we use the following criterion. Given two models $\mathscr{M}_0 \subset \mathscr{M}$ of an arbitrary theory, we say that $\mathscr{M}_0$ is *1-e.c.*[1] in $\mathscr{M}$ if any quantifier-free definable subset of M, defined with parameters in M_0, has a non-empty intersection with M_0.

Proposition 4.1.2 ([33, Corollary 3.1.12])**.** *Let T be an $\mathcal{L}$-theory such that*

1. *(T has algebraically prime models) for all $\mathscr{M} \models T$ and all $\mathscr{A} \subset \mathscr{M}$, there exists a model $\overline{\mathscr{A}}$ of T such that for all $\mathscr{N} \models T$, any embedding $f : \mathscr{A} \to \mathscr{N}$ extends to an embedding $\bar{f} : \overline{\mathscr{A}} \to \mathscr{N}$;*
2. *(T is 1-e.c.) for all $\mathscr{M}_0, \mathscr{M} \models T$, if $\mathscr{M}_0 \subset \mathscr{M}$ then $\mathscr{M}_0$ is 1-e.c. in $\mathscr{M}$.*

Then T has quantifier elimination.

The proof of Theorem 4.1.1 will be a consequence of Proposition 4.1.2 and the work done in the following sections. In Section 4.2, we prove several direct consequences of T_R regarding equations of the form $\mathsf{f}_1(x_1) + \cdots + \mathsf{f}_n(x_n) = a$. In Section 4.3, we give a detailed construction of algebraically prime models of T_R. Finally, we show in Section 4.4 that T_R is 1-e.c.

Corollary 4.1.3. *The $\mathcal{L}$-structure $\mathscr{Z}_{R,\mathcal{L}}$ is a prime model of T_R. In particular T_R is complete.*

Proof. Since $\mathscr{Z}_{R,\mathcal{L}}$ is an algebraically prime model and T_R has quantifier elimination, $\mathscr{Z}_{R,\mathcal{L}}$ is a prime model. Therefore, T_R is complete. □

4.2 Equations in T_R

In order to show that T_R is 1-e.c., we need a good understanding of quantifier-free definable subsets of the domains of models of T_R. As those sets are boolean combinations of equations and Im predicates, we study in this section sets definable by equations and Im predicates in several variables which range in R. We do this because we shall see in our construction of algebraically prime models that elements of a model of T can be written as a sum of elements in R and possibly an element in a small model.

Definition 4.2.1. A operator f is said to be *trivial* if $\mathsf{f} = \mathsf{f}_Q$ for some $Q \in \mathrm{Triv}$.

The addition of S and S^{-1} to our language allows to partition R into orbits.

[1]This abbreviation stands for *existentially closed* for formulas in 1 free variable.

Definition 4.2.2. Let $\mathscr{M} \models T_R$ and $a, b \in R$. The *orbit* of a is the set $\{S^k(a) | k \in \mathbf{Z}\}$ and is denoted by $\mathrm{Orb}(a)$. We say that a and b are in the same orbit if and only if $b \in \mathrm{Orb}(a)$.

The relation "a and b are in the same orbit" is an equivalence relation.

Let us reiterate a comment we made before Proposition 3.3.1. Consider the equation $\sum_{i=1}^n \mathsf{f}_i(x_i) = 0$. Let $\mathscr{M} \models T_R$ and $\bar{b} \in M^n$ be a solution of this equation. Then there exists $(P_1, \ldots, P_\ell)$, a partition of $[n]$, such that for all $i \in [\ell]$, $\bar{b}_{P_i}$ is a non-degenerate solution of $\sum_{i=1}^n \mathsf{f}_i(x_i) = 0$. This is shown by induction on n. If $n = 1$, there is nothing to prove. For $n > 1$, if $\bar{b}$ is non-degenerate, we consider the partition $([n])$. If $\bar{b}$ is degenerate, then there exists $I \subset [n]$ such that $I \neq \emptyset$, $I \neq [n]$,

$$\sum_{i \in I} \mathsf{f}_i(b_i) = 0 \text{ and } \sum_{i \notin I} \mathsf{f}_i(b_i) = 0.$$

Thus we may apply the induction hypothesis to the equations $\sum_{i \in I} \mathsf{f}_i(x_i) = 0$ and $\sum_{i \notin I} \mathsf{f}_i(x_i) = 0$ to conclude. To summarize, we have the following lemma, which states that the set of solutions of an equation can be decomposed as a union of sets of non-degenerate solutions of sub-equations.

Lemma 4.2.3. *Let $\bar{Q} \in \mathbf{Z}[X]^n$. Then*

$$T_R \models \forall y \forall \bar{x} \in R \left(\sum_{i=1}^n \mathsf{f}_i(x_i) = y \leftrightarrow \bigvee_{\bar{I} \in \mathsf{Part}([n])} \mathsf{S}^{\mathrm{nd}}_{\bar{Q}_{I_0}}(\bar{x}_{I_0}, y) \wedge \bigwedge_{j=1}^{|\bar{I}|-1} \mathsf{S}^{\mathrm{nd}}_{\bar{Q}_{I_j}}(\bar{x}_{I_j}, 0) \right). \quad \square$$

As we have already seen in the previous chapter, namely in Proposition 3.3.1, elements of a regular set cannot satisfy an homogeneous equation if they are too far away from each other, in the sense of the function S, unless they form a degenerate solution of this equation. In an arbitrary model of T, the same property holds for elements in different orbits. In the following lemma, we even show that for $n \in \mathbf{N}^{>1}$, a tuple with more than $n/2$ elements in different orbits is *never* a solution of an homogeneous equation with n variables.

Lemma 4.2.4. *Let $\mathscr{M} \models T_R$. Let $\bar{\mathsf{f}}$ be a n-tuple of non-trivial operators, $n > 1$, and let $b_1, \ldots, b_k \in R$, $1 < k \leq n$, be in different orbits.*

1. *If $k > n/2$, then for all $c_{k+1}, \ldots, c_n \in R$,*

$$\sum_{i=1}^k \mathsf{f}_i(b_i) + \sum_{i=k+1}^n \mathsf{f}_i(c_i) \neq 0;$$

2. *If $k \leq n/2$, then for all $c_{k+1}, \ldots, c_n \in R$, the elements $b_1, \ldots, b_k, c_{k+1}, \ldots, c_n$ do not form a non-degenerate solution of the equation $\sum_{i=1}^n \mathsf{f}_i(x_i) = 0$. Moreover, if*

$\sum_{i=1}^{k} \mathsf{f}_i(b_i) + \sum_{i=k+1}^{n} \mathsf{f}_i(c_i) = 0$, *then for all* $i \in [k]$ *there exists a non-empty* $P_i \subset \{k+1,\ldots,n\}$ *such that* $P_i \cap P_{i'} = \emptyset$ *for all* $i \neq i' \in [k]$ *and for all* $i \in [k]$ $(b_i, c_j \mid j \in P_i)$ *is a non-degenerate solution of*

$$\mathsf{f}_i(x_i) + \sum_{j\in P_i} \mathsf{f}_j(x_j) = 0.$$

Proof. Let $c_{k+1},\ldots,c_n \in R$. By (Ax.8), $b_1,\ldots,b_k,c_{k+1},\ldots,c_n$ cannot be a non-degenerate solution of

$$\sum_{i=1}^{n} \mathsf{f}_i(x_i) = 0,$$

since, for instance, b_1 is not in the same orbit as b_2. Suppose $b_1,\ldots,b_k,c_{k+1},\ldots,c_n$ is a degenerate solution of

$$\sum_{i=1}^{n} \mathsf{f}_i(x_i) = 0.$$

Then there exists a partition $(P_1,\ldots,P_\ell)$ of $[n]$ such that for all $j \in [\ell]$ the tuple $(b_i \mid i \in P_j \cap [k]), (c_i \mid i \in P_j \cap \{k+1,\ldots,n\})$ is a non-degenerate solution of

$$\sum_{i\in P_j\cap[k]} \mathsf{f}_i(x_i) + \sum_{i\in P_j\cap\{k+1,\ldots,n\}} \mathsf{f}_i(x_i) = 0.$$

Since $b_1,\ldots,b_k$ are in different orbits, we must have, by (Ax.8), $|P_j \cap [k]| \leq 1$ for all $j \in [\ell]$. Also, since all operators involved are non-trivial, we must have $|P_j \cap \{k+1,\ldots,n\}| > 0$ for all $j \in [\ell]$. This implies in particular that $k \leq n/2$ and finishes the proof of the lemma. □

We now show that (Ax.8) is true for non-homogeneous equations.

Proposition 4.2.5. *Let* $\mathscr{M} \models T_R$, $\bar{Q} \in \mathbf{Z}[X]^n$ *and* $a \in M$, $a \neq 0$. *Then there exist* $\bar{b}_1,\ldots,\bar{b}_k \in R$ *such that*

$$\mathscr{M} \models \forall \bar{x} \in R \left(\mathsf{S}^{\mathrm{nd}}_{\bar{Q}}(\bar{x},a) \rightarrow \bigvee_{j=1}^{k} \bigwedge_{i=1}^{n} x_i = b_{ji} \right).$$

Proof. Let $\mathsf{f}_i = \mathsf{f}_{Q_i}$ for all $i \in [n]$. Assume there exist infinitely many distinct non-degenerate solutions $\bar{b}_i \in M^n$, $i \in \mathbf{N}$, of the equation

$$\mathsf{f}_1(x_1) + \cdots + \mathsf{f}_n(x_n) = a.$$

We will reach a contradiction using (Ax.8) applied to the equation

$$\sum_{i=1}^{n} \mathsf{f}_i(x_i) - \sum_{i=n+1}^{2n} \mathsf{f}_{i-n}(x_i) = 0,$$

which we denote by $\varphi(\bar{x})$.

We have that for all $i \in \mathbf{N}$, the tuple $(\bar{b}_0, \bar{b}_i)$ is a solution of $\varphi(\bar{x})$. We may assume that there exists a partition $I = (I_1, \ldots, I_\ell)$ of $[2n]$ such that for all $i \in \mathbf{N}$ and all $j \in [\ell]$, $(b_{0k} \mid k \in I_j \cap [n])$, $(b_{ik} \mid k+n \in I_j \setminus [n])$ is a non-degenerate solution of the equation

$$\sum_{i \in I_j \cap [n]} \mathsf{f}_i(x_i) - \sum_{i \in I_j \setminus [n]} \mathsf{f}_{i-n}(x_i) = 0.$$

By non-degeneracy and the fact that $a \neq 0$, we have for all $j \in [\ell]$ that $I_j \cap [n] \neq \emptyset$ and $I_j \setminus [n] \neq \emptyset$.

By (Ax.8) for all $j \in [\ell]$, there is a finite set $E_j \subset \mathbf{Z}^{|I_j|}$ such that for all $i \in \mathbf{N}$

$$\bigvee_{\bar{m} \in E_j} \left(\bigwedge_{k \in I_j \cap [n]} b_{0k_0} = S^{m_k}(b_{0k}) \wedge \bigwedge_{k \in I_j \setminus [n]} b_{0k_0} = S^{m_k}(b_{ik}) \right),$$

where $k_0 = \min I_j$.

But this is a contradiction since the set defined by the formula

$$\bigvee_{\bar{m} \in E_j} \left(\bigwedge_{k \in I_j \cap [n]} b_{0k_0} = S^{m_k}(b_{0k}) \wedge \bigwedge_{k \in I_j \setminus [n]} b_{0k_0} = S^{m_k}(x_k) \right)$$

is finite for all $j \in [\ell]$. □

As a corollary, by compactness, for a fixed model of T_R, we obtain a uniform bound on the number of non-degenerate solutions of a given equation.

Corollary 4.2.6. *Let $\bar{Q} \in (\mathbf{Z}[X] \setminus \mathrm{Triv})^n$ and $\mathscr{M} \models T_R$. Then there exist $k \in \mathbf{N}$ such that*

$$\mathscr{M} \models \forall y \left(y \neq 0 \to \exists \bar{x}_1, \ldots, \bar{x}_k \in R \, \forall \bar{z} \in R \left(\mathsf{S}^{\mathrm{nd}}_{\bar{Q}}(\bar{z}, y) \to \bigvee_{j=1}^{k} \bigwedge_{i=1}^{n} z_i = x_{ji} \right) \right).$$

In order to show that elements of a model of T_R are sums of elements in R and possibly an element in a smaller model, we will need to know when such a sum is again a element in R. The next proposition states that this cannot happen for sums of elements in different orbits.

Proposition 4.2.7. *Let $\mathscr{M}, \mathscr{M}_0 \models T_R$ such that $\mathscr{M}_0 \subset \mathscr{M}$ and $n \in \mathbf{N}^{>1}$. Let $\bar{\mathsf{f}}$ be a tuple of n non-trivial operators, $b_1, \ldots, b_n \in R(M) \setminus R(M_0)$ in different orbits and $a \in M_0$. Then*

$$\sum_{i=1}^{n} \mathsf{f}_i(b_i) + a \notin R(M).$$

Proof. Suppose, towards a contradiction, that $\sum_{i=1}^{n} \mathsf{f}_i(b_i) + a = b_{n+1} \in R(M)$. Assume that $\mathsf{f}_i = \mathsf{f}_{Q_i}$, where $Q_i \in \mathbf{Z}[X]$. Consider the formula $\mathrm{Im}_{[Q]}(y)$, where $[Q] = (Q_1, \ldots, Q_n, -1)$. Then we have $\mathscr{M} \models \mathrm{Im}_{[Q]}(-a)$. As $\mathscr{M}_0 \subset \mathscr{M}$, $\mathscr{M}_0 \models T_R$ and $a \in M_0$, we also have $\mathscr{M}_0 \models \mathrm{Im}_{[Q]}(-a)$. By (Ax.6), there exist $b_{n+2}, \ldots, b_{2n+2} \in R(M_0)$ such that $\sum_{i=1}^{n} \mathsf{f}_i(b_i) - b_{n+1} = \sum_{i=1}^{n} \mathsf{f}_i(b_{n+1+i}) - b_{2n+2} = -a$. By Lemma 4.2.4 applied to

$$\sum_{i=1}^{n} \mathsf{f}_i(x_i) - x_{n+1} - \sum_{i=1}^{n} \mathsf{f}_i(x_{n+1+i}) + x_{2n+2} = 0,$$

and $b_1, \ldots, b_n$, for all $i \in [n]$, there exists $J_i \subset \{n+1, \ldots, 2n+2\}$ such that $b_i, (b_j)_{j \in J_i}$ is a non-degenerate solution to the corresponding equation, namely

$$\mathsf{f}_i(b_i) - \sum_{j \in J_i} \mathsf{f}_j(b_j) = 0,$$

where f_{2n+2} is the operator associated to the constant polynomial -1. We furthermore have $J_i \cap J_{i'} = \emptyset$ for all $i \neq i' \in [n]$ and $J_i \neq \emptyset$ for all $i \in [n]$. For all $i \in [n]$ we have in particular, in virtue of (Ax.8), that for all $j \in J_i$, b_j is in the orbit of b_i. This enforces that for all $i \in [n]$, for all $j \in J_i$, $b_j \notin M_0$ since $b_i \notin M_0$. As a result $J_i = \{n+1\}$ for all $i \in [n]$. But then, by (Ax.8), b_1 and b_2 are in the same orbit, a contradiction. □

We apply our understanding of equations satisfied by elements of R to understand sets defined by Im predicates. Let $\mathscr{M} \models T$. Let $[Q] \in \mathbf{Z}[X]^{n \times m}$, $[Q'] \in \mathbf{Z}[X]^{n \times m'}$ and D a set of divisibility conditions of size m. Let $\bar{a} \in M^n$. We want to understand the set N defined by

$$\mathrm{Im}_{[Q],D}\left(a_1 + \sum_{i=1}^{m'} \mathsf{f}_{Q'_{1i}}(x_i), \ldots, a_n + \sum_{i=1}^{m'} \mathsf{f}_{Q'_{ni}}(x_i)\right). \tag{4.1}$$

This will be needed to show that T_R is 1-e.c, in order to reduce the complexity of formulas. In particular, we shall explain how we can separate the variables from the constants in (4.1).

We want to show that the formula (4.1) expresses two things for a tuple $\bar{b} \in R(M)^{m'}$:

1. there exists $J' \subset [m']$ such that $\bar{b}_{J'}$ belongs to a finite set depending only on $[Q], [Q'], D$ and $\bar{a}$;
2. for all $J \subset [m']$, $\bar{b}_J$ satisfy a finite number of recurrence relations and congruence relations again depending only on $[Q], [Q'], D$ and $\bar{a}$. Also, when $J_1, J_2 \subset [m]$ have a non-empty intersection, the above conditions on $\bar{b}_{J_1}$ and $\bar{b}_{J_2}$ must be consistent.

To do so, for all $J_{01},\ldots,J_{0n} \subset [m]$ and $J_{11},\ldots,J_{1n} \subset [m']$, we let $\mathsf{S}^{\mathrm{nd}}_{[Q],[Q'],\bar{J}_0,\bar{J}_1}(\bar{a})$ be the set defined by the formula

$$\exists \bar{z} \in R\, \varphi_D(\bar{z}) \wedge \bigwedge_{i\in[n]} \left(\sum_{j\in J_{0i}} \mathsf{f}_{Q_{ij}}(z_j) = a_i + \sum_{j\in J_{1i}} \mathsf{f}_{Q'_{ij}}(x_i) \wedge \bar{z}_{J_{0i}}\bar{x}_{J_{1i}} \text{ is non-degenerate} \right).$$

Notice that by Proposition 4.2.5, the set $\mathsf{S}^{\mathrm{nd}}_{[Q],[Q'],\bar{J}_0,\bar{J}_1}(\bar{a})$ is finite if $a_i \neq 0$ for some $i \in [n]$.

Recall that by (Ax.6), the formula (4.1) is satisfied by some $\bar{b} \in R(M)^{m'}$ if and only if there is $\bar{z} \in R(M)^m$ such that the following system of equations and congruence relations is satisfied:

$$\begin{cases} \mathsf{f}_{Q_{11}}(z_1) + \ldots + \mathsf{f}_{Q_{1m}}(z_m) = a_1 + \mathsf{f}_{Q'_{11}}(b_1) + \cdots + \mathsf{f}_{Q_{1m'}}(b_{m'}) \\ \vdots \\ \mathsf{f}_{Q_{n1}}(z_1) + \ldots + \mathsf{f}_{Q_{nm}}(z_m) = a_n + \mathsf{f}_{Q'_{n1}}(b_1) + \cdots + \mathsf{f}_{Q_{nm'}}(b_{m'}) \\ D_{\ell_1}(\mathsf{f}_{P_1}(z_1) + k_1), \ldots, D_{\ell_m}(\mathsf{f}_{P_m}(z_m) + k_m). \end{cases}$$

Now, for each $i \in [n]$, choose, according to Lemma 4.2.3, $\bar{J}_i = (J_{i0},\ldots,J_{i\ell_1}) \in \mathsf{Part}([m])$, $\bar{J}'_i = (J'_{i0},\ldots,J'_{i\ell_2}) \in \mathsf{Part}([m'])$ and $K_i \in [\ell_1]\times[\ell_2]$ such that for all $i \in [n]$:

1. $(\bar{z}_{J_{i0}}, \bar{b}_{J'_{i0}})$ is a non-degenerate solution of
$$\sum_{j\in J_{i0}} \mathsf{f}_{Q_{ij}}(y_j) = a_i + \sum_{j\in J'_{i0}} \mathsf{f}_{Q'_{ij}}(x_j);$$
2. for all $(s_1,s_2) \in K_i$, $(\bar{z}_{J_{is_1}}, \bar{b}_{J'_{is_2}})$ is a non-degenerate solution of
$$\sum_{j\in J_{is_1}} \mathsf{f}_{Q_{ij}}(y_j) = \sum_{j\in J'_{is_2}} \mathsf{f}_{Q'_{ij}}(x_j);$$
3. for all $(s_1,s_2) \in [\ell_1]\times[\ell_2]$ such that s_1 (resp. s_2) does not appear in the first (reps. second) coordinate of an element of K_i, $\bar{z}_{J_{is_1}}$ and $\bar{b}_{J'_{is_2}}$ are respectively non-degenerate solutions of
$$\sum_{j\in J_{is_1}} \mathsf{f}_{Q_{ij}}(y_j) = 0 \text{ and } 0 = \sum_{j\in J'_{is_2}} \mathsf{f}_{Q'_{ij}}(x_j).$$

This decomposition of each equation in the system shows that $\bar{b}_{J'}$ is in $\mathsf{S}^{\mathrm{nd}}_{[Q],[Q'],\bar{J}_0,\bar{J}_1}(\bar{a})$, where $J' = \bigcup_{i\in[n]} J'_{i0}$. For the homogeneous equations above, we may apply (Ax.8) to obtain the desired relations between $\bar{z}_{J_{is_1}}$ and $\bar{b}_{J'_{is_2}}$ whenever $(s_1,s_2) \in K_i$.

To summarize, we state the following corollary, which is an explicit statement of the above discussion.

Corollary 4.2.8. *Let $[Q] \in \mathbf{Z}[X]^{n\times m}$ and D a set of divisibility of size m. Let $[Q'] \in \mathbf{Z}[X]^{n\times m'}$. Let $\mathscr{M} \models T_R$ and $\bar{a} \in M^n$. Then for all $\bar{b} \in R(M)^{m'}$*

$$\mathscr{M} \models \mathrm{Im}_{[Q],D}\left(a_1 + \sum_{i=1}^{m'} \mathsf{f}_{Q'_{1i}}(b_i), \ldots, a_n + \sum_{i=1}^{m'} \mathsf{f}_{Q'_{ni}}(b_i)\right)$$

if and only if for all $i \in [n]$ there are $(J_{i0}, \ldots, J_{is}) \in \mathsf{Part}([m])$ and $(J'_{i0}, \ldots, J'_{is'}) \in \mathsf{Part}([m'])$ and $K_i \subset [s] \times [s']$ such that for all $s_1 \in [s]$ there is at most one $s_2 \in [s']$ such that $(s_1, s_2) \in K_i$ and

$$\mathscr{M} \models \bar{b}_{J'} \in \mathsf{S}^{\mathrm{nd}}_{[Q],[Q'],\bar{J}_0,\bar{J}'_0}(\bar{a}) \tag{4.2}$$

$$\wedge \bigwedge_{i\in[n]} \bigwedge_{(s_1,s_2)\in K_i} \left(\sum_{j\in J_{is_1}} \mathsf{f}_{Q_{ij}}(S^{k_{ij}}(b_{j^*})) = \sum_{j\in J'_{is_2}} \mathsf{f}_{Q'_{ij}}(S^{k'_{ij}}(b_{j^*})) \right. \tag{4.3}$$

$$\left. \wedge \bigwedge_{j\in J'_{is_2}} b_j = S^{k'_{ij}}(b_{j^*}) \bigwedge_{j\in J_{is_1}} D_{\ell_j}(\mathsf{f}_{P_j}(S^{k_{ij}}(b_{j^*})) + k'_j) \right) \tag{4.4}$$

$$\wedge \bigwedge_{i\in[n]} \bigwedge_{(s_1,s_2)\in \tilde{K}_i} \left(0 = \sum_{j\in J'_{is_2}} \mathsf{f}_{Q'_{ij}}(S^{k_{ij}}(b_{j^*})) \wedge \bigwedge_{j\in J'_{is_2}} b_j = S^{k_{ij}}(b_{j^*}) \right) \tag{4.5}$$

$$\wedge \bigwedge_{i\in[n]} \bigwedge_{(s_1,i',s'_1,s'_2)\in K'_i} \left(\sum_{j\in J_{is_1}} \mathsf{f}_{Q_{ij}}(S^{k_{ij}}(b_{j^*})) = 0 \wedge \bigwedge_{j\in J_{is_1}} D_{\ell_j}(\mathsf{f}_{P_j}(S^{k_{ij}}(b_{j^*})) + k'_j) \right) \tag{4.6}$$

$$\wedge \mathrm{Im}_{[\tilde{Q}],D}(0), \tag{4.7}$$

where $J' = \cup_{i\in[n]} J'_{i0}$, for all $i \in [n]$, $j^* = \min J'_{is_2}$ and

1. *if $(s_1, s_2) \in K_i$, $\bar{k}_i$ and $\bar{k}'_i$ are given by (Ax.8) applied to the operators in (4.3);*
2. *$(s_1, s_2) \in \tilde{K}_i$ if and only if s_1 (resp. s_2) does not appear in the first (reps. second) coordinate of an element of K_i. In this case, $\bar{k}_i$ is given by (Ax.8) applied to the operators in (4.5),*
3. *$(s_1, i', s'_1, s'_2) \in K'_i$ if and only if $(s_1, s_2) \notin K_i$ for all $s_2 \in [n]$, $J_{is_1} \cap J_{i's'_1} \neq \emptyset$, $(s'_1, s'_2) \in K_i$ and $j^* = \min J_{i's'_2}$. In this case, $\bar{k}_i$ is given by (Ax.8) applied to the operators in (4.6),*

and $[\tilde{Q}]$ is the matrix defined by $\tilde{Q}_{ij} = Q_{ij}$ if Q_{ij} does not appear in (4.2)-(4.6) and $\tilde{Q}_{ij} = 0$ otherwise. □

4.3 T_R has algebraically prime models

Let $\mathscr{M} \models T_R$ and $\mathscr{A} \subset \mathscr{M}$. For $X \subset M$, we let $\mathrm{div}(X)$ be the *divisible closure of X in $\mathscr{M}$*, that is the substructure generated by $\{d \mid nd \in X \text{ for some } n \in \mathbf{N}^{>0}\}$. (When X

is the domain of an $\mathcal{L}$-substructure, $\mathrm{div}(X)$ is the divisible closure of X in the group theoretic sense.) The construction of the algebraically prime model over $\mathscr{A}$, denoted $\overline{\mathscr{A}}$, is done as follows. Let $\bar{\mathsf{f}}$ be a n-tuple of non-trivial operators. Call an n-tuple $\bar{b} \in R(M)$ $\bar{\mathsf{f}}$-*good* if

1. $b_i \notin A$ for all $i \in [n]$;
2. $\mathsf{f}_1(b_1) + \cdots + \mathsf{f}_n(b_n) \in \mathscr{A}$;
3. $b_i \notin \mathrm{Orb}(b_j)$ whenever $j \neq i$.

Let $\tilde{\mathscr{A}}$ be the substructure generated by $\mathscr{A}$ and $\bar{\mathsf{f}}$-good tuples of elements of $R(M)$, for all tuples $\bar{\mathsf{f}}$ of non-trivial operators. This structure will satisfy all axioms of T_R except the definition of the symbols D_n. So our algebraically prime model over $\mathscr{A}$ will be $\overline{\mathscr{A}} = \mathrm{div}(\tilde{\mathscr{A}})$.

Lemma 4.3.1. *$\overline{\mathscr{A}}$ is a model of T_R.*

Proof. We begin with a description of elements in $\tilde{\mathscr{A}}$. Assume $\tilde{\mathscr{A}} = \langle A, (b_\lambda)_{\lambda<\kappa} \rangle$, where $b_\lambda \notin \mathrm{Orb}(b_{\lambda'})$ for all $\lambda \neq \lambda'$ and each b_λ appears in a good tuple. We want to show that any $d \in \tilde{\mathscr{A}}$ can be put in the form $a + \sum_{i=1}^n \mathsf{f}_i(b_{\lambda_i})$, where $\lambda_i \neq \lambda_j$ for all $i \neq j \in [n]$ and $a \in A$. Let $t(\bar{x}, y)$ be the term $y + \sum_{i=1}^n \mathsf{f}_i(x_i)$. We show that for all $a \in A$ and $b_{\lambda_1}, \ldots, b_{\lambda_n}$ in different orbits, either $S(t(\bar{b}, a)) = t(\bar{b}, a)$ or $t(\bar{b}, a) = S^k(b_\lambda)$ for some $\lambda < \kappa$ and $k \in \mathbf{Z}$. Assume $S(t(\bar{b}, a)) \neq t(\bar{b}, a)$. This implies that $b = t(\bar{b}, a) \in R(M)$. Then, since $a = b - \sum_{i=1}^n \mathsf{f}_i(b_{\lambda_i})$, either b is in the orbit of b_{λ_i} for some $i \in [n]$ or $(b, \bar{b})$ is an $(x, -\bar{\mathsf{f}})$-good tuple. This shows that $b = S^k(b_\lambda)$ for some $\lambda < \kappa$ and $k \in \mathbf{Z}$. Thus every element $\tilde{\mathscr{A}}$ is of the form $a + \sum_{i=1}^n \mathsf{f}_i(b_{\lambda_i})$.

We now do the same job for elements in $\overline{\mathscr{A}}$.

Claim 4.3.2. Let $d \in \overline{\mathscr{A}}$. Then there exist $a \in \tilde{\mathscr{A}}$ and $n \in \mathbf{N}^{>0}$ such that $nd = a$.

Proof. Let X be the set $\{d \mid nd \in \tilde{A}$ for some $n \in \mathbf{N}^{>0}\}$. We first notice that for all $\bar{d} \in X^k$ and $\bar{m} \in \mathbf{Z}^k$, there is $n \in \mathbf{N}$ such that $n(m_1 d_1 + \cdots + m_k d_k) \in \tilde{A}$ (just take n to be the product of the witnesses of the fact that $\bar{d} \in X^k$). So to conclude, it is enough to show that for all terms $t(\bar{x})$, $|x| = k$, and all $\bar{d} \in X^k$ either $t(\bar{d}) \in \tilde{A}$ or there is $\bar{m} \in \mathbf{Z}^k$ such that $t(\bar{d}) - (m_1 d_1 + \cdots + m_k d_k) \in \tilde{A}$. This is done by induction on the complexity of terms (the complexity being here the number of occurrences of the symbols S and S^{-1}).

Let $t(\bar{x})$ be a term, $|x| = k$. Assume that $t(\bar{x}) = t_1(\bar{x}) + S^\epsilon(t_2(\bar{x}))$, where $\epsilon \in \{-1, 1\}$ and t_1 and t_2 are terms such that for all $d \in X^k$, either $t_i(\bar{d}) \in \tilde{A}$ or there is $\bar{m} \in \mathbf{Z}^k$ such that $t_i(\bar{d}) - (m_1 d_1 + \cdots + m_k d_k) \in \tilde{A}$, $i \in [2]$. Let $\bar{d} \in X^k$ and assume that $t(\bar{d}) \notin \tilde{A}$. Then we have at least one of $t_1(\bar{d})$ or $S^\epsilon(t_2(\bar{d}))$ not in $\tilde{A}$. Let us show that if $S^\epsilon(t_2(\bar{d})) \notin \tilde{A}$, then $S^\epsilon(t_2(\bar{d})) = t_2(\bar{d})$. Indeed, since $\bar{d} \in X^k$, there exists $n \in \mathbf{N}$ such that $nt_2(\bar{d}) \in \tilde{A}$. Thus, if $S^\epsilon(t_2(\bar{d})) = t_2(\bar{d})$ is not true, then $t_2(\bar{d}) \in R(M)$

and this imply that $t_2(\bar{d})$ is (nx)-good, in contradiction with our assumption that $t_2(\bar{d}) \notin \tilde{A}$. Therefore, if $S^{\epsilon}(t_2(\bar{d})) \notin \tilde{A}$, $t(\bar{d}) = t_1(\bar{d}) + t_2(\bar{d})$, and the result follows by induction. If $S^{\epsilon}(t_2(\bar{d})) \in \tilde{A}$, then $t_1(\bar{d}) \notin \tilde{A}$. Then, by induction, there is $\bar{m} \in \mathbf{Z}^k$ such that $t_1(\bar{d}) - (m_1 d_1 + \cdots + m_k d_k) \in \tilde{A}$. But then $t(\bar{d}) - (m_1 d_1 + \cdots + m_k d_k) = t_1(\bar{d}) - (m_1 d_1 + \cdots + m_k d_k) + S^{\epsilon}(t_2(\bar{d})) \in \tilde{A}$. □

Let us finally show that $\overline{\mathscr{A}} \models T_R$. The only axiom scheme that requires details is (Ax.6) – the defining axioms for the divisibility predicates are true since we took the divisible closure of $\tilde{\mathscr{A}}$ and the others are universal and thus true in any substructure. So assume that $\overline{\mathscr{A}} \models \mathrm{Im}_{[Q],D}(d_1, \ldots, d_n)$, where $[Q] \in (\mathbf{Z}[X])^{n \times m}$ and D is a set of divisibility conditions of size m. By Claim 4.3.2,there exists $m \in \mathbf{N}^{>0}$ such that $md_1, \ldots, md_n \in \tilde{A}$: for all $i \in [n]$, $md_i = a_i + \sum_{j=1}^{k} \mathsf{f}'_{ij}(b_{\lambda_{ij}})$, for some $a_i \in A$, some tuple $\bar{\mathsf{f}}'_i$ of operators and $\lambda_{i1}, \ldots, \lambda_{ik} < \kappa$. Since $\mathscr{M} \models T_R$ and we can find $b_1, \ldots, b_m \in R(M)$ such that

$$\mathscr{M} \models \bigwedge_{i \in [n]} \sum_{j \in [m]} \mathsf{f}_{Q_{ij}}(b_j) = d_i \wedge \bigwedge_{i \in [m]} D_{\ell_i}(\mathsf{f}_{Q_i}(b_i) + k_i).$$

This is equivalent to

$$\mathscr{M} \models \bigwedge_{i \in [n]} \sum_{j \in [m]} m\mathsf{f}_{Q_{ij}}(b_j) = md_i \wedge \bigwedge_{i \in [m]} D_{\ell_i}(\mathsf{f}_{Q_i}(b_i) + k_i).$$

We may assume that for all $j \in [m]$ there exists $i \in [n]$ such that $\mathsf{f}_{Q_{ij}}$ is non-trivial. Indeed if for some $j \in [m]$ the operator $\mathsf{f}_{Q_{ij}}$ is trivial for all $i \in [n]$, we may replace, by (Ax.5) and (Ax.7), b_j by any $b'_i \in R(\mathbf{Z})$ such that $\mathscr{Z}_R \models D_{\ell_j}(\mathsf{f}_j(b'_j) + k_j)$, which is possible since $\mathscr{Z}_R \subset \mathscr{A}$ and $\mathscr{Z}_R \models T_R$. Now let $\ell \in [m]$. Let $i \in [n]$ such that $\mathsf{f}_{Q_{i\ell}}$ in non-trivial. Then, we have

$$\sum_{j \in [m]} m\mathsf{f}_{Q_{ij}}(b_j) = a_i + \sum_{j \in [k]} \mathsf{f}'_{ij}(b_{\lambda_{ij}}).$$

Thus, there exists $J \in [m] \setminus \{\ell\}$ and $J' \subset [k]$ such that

$$m\mathsf{f}_{\ell}(b_{\ell}) + \sum_{j \in J} m\mathsf{f}_{Q_{ij}}(b_j) = a_i + \sum_{j \in J'} \mathsf{f}'_{ij}(b_{\lambda_{ij}})$$

or

$$m\mathsf{f}_{\ell}(b_{\ell}) + \sum_{j \in J} m\mathsf{f}_{Q_{ij}}(b_j) = \sum_{j \in J'} \mathsf{f}'_{ij}(b_{\lambda_{ij}}),$$

in a non-degenerate way. So:

1. if $J' \neq \emptyset$ and there exists $j \in J'$ such that $\mathsf{f}'_{Q_{ij}}$ is non-trivial, then b_{ℓ} is in the orbit of $b_{\lambda_{ij}}$. Hence $b_{\ell} \in \overline{A}$;

2. if $J' = \emptyset$ or $\mathsf{f}'_{Q_{ij}}$ is trivial for all $j \in J'$, then we have

$$m\mathsf{f}_\ell(b_\ell) + \sum_{j\in J} m\mathsf{f}_{Q_{ij}}(b_j) \in \{0, a_i\}.$$

In particular b_ℓ appears in a good tuple. In particular, $b_\ell \in \overline{A}$.

Thus $\overline{\mathscr{A}}$ satisfies (Ax.8). □

Let us show that any embedding $f : \mathscr{A} \to \mathscr{N}$ extends to an embedding $\bar{f} : \overline{\mathscr{A}} \to \mathscr{N}$.

Lemma 4.3.3. *Let $f : \mathscr{A} \to \mathscr{N}$ be an $\mathcal{L}$-embedding. Then f extends to an $\mathcal{L}$-embedding $\bar{f} : \overline{\mathscr{A}} \to \mathscr{N}$.*

Proof. Let $\mathcal{L}_0$ be the language $\{+,-,0,1,R\} \cup \mathcal{L}_S$. We first extend f to an $\mathcal{L}_0$-embedding $\tilde{f} : \tilde{\mathscr{A}} \to \mathscr{N}$. Let q be the partial type

$$\begin{aligned}
&\{\mathsf{f}_1(x_{\lambda_1}) + \cdots + \mathsf{f}_n(x_{\lambda_n}) = f(a) \mid \mathscr{M} \models \mathsf{f}_1(b_{\lambda_1}) + \cdots + \mathsf{f}_n(b_{\lambda_n}) = a, a \in A, \bar{\mathsf{f}} \text{ non-trivial}\}\\
&\cup\{x_\lambda \neq f(a) \mid \lambda < \kappa, a \in A\}\\
&\cup\{S^k(x_{\lambda_1}) \neq x_{\lambda_2} \mid \lambda_1 \neq \lambda_2, k \in \mathbf{Z}\}\\
&\cup\{D_\ell(\mathsf{f}(x_\lambda) + k) \mid \mathscr{M} \models D_\ell(\mathsf{f}(b_\lambda) + k), 0 \leq k < \ell \in \mathbf{N}\}.
\end{aligned}$$

We claim that q is finitely consistent in $\mathscr{N}$. Let Δ be a finite part of q. We may assume the conjunction of the formulas in Δ is of the form

$$\bigwedge_{i\in I_1} \mathsf{f}_{i1}(x_{\lambda_1}) + \cdots + \mathsf{f}_{in}(x_{\lambda_n}) = f(a_i) \wedge \bigwedge_{i\in[n]} D_{\ell_i}(\mathsf{f}_i(x_{\lambda_i}) + k_i)$$
$$\wedge \bigwedge_{i\in I_2; j\in[n]} x_{\lambda_j} \neq f(a_{ij}) \wedge \bigwedge_{i,j\in[n]; k\in I_3} S^k(x_{\lambda_i}) \neq x_{\lambda_j},$$

where I_1, I_2 and I_3 are non-empty.

By (Ax.6), there exists $\bar{b}' \in R(N)^n$ such that

$$\bigwedge_{i\in I_1} \mathsf{f}_{i1}(b'_1) + \cdots + \mathsf{f}_{in}(b'_n) = f(a_i) \wedge \bigwedge_{i\in[n]} D_{\ell_i}(\mathsf{f}_i(b'_i) + k_i).$$

Assume towards a contradiction that $\bar{b}'$ is not a realization of Δ. Then we have that, for some $i_1 \in I_2$, $j_1, i_2, j_2 \in [n]$ and $k \in I_3$,

$$b'_{j_1} = f(a_{i_1 j_1}) \vee S^k(b'_{i_2}) = b'_{j_2}.$$

So, again using (Ax.6), we can find $\bar{b}'' \in R(M)^n$ such that

$$\bigwedge_{i\in I_1} \mathsf{f}_{i1}(b''_1) + \cdots + \mathsf{f}_{in}(b''_n) = a_i \wedge \bigwedge_{i\in[n]} D_{\ell_i}(\mathsf{f}_i(b''_i) + k_i) \wedge (b''_{j_1} = a_{i_1 j_1} \vee S^k(b''_{i_2}) = b''_{j_2}).$$

Let us show that this contradicts the fact that $b_{\lambda_1}, \ldots, b_{\lambda_n}$ is a good tuple:

1. assume first that $b''_{j_1} = a_{i_1 j_1}$. Let $i \in I_1$. Since $\mathsf{f}_{i1}(b''_1) + \cdots + \mathsf{f}_{in}(b''_n) = \mathsf{f}_{i1}(b_{\lambda_1}) + \cdots + \mathsf{f}_{in}(b_{\lambda_n})$ and $b_{\lambda_i} \notin \mathrm{Orb}(b_{\lambda_j})$ for all $i \neq j \in [n]$, we have by Lemma 4.2.4 and (Ax.8) that $b_{\lambda_i} \in \mathrm{Orb}(b''_{j_1})$ for some $i \in [n]$. Thus, $b_{\lambda_i} \in A$ since $\mathscr{A}$ is closed under S and S^{-1}. This contradicts the first clause in the definition of a good tuple;
2. second, assume that $S^k(b''_{i_2}) = b''_{j_2}$. As in the previous case, we get that $b_{\lambda_{i_1}} \in \mathrm{Orb}(b_{\lambda_{j_2}})$, contradicting the third clause in the definition of a good tuple.

Hence $\bar{b}'$ is a realization of Δ. Thus q is finitely consistent in $\mathscr{N}$ and so realized in an elementary extension $\mathscr{N}^*$ of $\mathscr{N}$ by some $(b'_\lambda)_{\lambda<\kappa}$. Let us show that $(b'_\lambda)_{\lambda<\kappa}$ is in $\mathscr{N}$. Let $\lambda < \kappa$. By definition, b_λ appears in a $\bar{\mathsf{f}}$-good tuple: there exist $b_{\lambda_2}, \ldots, b_{\lambda_n} \in R(M)\backslash A$ and $a \in A$ such that $\mathsf{f}_1(b_\lambda) + \mathsf{f}_2(b_{\lambda_2}) + \cdots + \mathsf{f}_n(b_{\lambda_n}) = a$. The same holds for $b'_\lambda, b'_{\lambda_2}, \ldots, b'_{\lambda_n}$ and $f(a)$. Furthermore, we have that $\mathscr{N} \models \mathrm{Im}_{\bar{Q}}(f(a))$, where $\mathsf{f}_i = \mathsf{f}_{Q_i}$. Since $\mathscr{N} \models T_R$, there are $d_1, \ldots, d_n \in R(N)$ such that

$$\sum_{i=1}^{n} \mathsf{f}_i(d_i) = f(a).$$

Hence, by Lemma 4.2.4, b'_λ is in the orbit of d_i for some $i \in [n]$: this shows that $b'_\lambda \in N$.

Since for all $\lambda_1 \neq \lambda_2$ and all $z \in \mathbf{Z}$, the formula $S^k(x_{\lambda_1}) \neq x_{\lambda_2}$ is in q, we have that for all $\lambda_1 \neq \lambda_2$, $b'_{\lambda_1} \notin \mathrm{Orb}(b'_{\lambda_2})$. Likewise, we have that $b'_\lambda \notin f(A)$ for all $\lambda < \kappa$. This shows that $(b'_\lambda)_{\lambda<\kappa}$ realizes the quantifier-free type of $(b_\lambda)_{\lambda<\kappa}$ over A in $\mathcal{L}_0$. Hence the map $\tilde{f}$ defined on $\tilde{\mathscr{A}}$ by $a + \sum_{i=1}^{n} \mathsf{f}_i(b_{\lambda_i}) \mapsto f(a) + \sum_{i=1}^{n} \mathsf{f}_i(b'_{\lambda_i})$ is an $\mathcal{L}_0$-embedding.

Now we extend $\tilde{f}$ to an $\mathcal{L}$-embedding $\bar{f} : \overline{\mathscr{A}} \to \mathscr{N}$. Recall that for all $d \in \overline{\mathscr{A}}\backslash\tilde{\mathscr{A}}$, there exist $a \in \mathscr{A}$, $\bar{\mathsf{f}}$ a tuple of non-trivial operators, $b_{\lambda_1}, \ldots, b_{\lambda_n}$ and $n \in \mathbf{N}^{>0}$ such that $nd = a + \sum_{i=1}^{n} \mathsf{f}_i(b_{\lambda_i})$. Because $\tilde{f}$ preserves $\mathcal{L}_0$, $\tilde{f}(nd)$ is divisible by n: by (Ax.1) there exists a unique d^* such that $\tilde{f}(nd) = nd^*$ (uniqueness follows from the fact that models of T_1 are torsionless). We extend $\tilde{f}$ by the rule $\bar{f}(d) = d^*$. So $\bar{f}$ respects the divisibility predicates. And since the Im predicates are definable by $\mathcal{L}_0 \cup \{D_n \mid n \in \mathbf{N}^{>1}\}$-formulas, we get that $\bar{f}$ is indeed an $\mathcal{L}$-embedding. □

4.4 T_R is 1-e.c.

We now show that T_R is 1-e.c. using the material of the two previous sections. We distinguish two cases for two models $\mathscr{M}_0 \subset \mathscr{M}$ of T_R. In the first case, we assume that $R(M) = R(M_0)$. This allows to work almost as if we were in $\mathrm{Th}(\mathbf{Z}, +, 0, 1, D_n \mid n \in \mathbf{N}^{>1})$. The only extra ingredient needed to make things work here is an analogue of Proposition 3.4.4, which states that the image of sums of operators of the form f_Q in $\mathbf{N}$ remains non piecewise syndetic. In the second case, where $R(M) \supsetneq R(M_0)$, we use our construction of algebraically prime models to reduce to the case where $\mathscr{M} = \overline{\mathscr{A}}$, where $\mathscr{A}$ is the substructure generated by $\mathscr{M}_0$ and $R(M)$.

We begin with a lemma on terms.

Lemma 4.4.1. *Let $\mathscr{M},\mathscr{M}_0 \models T_R$ such that $\mathscr{M}_0 \subset \mathscr{M}$. Let $t(x,\bar{y})$ be an $\mathcal{L}$-term, with $\bar{y}$ of size n. Then for all $d \in M \setminus M_0$ and $\bar{a} \in M_0^n$ one of the following holds:*

1. *if $md + a = b \in R(M)$ for some $m \in \mathbf{Z} \setminus \{0\}$ and $a \in M_0$, then there exist $Q \in \mathbf{Z}[X]$, $m' \in \mathbf{Z}$ and $a' \in M_0$ such that $t(d,\bar{a}) = \mathsf{f}_Q(b) + m'd + a'$;*
2. *if for all $m \in \mathbf{Z} \setminus \{0\}$ and all $a \in M_0$, $md + a \notin R(M)$, then there exist $m' \in \mathbf{Z}$ and $a' \in M_0$ such that $t(d,\bar{a}) = m'd + a'$.*

Proof. Let $d \in C \setminus M_0$.

1. Assume $md + a = b \in R(M)$ for some $m \in \mathbf{Z} \setminus \{0\}$ and $a \in M_0$. It is enough to show that for all $Q \in \mathbf{Z}[X]$ $m' \in \mathbf{Z} \setminus \{0\}$ and all $a' \in M_0$, if $\mathsf{f}_Q(b) + m'd + a' = b' \in R(M) \setminus R(M_0)$, then there exists $k \in \mathbf{Z}$ such that $\mathsf{f}_Q(b) + m'd + a' = S^k(b)$. Notice that $\mathsf{f}_Q(b) + m'd + a' = b'$ is equivalent to $m\mathsf{f}_Q(b) + m'd + ma' - m'a = mb'$. Let $\mathsf{f}'(x) = m\mathsf{f}_Q(x) + m'x$, so that $\mathsf{f}'(b) - mb' = m'a - ma'$. Notice that f' is non-trivial, since $b' \in R(M) \setminus R(M_0)$. Since $\mathscr{M}_0$ is a model of T_R, we can find $b_0, b_0' \in R(M_0)$ such that $\mathsf{f}'(b_0) - mb_0' = m'a - ma'$. Since $b, b' \in R(M) \setminus R(M_0)$, this implies by Lemma 4.2.4 that $m'a - ma' = 0$. As f' is non-trivial and $m \in \mathbf{Z} \setminus \{0\}$, (b,b') is a non-degenerate solution of $\mathsf{f}'(x) - my = 0$. So, by (Ax.8), there exists $k \in \mathbf{Z}$ such that $b' = S^k(b)$, which is what we wanted.
2. Assume that for all $m \in \mathbf{Z} \setminus \{0\}$ and all $a \in M_0$, $md + a \notin R(M)$. In that case $S(md + a) = md + a$ for all $m \in \mathbf{Z} \setminus \{0\}$ and all $a \in M_0$. This is enough to conclude. □

Let us treat the first case.

Proposition 4.4.2. *Let $Q_1, \dots, Q_k \in \mathbf{Z}[X]$. Then for all $\mathscr{M} \models T_R$, $\{z \in \mathbf{N} \mid \mathscr{M} \models \exists \bar{x} \in R\ \mathsf{f}_{Q_1}(x_1) + \cdots + \mathsf{f}_{Q_k}(x_k) = z\}$ is not piecewise syndetic.*

Proof. This is an immediate consequence of the following observation: $\{z \in \mathbf{N} \mid \mathscr{M} \models \exists \bar{x} \in R\ \mathsf{f}_{Q_1}(x_1) + \cdots + \mathsf{f}_{Q_k}(x_k) = z\} = \{z \in \mathbf{N} \mid \mathscr{Z}_{R,\mathcal{L}} \models \exists \bar{x} \in R\ \mathsf{f}_{Q_1}(x_1) + \cdots + \mathsf{f}_{Q_k}(x_k) = z\}$ and Proposition 3.4.4. □

Proposition 4.4.3. *Let $\mathscr{M},\mathscr{M}_0 \models T_R$ such that $\mathscr{M}_0 \subset \mathscr{M}$. Assume that $R(M_0) = R(M)$. Then $\mathscr{M}_0$ is 1-e.c. in $\mathscr{M}$.*

Proof. Let $\varphi(x,\bar{y})$ be a quantifier-free formula such that $\mathscr{M} \models \varphi(b,\bar{a})$ for some $b \in M \backslash M_0$ and $\bar{a} \in M_0$. We will show that there exists $b_0 \in M_0$ such that $\mathscr{M}_0 \models \varphi(b_0,\bar{a})$. Let us simplify φ.

First, by Lemma 4.4.1, we fave that for all $\mathcal{L}$-terms $t(x,\bar{y})$, $\bar{y}$ of size n, for all $b \in M \setminus M_0$ and all $\bar{a} \in M_0^n$, there are $n \in \mathbf{Z}$ and $a \in M_0$ such that $t(b,\bar{a}) = nb + a$. All this can be tracked by a conjunction of negation of Im predicate, so we may assume, at the cost of adding this conjunction, that all terms involved in φ are of the form $nx + a$.

Now we look at the atomic formulas satisfied by elements in $M \setminus M_0$ with parameters in M_0. Let $b \in M \setminus M_0$, $n_1, \dots, n_k \in \mathbf{Z}$, $a_1, \dots, a_k \in M_0$, $[Q] \in (\mathbf{Z}[X])^{k\times m}$ and D be a set of divisibility conditions of size m. Since $R(M) = R(M_0)$, we have $\mathscr{M} \models \mathrm{Im}_{[Q],D}(n_1 b + a_1, \dots, n_k b + a_k)$ if and only if $n_1 = \cdots = n_k = 0$ and $\mathscr{M} \models \mathrm{Im}_{[Q],D}(\bar{a})$. Likewise, for all $n \in \mathbf{Z}$ and $a \in M_0$, we have $\mathscr{M} \models nb + a = 0$ if and only if $n = 0$ and $\mathscr{M} \models a = 0$.

Thus, after writing $\varphi(x, \bar{y})$ in its equivalent disjunctive normal form, we may select a conjunctive clause satisfied by $(b, \bar{a})$ and then assume that $\varphi(x, \bar{a})$ is of the form

$$\begin{aligned} &\bigwedge_{i\in I_1} n_i x + a'_i \neq 0 \\ \wedge &\bigwedge_{i\in I_2} \neg\mathrm{Im}_{[Q]_i, D_i}(n_{i1}x + a'_{i1}, \dots, n_{im_i}x + a'_{im_i}) \\ \wedge &\bigwedge_{i\in I_3} D_{\ell_i}(n_i x + k_i), \end{aligned}$$

where for all $i \in I_1$, $n_i \in \mathbf{Z} \setminus \{0\}$ and $a'_i = t_i(\bar{a})$ for some $\mathcal{L}$-term $t_i(\bar{y})$, for all $i \in I_2$, $m_i \in \mathbf{N}^{>0}$, $\bar{n}_i \in (\mathbf{Z} \setminus \{0\})^{m_i}$ and $a'_{ij} = t_{ij}(\bar{a})$ for some $\mathcal{L}$-term $t_{ij}(\bar{y})$ and for all $i \in I_3$, $n_i \in \mathbf{Z} \setminus \{0\}$, $\ell_i \in \mathbf{N}^{>1}$ and $0 \leq k_i < \ell_i$.

Let us finally show that $\varphi(M_0, \bar{a})$ is not empty. By model completeness of the theory $\mathrm{Th}(\mathbf{Z}, +, 0, 1, D_n | 1 < n \in \mathbf{N})$, there exists $b_0 \in M_0$ such that

$$\mathscr{M}_0 \models \bigwedge_{i\in I_1} n_i b_0 + a'_i \neq 0 \wedge \bigwedge_{i\in I_3} D_{\ell_i}(n_i b_0 + a'_i).$$

However, $\mathscr{M}_0$ may not satisfy $\varphi(b_0, \bar{a})$. But this can be overcome in the following way. Let

$$X_1 = \left\{ m \in \mathbf{N} \,\middle|\, \mathscr{M}_0 \models \bigwedge_{i\in I_3} n_i(b_0 + m) + a'_i \neq 0 \right\},$$

$$X_2 = \left\{ m \in \mathbf{N} \,\middle|\, \mathscr{M}_0 \models \bigwedge_{i\in I_2} \neg\mathrm{Im}_{[Q]_i, D_i}(n_{i1}(b_0 + m) + a'_{i1}, \dots) \right\},$$

and

$$X_3 = \left\{ m \in \mathbf{N} \,\middle|\, \mathscr{M}_0 \models \bigwedge_{i\in I_3} D_{\ell_i}(n_i m) \right\}.$$

We want to show that the set $X = X_1 \cap X_2 \cap X_3$ is not empty, thus ensuring that $\mathscr{M}_0 \models \varphi(b_0 + m, \bar{b})$ for some $m \in \mathbf{N}$. Suppose otherwise that $X = \emptyset$. This implies that $X_3 \subset \mathbf{N} \setminus (X_1 \cap X_2)$. But then, since X_3 is piecewise syndetic, $\mathbf{N} \setminus (X_1 \cap X_2)$ is piecewise syndetic. Hence, by Brown's Lemma (see Theorem 3.4.3), $\mathbf{N} \setminus X_2$ is piecewise syndetic, $\mathbf{N} \setminus X_1$ being finite. But, by Proposition 4.4.2, this is not possible since $\mathbf{N} \setminus X_2$ is in the image of a sum of operators. $\square$

In order to establish that T_R is 1-e.c., we need a final lemma on Im predicates, in which we show that a conjunction of Im predicates is equivalent to an Im predicate.

Lemma 4.4.4. *For all $[Q]_1 \in \mathbf{Z}^{n_1 \times m_1}, \ldots, [Q]_\ell \in \mathbf{Z}^{n_\ell \times m_\ell}$ and sets of divisibility conditions $D_1, \ldots, D_\ell$, there exists $[Q] \in \mathbf{Z}^{(n_1 \cdots n_\ell) \times (m_1 \cdots m_\ell)}$ and a set of divisibility conditions D such that*

$$T_R \models \forall \bar{y}_1, \ldots, \bar{y}_\ell \left(\bigwedge_{i \in [\ell]} \mathrm{Im}_{[Q]_i, D_i}(\bar{y}_i) \leftrightarrow \mathrm{Im}_{[Q], D}(\bar{y}_1, \ldots, \bar{y}_\ell) \right).$$

Proof. Just take $[Q] = [Q]_1 \oplus \cdots \oplus [Q]_\ell$ ($\oplus$ denotes the direct sum) and $D = D_1 \sqcup \cdots \sqcup D_\ell$. $\square$

Theorem 4.4.5. *The theory T_R is 1-e.c.*

Proof. Let us show that for all $\mathscr{M}, \mathscr{M}_0 \models T_R$ such that $\mathscr{M}_0 \subset \mathscr{M}$, then $\mathscr{M}_0$ is 1-e.c. in $\mathscr{M}$. Let $\mathscr{M}, \mathscr{M}_0 \models T_R$ such that $\mathscr{M}_0 \subset \mathscr{M}$. Two cases are possible: either $R(M_0) = R(M)$ or $R(M_0) \subsetneq R(M)$. The first case has been proved in Proposition 4.4.3. So let us assume that we are in the second case.

By Lemma 4.3.1, we may assume that $\mathscr{M} = \overline{\mathscr{A}}$ where $\mathscr{A}$ is the substructure of $\mathscr{M}$ generated by $M_0 \cup R(M)$. Recall that by the proof of Lemma 4.3.1, any element d of $\mathscr{M}$ is such that $nd = a + \sum_{i=1}^{\ell} \mathsf{f}_i(b_i)$, where $n \in \mathbf{N}$, $a \in M_0$ and $b_1, \ldots, b_\ell \in R(M) \backslash R(M_0)$ are in different orbits. Our strategy is to establish that $\mathscr{M}_0$ is 1-e.c. in $\mathscr{M}$ from the fact that for all tuples $\bar{b}$ of elements of $R(M) \backslash R(M_0)$ in different orbits, all $\bar{a} \in M_0$ and all $\varphi(\bar{x}, \bar{y})$, $\mathscr{M} \models \varphi(\bar{b}, \bar{a})$ implies $\mathscr{M}_0 \models \exists \bar{x} \in R\, \varphi(\bar{x}, \bar{a})$.

Let $\varphi(x, \bar{y})$ an $\mathcal{L}$-formula, with $\bar{y}$ of size k, such that $\mathscr{M} \models \varphi(d, \bar{a})$, for some $d \in M \setminus M_0$ and $\bar{a} \in M_0^k$. Using Lemma 4.4.1, the fact that, since $d \in M \setminus M_0$, $nd = a + \sum_{i=1}^{\ell} \mathsf{f}_i(b_i)$, for some $n \in \mathbf{N}^{>0}$, $a \in M_0$, $b_1, \ldots, b_\ell$ in different orbits and $\mathsf{f}_1, \ldots, \mathsf{f}_\ell$ non-trivial, $D_{n_1 n_2}(n_1 x) \leftrightarrow D_{n_2}(x)$ and $\mathrm{Im}_{[Q], D}(\bar{y}) \leftrightarrow \mathrm{Im}_{n[Q], D}(n\bar{y})$, we may assume that $\varphi(x, \bar{a})$ is of the form

$$\begin{aligned} &\bigwedge_{i \in I_1} m_i \left(a + \sum_{i=1}^{\ell} \mathsf{f}_i(x_i) \right) + na_i' \neq 0 \wedge \bigwedge_{i \in I_2} D_{n\ell_i} \left(m_i \left(a + \sum_{i=1}^{\ell} \mathsf{f}_i(x_i) \right) + ns_i \right) \\ &\wedge \bigwedge_{i \in I_3} \mathrm{Im}_{n[Q]_i, D_i} \left(m_{i1} \left(a + \sum_{i=1}^{\ell} \mathsf{f}_i(x_i) \right) + na_{i1}', \ldots, m_{ik_i} \left(a + \sum_{i=1}^{\ell} \mathsf{f}_i(x_i) \right) + na_{ik_i}' \right) \\ &\wedge \bigwedge_{i \in I_4} \neg \mathrm{Im}_{n[Q]_i, D_i} \left(m_{i1} \left(a + \sum_{i=1}^{\ell} \mathsf{f}_i(x_i) \right) + na_{i1}', \ldots, m_{ik_i} \left(a + \sum_{i=1}^{\ell} \mathsf{f}_i(x_i) \right) + na_{ik_i}' \right). \end{aligned}$$

where, for all $i \in I_1 \cup I_2$, $a_i' = t_i(\bar{a})$ for some $\mathcal{L}$-term $t_i(\bar{y})$, $m_i \in \mathbf{Z} \setminus \{0\}$, $\ell_i \in \mathbf{N}^{>1}$, $0 \leq s_i < \ell_i$ and for all $i \in I_3 \cup I_4$, $k_i \in \mathbf{N}^{>0}$, $a_{ij}' = t_i(\bar{a})$ for some $\mathcal{L}$-term $t_{ij}(\bar{y})$,

$\bar{m}_i \in (\mathbf{Z} \setminus \{0\})^{k_i}$, $[Q]_i \in (\mathbf{Z}[X])^{k_i \times k'_i}$ and D_i is a set of divisibility conditions. We may also assume that $|I_3| \leq 1$ by Lemma 4.4.4.

Furthermore, we may replace

$$\bigwedge_{i \in I_2} D_{n\ell_i} \left(m_i \left(a + \sum_{i=1}^{\ell} \mathsf{f}_i(x_i) \right) + ns_i \right)$$

by

$$\bigwedge_{i \in [\ell]} D_{\ell'_i} \left(m'_i \mathsf{f}_i(x_i) + s'_i \right),$$

where for all $i \in [\ell]$, $\ell_i \in \mathbf{N}^{>1}$, $m'_i \in \mathbf{Z}$ and $0 \leq s'_i < \ell'_i$. Finally, by Lemma 4.2.3 and Corollary 4.2.8, we may assume that $\tilde{\varphi}(\bar{x}, \bar{a})$ is of the form

$$\bigwedge_{i \in [n]} \bigwedge_{j \in J} D_{\ell_{ij}}(\mathsf{f}_{Q_{ij}}(S^{k_j}(x_i)) + k_{ij})$$
$$\wedge \bigwedge_{(i,j) \in K_1} \mathsf{f}_{Q'_j}(x_i) = 0 \wedge \bigwedge_{(i,j) \in K_2} \mathsf{f}_{Q'_j}(x_i) \neq 0 \wedge \bigwedge_{i \in I} \bar{x}_{J_i} \notin F_i,$$

where, for all $i \in I$, F_i is a finite set of $|J_i|$-tuples in M_0. But then, by (Ax.5) and (Ax.7), we may find a realization $\bar{b}_0$ of $\tilde{\varphi}(\bar{x}, \bar{a})$ in $R(M_0)$, as desired. □

4.5 Superstability of T_R

From the quantifier elimination of T_R, we deduce, by means of counting types, that it is superstable.

Theorem 4.5.1. *The theory T_R is superstable.*

Proof. Let $\mathscr{C}$ be a monster model of T_R and let $A \subset C$ be a small set of parameters. We want to show that $|S_1(A)| \leq \max\{2^{\aleph_0}, |A|\}$. Without loss of generality, we may assume that A is the domain of a model $\mathscr{A}$. By quantifier elimination (see Theorem 4.1.1), any type $p(x)$ over A is determined by the set of atomic formulas it contains. Let $\mathcal{L}_1 = \mathcal{L}_g \cup \mathcal{L}_S$ and $\mathcal{L}_2$ be $\mathcal{L} \setminus \{D_n \mid n > 1\}$. Let $p_{|\mathcal{L}_i}$ denote the restriction of p to $\mathcal{L}_i$, so that $p(x) = p_{|\mathcal{L}_1}(x) \cup p_{|\mathcal{L}_2}(x)$. We may assume that $p(x)$ does not contain a formula of the form $x = a$ for some $a \in M$. We consider two cases:

(Case 1) there exist $m \in \mathbf{Z} \setminus \{0\}$ and $a \in A$ such that $R(mx + a) \in p(x)$;

(Case 2) for all $m \in \mathbf{Z} \setminus \{0\}$ and all $a \in A$, $R(mx + a) \notin p(x)$.

By Lemma 4.4.1, we may assume in the rest of the proof that the terms (with parameters in A) that appear in formulas are of the form $\mathsf{f}_Q(mx + a) + m'x + a'$, where $m' \in \mathbf{Z}$, $a' \in A$ and

1. $m \in \mathbf{Z} \setminus \{0\}$ and $a \in A$ are fixed when we are in (Case 1);
2. $Q = 0$ when we are in (Case 2).

Claim 4.5.2. The number of types of the form $p_{|\mathcal{L}_1}(x)$ is at most $\max\{2^{\aleph_0}, |A|\}$.

Proof. Indeed, any formula of the form $D_n(\mathsf{f}_Q(mx+a)+m'x+a')$ is equivalent to a formula of the form $D_n(\mathsf{f}_Q(mx+a)+m'x+k)$, where $k \in \mathbf{Z}$ is such that $D_n(a'-k)$. In (Case 2), we know that a formula of the form $m'x+a'=0$ is never in $p(x)$, unless $m'=0$ and $a'=0$. Let us now look at equations when we are in (Case 1).

Assume that $\mathsf{f}_Q(mx+a)+m'x+a' = 0 \in p_{|\mathcal{L}_1}(x)$, where $m' \in \mathbf{Z}$ and $a' \in A$. Then, by (Ax.6), $\mathrm{Im}_{mQ,m'X}(m'a-ma')$ holds in $\mathscr{A}$. Thus there exists $b' \in R(A)$ such that $m\mathsf{f}(mx+a)+m'(mx+a) = m\mathsf{f}(b')+m'b'$. This implies, by Lemma 4.2.4 that $m\mathsf{f}(mx+a)+m'(mx+a) = 0$. Hence $m'a-ma' = 0$. So the only equations that appear in $p(x)$ are of the form $\mathsf{f}(mx+a)=0$. □

By the previous claim, it remains to show that the number of types of the form $p_{|\mathcal{L}_2}(x)$ is at most $\max\{2^{\aleph_0}, |A|\}$. So we need to look at formulas of the form $\mathrm{Im}_{[Q],D}(\mathsf{f}_1(mx+a)+m_1x+a_1,\ldots,\mathsf{f}_k(mx+a)+m_kx+a_k)$. For simplicity, we only look at the case $k=1$. We may restrict ourselves to formulas of the form $\mathrm{Im}_{\bar{Q},D}(\mathsf{f}(mx+a)+a')$ in (Case 1) and $\mathrm{Im}_{\bar{Q},D}(m'x+a')$ in (Case 2). We want to show that in both cases, we can separate the parameters from the variable, in the same way we did for divisibility conditions. This will be enough to conclude. For (Case 1), this is a consequence of Corollary 4.2.8. For (Case 2), we have the following claim.

Claim 4.5.3. Assume we are in (Case 2). Let $\bar{Q} \in \mathbf{Z}[X]^n$ and $m \in \mathbf{Z}\setminus\{0\}$. Then there exists at most one $a_{\bar{Q}} \in A$ such that $\sum_{i=1}^n \mathsf{f}_{Q_i}(x_i) = mx + a_{\bar{Q}}$ has a non-degenerate solution in $R(C)\setminus R(A)$.

Proof. Assume that there exists another $a' \in A$ that satisfies the claim. Then we have $\mathscr{A} \models \mathrm{Im}_{\bar{\mathsf{f}},-\bar{\mathsf{f}}}(a_{\bar{Q}}-a')$. Thus, we can find tuples $\bar{b}_1, \bar{b}_2 \in R(C)\setminus R(A)$ and $\bar{b}'_1, \bar{b}'_2 \in R(A)$ such that

$$\sum_{i=1}^n \mathsf{f}_{Q_i}(b_{1i}) - \mathsf{f}_{Q_i}(b_{2i}) - (\mathsf{f}_{Q_i}(b'_{1i}) - \mathsf{f}_{Q_i}(b'_{2i})) = 0.$$

But this can happen only if $a_{\bar{Q}} = a'$ by Lemma 4.2.4. □

As a consequence, we get that in (Case 2), a formula of the form $\mathrm{Im}_{\bar{Q},D}(mx+a)$ is in $p_{|\mathcal{L}_2}(x)$ if and only if some disjunction of formulas of the form

$$\mathrm{Im}_{\bar{Q}_I,D}(mx+a_{\bar{Q}_I}) \wedge \mathrm{Im}_{\bar{Q}_{[n]\setminus I},D}(a-a_{\bar{Q}_I})$$

is in $p_{|\mathcal{L}_2}(x)$. This proves that the number of types of the form $p_{|\mathcal{L}_2}(x)$ in (Case 2) is at most $\max\{|A|, 2^{\aleph_0}\}$. We conclude that $|S_1(A)| \leq \max\{|A|, 2^{\aleph_0}\}$. □

Since we do not know whether S and 1 are definable in $\mathcal{L}_R$ we cannot deduce directly that $\mathrm{Th}(\mathscr{Z}_R)$ is superstable. Nevertheless we can recover superstability of $\mathrm{Th}(\mathscr{Z}_R)$ (in $\mathcal{L}_R$) from the superstability of T_R using the following consequence of a variation of elementary amalgamation.

Theorem 4.5.4 ([25, Corollary 6.6.2]). *Let $\mathcal{L}_1 \subset \mathcal{L}_2$ be two languages and T an $\mathcal{L}_2$-theory. Let $\mathscr{A}$ be an $\mathcal{L}_1$-structure. Then $\mathscr{A} \models T_{\mathcal{L}_1}$ if and only if for some model $\mathscr{M}$ of T, $\mathscr{A} \prec \mathscr{M}_{\mathcal{L}_1}$. Here $T_{\mathcal{L}_1}$ is the set $\mathcal{L}_1$-consequences of T and $\mathscr{M}_{\mathcal{L}_1}$ is the $\mathcal{L}_1$-reduct of $\mathscr{M}$.*

Corollary 4.5.5. $\mathrm{Th}(\mathscr{Z}_R)$ *is superstable.*

Proof. Applying Theorem 4.5.4 to $\mathcal{L}_1 = \mathcal{L}_R$, $\mathcal{L}_2 = \mathcal{L}$ and $T = T_R$, we obtain that for all model $\mathscr{A}$ of $T_{\mathcal{L}_1} = \mathrm{Th}(\mathscr{Z}_R)$, there is a model $\mathscr{M}$ of T such that $\mathscr{A} \prec \mathscr{M}_{\mathcal{L}_1}$. In particular, we have that $|S_1^{\mathcal{L}_R}(A)| \leq |S_1^{\mathcal{L}}(A)| \leq \max\{2^{\aleph_0}, |A|\}$, where the last inequality comes from Theorem 4.5.1. □

4.6 Decidability of T_R

As a consequence of the fact that the theory of $\mathscr{Z}_R$ is axiomatized by T_R when R is enumerated by a regular sequence (r_n), we get the following decidability result. First let us recall some terminology.

Definition 4.6.1. Let (r_n) be a sequence in $\mathbf{Z}$:

1. (r_n) is *congruence periodic* if for all $k \in \mathbf{N}^{>1}$, there exist constants $m, p \in \mathbf{N}$ such that the sequence $(r_n)_{n \geq m}$ is periodic modulo k with period p;
2. (r_n) is *effectively congruence periodic* if there is a recursive function $f : \mathbf{N}^{>1} \to \mathbf{N} \times \mathbf{N}$ such that for all $k \in \mathbf{N}^{>1}$, $(r_n)_{n \geq m}$ is periodic modulo k of period p, where $(m, p) = f(k)$;
3. (r_n) has *an effective Kepler limit $\theta \in \mathbf{R}^{\geq 0}_{\infty}$* if there is a recursive function $\delta : \mathbf{N} \to \mathbf{N}$ such that $\forall n \in \mathbf{N} \forall m \geq \delta(n)\, |r_n / r_{n+1} - \theta^{-1}| \leq 1/2^n$, where by convention $\infty^{-1} = 0$.

Theorem 4.6.2. *Let R be a regular set and let (r_n) be a strictly increasing enumeration of R such that (r_n) has an effective Kepler limit and is effectively congruence periodic. Assume that the sets in (Ax.8) can be computed effectively. Then the $\mathcal{L}$-theory T_R is decidable.*

Proof. Indeed, under these assumptions, the constants that appear in (Ax.7) can be computed effectively. To see this, we first use the remarks after the proof of Proposition 3.2.2 to show that for all $Q \notin \mathrm{Triv}$, we can compute $\ell = \ell(Q)$ such that $\forall x (x > \ell \to \mathrm{f}_Q(x) \neq 0)$. Then using the proof of Corollary 3.3.4 and our assumption that (Ax.8) is effective, we can compute $k(Q)$ for all $Q \notin \mathrm{Triv}$.

For (Ax.6), effectiveness follows from the effective periodicity of R. Thus, T_R is recursively axiomatizable. And since T_R is complete, we may conclude that T_R is decidable. □

Because the proof of Proposition 3.3.1 is not effective, we cannot *a priori* remove the assumption on (Ax.8). However, we believe that the analysis of expansion of Presburger arithmetic in the next chapter could provide a way to demonstrate Proposition 3.3.1 effectively. Indeed, we show in the next chapter that inequalities of the form $\mathsf{f}_1(n_1) + \cdots + \mathsf{f}_k(n_k) < a$ can be dealt with as follows:

1. for each $i \in [k]$, we look at solutions of $\mathsf{f}_1(n_1) + \cdots + \mathsf{f}_k(n_k) < a$ such that n_i is far away from n_j, $j \neq i$, in the sense that $n_i - n_j$ is bigger than some constant depending only on f_i (and this constant can be obtained effectively as long as the Kepler limit is effective). We say in that case that n_i dominates n_j, $j \neq i$. Then we frame a by two consecutive images of f_i, say $\mathsf{f}(n_0) \leq a < \mathsf{f}(n_0 \pm 1)$, and reduce the satisfaction of $\mathsf{f}_1(n_1) + \cdots + \mathsf{f}_k(n_k) < a$ to the relative position of n_1 and n_0;
2. then we treat the case where none of the n_i dominates the other variables. In that case, we repeat the analysis in the first case by replacing n_i with $n_j + k$ for some k. This process then reduces to the case where $k = 1$.

We could apply this to equations $\mathsf{f}_1(n_1) + \cdots + \mathsf{f}_k(n_k) = 0$ using the fact that it is equivalent to

$$(\mathsf{f}_1(n_1) + \cdots + \mathsf{f}_k(n_k) < 1) \wedge (-\mathsf{f}_1(n_1) - \cdots - \mathsf{f}_k(n_k) < 1).$$

As we haven't checked this in details, we leave the question of the necessity of our assumption on (Ax.8) open.

CHAPTER

5 Expansion of $(\mathbf{Z},+,0,<)$ by a sparse sequence

In [49], A. L. Semenov studied various expansions of Presburger arithmetic, that is $\mathrm{Th}(\mathbf{Z},+,0,<)$. The focus there is on decidability issues for those expansions. One important class of expansions studied in [49] is the class of expansions by a sparse set. These sets are enumerated by fast growing sequences and, as we shall see in the next section, regular sets are particular instances of sparse sets. Among other results, it is shown that when R is a congruence periodic sparse set, $\mathrm{Th}(\mathbf{Z},+,0,<,R)$ is decidable if and only if R is effectively sparse and effectively congruence periodic (see [49, Corollary 2]). For regular sets, being effectively sparse means that it has an effective Kepler limit. As a result $\mathrm{Th}(\mathbf{Z},+,0,<,A_q)$ and $\mathrm{Th}(\mathbf{Z},+,0,<,\mathrm{Fib})$ are decidable. More generally, $\mathrm{Th}(\mathbf{Z},+,0,<,\{r_n \mid n\in\mathbf{N}\})$ is decidable whenever (r_n) is a regular sequence that have an effective algebraic Kepler limit. Indeed, in that case (r_n) is a linear recurrence sequence and it is congruence periodic by Proposition 1.3.12 and in fact the congruence periodicity is effective.

The proof of [49, Corollary 2] relies on [49, Theorem 3], where it is established that a certain theory associated to $\mathrm{Th}(\mathbf{Z},+,0,<,R)$ is existential. The results we present here, specifically Theorem 5.5.3, use the techniques of the proof of [49, Theorem 3].

Expansions of Presburger arithmetic by sparse sets have also been studied by F. Point in [41], where another proof of [49, Corollary 2] is given using a quantifier elimination result, under the assumption of congruence periodicity (see [41, Proposition 9]). This last quantifier result allowed us, in collaboration with F. Point, to prove that expansions of Presburger by a congruence periodic sparse set is dependent (see [29, Theorem 2.32]).

It is the purpose of this chapter to give another proof of [29, Theorem 2.32], without using the congruence periodicity assumption. The cost of this operation is the necessity of another quantifier elimination result, and this is done by revisiting the proof of [49, Theorem 3]. We give such a quantifier elimination, in Theorem 5.5.3, in a definitional expansion $\mathcal{L}_<$ of $\{+,-,0,1,<,R,D_n \mid n\in\mathbf{N}\}$ by adding the language $\mathcal{L}_S$ and analogues of the Im predicates we used in the previous chapter. The Im predicates we add here differ in two ways from the ones used in Chapter 4:

1. we use systems of inequalities of the form $\mathsf{f}_1(x_1) + \cdots + \mathsf{f}_n(x_n) > y$ instead of equalities;
2. the divisibility conditions used in Chapter 4 to define Im predicates are replaced by *arbitrary* formulas in the language $\{S, S^{-1}, c, <, D_{n,k} \mid k < n \in \mathbf{N}\}$, with quantifiers relativized to R and $D_{n,k}(x)$ interpreted as $D_n(x+k)$.

While the last difference may seem a bit much, it will allow us to show that negation of an Im predicate is equivalent to a finite disjunction of Im predicates: this is a nice property that we did not have in Chapter 4.

Using quantifier elimination in $\mathcal{L}_<$, we then show that, for R a sparse set, the dependency of $\mathrm{Th}(\mathbf{Z}, +, 0, <, R)$ follows from the dependency of $\mathrm{Th}(\mathscr{R})$, where $\mathscr{R} = (R, S, S^{-1}, c, <, D_{n,k} \mid k < n \in \mathbf{N})$ (see Theorem 5.6.7). This is done by showing that the formulas that define Im predicates have honest definitions over R. Another essential ingredient in the proof of Theorem 5.6.7 is a separation of variables phenomenon in $\mathcal{L}_<$-terms.

This chapter is organized as follows. In Section 5.1, we define what are sparse sets and we show that regular sets are sparse. In Section 5.2, we introduce the theory $T_{R,<}$ that we use to axiomatize expansions of Presburger by sparse sets. Then in Section 5.3 we explain how to handle inequalities of the form $\mathsf{f}_1(x_1) + \cdots + \mathsf{f}_n(x_n) > y$ and then we use this to show, in Section 5.4, that the negation of an Im predicate is equivalent to a disjunction of Im predicates. Quantifier elimination is then established in Section 5.5. The dependency of $T_{<,R}$ is considered in Section 5.6. We finally end this chapter with Section 5.7 where the dependency of expansions of $(\mathbf{Q}, +, 0, <)$, $(\mathbf{R}, +, 0, <)$ and $(\mathbf{R}, +, 0, \lfloor \cdot \rfloor, <)$ by a sparse predicate is considered.

5.1 Sparse sets

We begin by introducing some notations. As in Chapter 3, given a sequence (r_n) of natural numbers and $Q \in \mathbf{Z}[X]$, f_Q denotes the function $n \mapsto a_0 r_n + a_1 r_{n+1} + \cdots + a_d r_{n+d}$, where $Q(X) = a_0 + a_1 X + \cdots + a_d X^d$. Given $Q \in \mathbf{Z}[X]$, $\mathsf{f}_Q = 0$ means $\{n \in \mathbf{N} \mid \mathsf{f}_Q(n) = 0\} = \mathbf{N}$, $\mathsf{f}_Q >_R 0$ means $\{n \in \mathbf{N} \mid \mathsf{f}_Q(n) > 0\}$ is cofinite and likewise $\mathsf{f}_Q <_R 0$ means $\{n \in \mathbf{N} \mid \mathsf{f}_Q(n) < 0\}$ is cofinite.

Definition 5.1.1. Let (r_n) be a sequence of natural numbers and $R \subset \mathbf{N}$.

1. We say that (r_n) is *sparse* if
   a) for all $Q \in \mathbf{Z}[X]$, either $\mathsf{f}_Q = 0$ or $\mathsf{f}_Q >_R 0$ or $\mathsf{f}_Q <_R 0$;
   b) for all $Q \in \mathbf{Z}[X]$, if $\mathsf{f}_Q >_R 0$ then there exists $k \in \mathbf{N}$ such that $\mathsf{f}_Q(n+k) - r_n > 0$ for all $n \in \mathbf{N}$.

2. We say that R is *sparse* if it is enumerated by a sparse sequence.

We note that R being sparse is independent of a choice of an enumeration by a sparse sequence (r_n). Indeed if (r_n) is sparse, then it must be ultimately strictly increasing, since $\mathsf{f}_Q >_R 0$, where $Q(X) = X - 1$. As a result, when given a sparse set R, we may assume that (r_n) is strictly increasing. We note that a sparse set is thus automatically infinite.

We end this section by giving a relation between sparse and regular sequences.

Lemma 5.1.2. *Let (r_n) be a regular sequence, then (r_n) is sparse.*

Proof. Let $\theta = \lim_{n\to\infty} r_{n+1}/r_n \in \mathbf{R}_{\infty}^{\geq 1}$. It is known that (r_n) is sparse if $\theta = \infty$ [49, § 3] and when $\lim_{n\to\infty} r_n/\theta^n \in \mathbf{R}^{>0}$ [41, §4]. Nevertheless, we shall treat all cases for convenience of the reader.

Let $Q \in \mathbf{Z}[X]$, of degree d. Recall that

$$\lim_{n\to\infty} \frac{\mathsf{f}_Q(n)}{r_{n+d}} = \begin{cases} a_d & \text{if } \theta = \infty \\ \theta^{-d} Q(\theta) & \text{otherwise.} \end{cases}$$

Therefore

1. if $\theta = \infty$, $\mathsf{f}_Q = 0$ if and only if $Q = 0$, $\mathsf{f}_Q >_R 0$ if and only if $a_d > 0$ and $\mathsf{f}_Q <_R 0$ if and only if $a_d < 0$;
2. if $\theta \in \mathbf{R}^{>1}$, $\mathsf{f}_Q = 0$ if and only if $Q(\theta) = 0$, $\mathsf{f}_Q >_R 0$ if and only if $Q(\theta) > 0$ and $\mathsf{f}_Q <_R 0$ if and only if $Q(\theta) < 0$.

Therefore the first condition in Definition 5.1.1 is satisfied.

Now let $Q \in \mathbf{Z}[X]$ such that $\mathsf{f}_Q >_R 0$. For $k \in \mathbf{N}$, let f_k be the function $n \mapsto \mathsf{f}_Q(n+k) - r_n$ and $Q_k(X) = X^k Q(X) - 1$. It is enough to find $k_0 \in \mathbf{N}$ such that $\mathsf{f}_{k_0} >_R 0$. Indeed, if that is the case, let $n_0 \in \mathbf{N}$ be such that $\mathsf{f}_{k_0}(n) > 0$ for all $n \geq n_0$. As a regular sequence is ultimately strictly increasing, we can find $n_1 \in \mathbf{N}$ such that $r_{n_1} > r_n$ for all $n < n_1$ and $r_{n+1} > r_n$ for all $n \geq n_1$. Set $k = \max\{k_0 + n_0, k_0 + n_1\}$. Then for all $n \in \mathbf{N}$, we have

$$\mathsf{f}_Q(n+k) > r_{n+k-k_0} > r_n,$$

because $n + k - k_0 \geq n_0$ and $n + k - k_0 \geq n_1$. Therefore the second condition of Definition 5.1.1 follows.

So let us find $k_0 \in \mathbf{N}$ such that $\mathsf{f}_{k_0} >_R 0$. If $\theta = \infty$, we may take $k_0 = 1$. Assume $\theta \in \mathbf{R}^{>1}$. We have that $\mathsf{f}_k >_R 0$ if and only if $Q_k(\theta) > 0$. As $Q(\theta) > 0$, we can find $k_0 \in \mathbf{N}$ such that $Q_{k_0}(\theta) = \theta^{k_0} Q(\theta) - 1 > 0$. As a result, we get that $\mathsf{f}_{k_0} >_R 0$. □

5.2 The theory $T_{R,<}$

We shall axiomatize expansions by a sparse set in a similar way as we axiomatized expansions by regular sets in Chapter 4. Let us fix a sparse set $R \subset \mathbf{N}$ that is sparse and enumerated by a (strictly increasing) sparse sequence (r_n) and consider $\mathscr{Z}_{R,<} = (\mathbf{Z}, +, -, 0, <, R)$. We can extract the following data from the definition of a sparse set:

1. $\mathbf{Z}[X]$ is partitioned in three sets Triv, Pos and Neg: for $Q \in \mathbf{Z}[X]$, $Q \in$ Triv if and only if $\mathsf{f}_Q = 0$, $Q \in$ Pos if and only if $\mathsf{f}_Q >_R 0$ and $Q \in$ Neg if and only if $\mathsf{f}_Q <_R 0$. Also for each $Q \in$ Pos (resp. each $Q \in$ Neg) we fix $n(Q) \in \mathbf{N}$ such that for all $n \in \mathbf{N}$ if $n > n(Q)$ then $\mathsf{f}_Q(n) > 0$ (resp. $\mathsf{f}_Q(n) < 0$). Also, we have that for all $Q \in \mathbf{Z}[X]$, $Q \in$ Pos if and only if $-Q \in$ Neg.
2. For each $Q \in$ Pos, we fix $k = k(Q) \in \mathbf{N}$ such that for all $n \in \mathbf{N}$, $\mathsf{f}_Q(n+k) - r_n > 0$.

Let us define the language in which we axiomatize $\mathrm{Th}(\mathscr{Z}_{R,<})$.

Recall that $\mathcal{L}_g$ is the language $\{+, -, 0, D_n \mid n \in \mathbf{N}^{>1}\}$ and $\mathcal{L}_S$ is the language $\{S, S^{-1}, c\}$. As we did in Chapter 4, these new symbols are interpreted as follows: c is interpreted as r_0, for all $n \in \mathbf{N}$, $S(r_n) = r_{n+1}$, $S^{-1}(r_{n+1}) = r_n$, $S^{-1}(c) = c$ and $S(z) = z = S^{-1}(z)$ for all $z \in \mathbf{Z} \setminus R$. For each $Q \in \mathbf{Z}[X]$, we let f_Q be the $\mathcal{L}_g \cup \mathcal{L}_S$-term $\sum_{i=0}^d a_i S^i(x)$, where $Q(X) = \sum_{i=0}^d$ and $S^0(x) = x$. We shall retain the terminology of Chapter 4 and call such terms *operators*. Given $Q \in \mathbf{Z}[X]$, f_Q is called *trivial* (resp. *positive*, *negative*) if $Q \in$ Triv (resp. $Q \in$ Pos, $Q \in$ Neg).

Let $[Q] \in \mathbf{Z}^{n\times m}$ and let $\varphi^<_{[Q]}(\bar{x}, \bar{y})$ be the formula

$$\bigwedge_{i\in[n]} \sum_{j\in[m]} \mathsf{f}_{Q_{ij}}(x_j) > y_i.$$

For $\varphi_0(\bar{x})$ a formula in the language $\mathcal{L}_0 = \{S, S^{-1}, c, <, D_{k,n} \mid k < n \in \mathbf{N}\}$, we let $\varphi_0^R(\bar{x})$ be the $\mathcal{L}_g \cup \mathcal{L}_S$-formula obtained by relativizing to the predicate R the quantifiers appearing in $\varphi_0(\bar{x})$ and replacing any occurrence of $D_{n,k}(t(\bar{x}))$ by $D_n(t(\bar{x}) + k)$, where $t(\bar{x})$ is an $\{S, S^{-1}, c\}$-term.

For each $[Q] \in \mathbf{Z}^{n\times m}$ and φ_0 an $\mathcal{L}_0$-formula, we associate an n-ary predicate $\mathrm{Im}_{[Q],\varphi_0}(\bar{y})$. In $\mathscr{Z}_{R,<}$, we interpret $\mathrm{Im}_{[Q],\varphi_0}(\bar{y})$ as follows: for all $\bar{m} \in \mathbf{Z}^{|\bar{y}|}$, $\mathrm{Im}_{[Q],\varphi_0}(\bar{m})$ if and only if $\exists \bar{x} \in R\big(\varphi^<_{[Q]}(\bar{x}, \bar{m}) \wedge \varphi_0^R(\bar{x})\big)$. These are the analogues of the Im predicates used in Chapter 4.

Our language $\mathcal{L}_<$ is thus defined as

$$\mathcal{L}_g \cup \{1, <\} \cup \mathcal{L}_S \cup \{R\} \cup \{\mathrm{Im}_{[Q],\varphi_0} \mid [Q] \in \mathbf{Z}^{n\times m} \text{ and } \varphi_0(\bar{x}) \text{ an } \mathcal{L}_0\text{-formula with } |\bar{x}| = n\}.$$

and we let $\mathscr{Z}_{R,\mathcal{L}_<}$ be the $\mathcal{L}_<$-expansion of $\mathscr{Z}_{R,<}$ we described.

We fix an axiomatization $T_{1,<}$ of $\mathrm{Th}(\mathbf{Z}, +, -, 0, 1, <, D_n \mid 1 < n \in \mathbf{N})$ (see [33, page 82]).

Let $T_{R,<}$ be the following set of axiom schemes:

(Ax.1) $T_{1,<}$;

(Ax.2) T_2^R;

(Ax.3) $c = r_0$ (that is c equals the term $\underbrace{1 + \cdots + 1}_{r_0 \text{ times}}$);

(Ax.4) $\forall x(\neg R(x) \to S(x) = x)$;

(Ax.5) $\forall x \in R(x < S(x))$;

(Ax.6) $\forall x \exists y \in R(x \geq c \to y \leq x < S(y))$;

(Ax.7) for all $[Q] \in \mathbf{Z}^{n \times m}$ and $\varphi_0(\bar{x})$, $|\bar{x}| = n$, an $\mathcal{L}_0$-formula,

$$\forall \bar{y}\Big(\mathrm{Im}_{[Q],\varphi_0}(\bar{y}) \leftrightarrow \exists \bar{x} \in R\big(\varphi^<_{[Q]}(\bar{x}, \bar{y}) \wedge \varphi_0^R(\bar{x})\big)\Big);$$

(Ax.8) for all $Q \in \mathbf{Z}[X]$, if $Q \in \mathrm{Triv}$, then we add the axiom

$$\forall x \in R\, \mathsf{f}_Q(x) = 0,$$

and if $Q \notin \mathrm{Triv}$, we add the axiom

$$(\forall x \in R(x > S^n(c) \to \mathsf{f}_Q(x) > 0)) \vee (\forall x \in R(x > S^n(c) \to \mathsf{f}_Q(x) < 0)),$$

where $n = n(Q)$ (see 1 on page 78);

(Ax.9) for all $Q \in \mathrm{Pos}$, we add the axiom

$$\forall x \in R\, \mathsf{f}_Q(S^k(x)) > x,$$

where $k = k(Q)$ (see 2 on page 78).

Remark 5.2.1. In what follows, we will use r_0 instead of c while treating formulas, as the meaning of r_0 is more explicit.

Our first goal is to prove that $T_{R,<}$ has quantifier elimination and is complete (see Theorem 5.5.3 and Corollary 5.5.4). The key ingredient in the proof of Theorem 5.5.3 is that the negation of an Im predicate is equivalent in $T_{R,<}$ to a disjunction of Im predicates. This is a key difference between T_R and $T_{R,<}$, and allow to reduce the technical work needed in Chapter 4 to construct algebraically prime models.

In Section 5.3, we give a detailed analysis of the satisfaction of formulas of the form $\mathsf{f}_1(x_1) + \cdots + \mathsf{f}_n(x_n) > y$ in $T_{R,<}$. We show there that we can manage these formulas easily after singling out a dominant variable among $\bar{x}$ and framing y with consecutive images of the operator corresponding to the dominant variable. This will allow us to prove by induction on n that the negation of Im predicates is a disjunction of Im predicates, in Section 5.4. Then we proceed in Section 5.5 with the proof of Theorem 5.5.3.

In what follows, we will use the following notations. For a n-tuple of variables $\bar{x}$ and $i \in [n]$, $\bar{x}^i$ denotes the tuple $(x_j \mid j \in [n] \setminus \{i\})$. For a term $t(\bar{x})$, the expression $|t(\bar{x})|$ stands for the absolute value of $t(\bar{x})$. Finally, the expression $z \in [x, y]$ (resp. $z \in]x, y[$) is a shorthand for the formula $x \leq z \leq y$ (resp. $x < z < y$).

We end this section with several consequences of the axioms of $T_{R,<}$. The next lemma states that an operator defined by a positive polynomial is eventually strictly increasing.

Lemma 5.2.2. *Let $Q \in \mathrm{Pos}$. Then there exists $\ell \in \mathbf{N}$ such that*

$$T_{R,<} \models \forall x \in R(x \geq S^\ell(r_0) \rightarrow \mathsf{f}_Q(x) < \mathsf{f}_Q(S(x)) \wedge \bigwedge_{i=0}^{\ell-1} \mathsf{f}_Q(S^i(r_0)) < \mathsf{f}_Q(S^\ell(r_0))).$$

Proof. Let $\mathsf{f}'(x)$ be the operator $\mathsf{f}_Q(S(x)) - \mathsf{f}_Q(x)$. Since $\mathsf{f}_Q >_R 0$, by (Ax.9), there exists $k \in \mathbf{N}$ such that $\mathsf{f}_Q(S^k(x)) > x$ for all $x \in R$. As a result, f_Q cannot be constant. Hence $\mathsf{f}' \neq_R 0$. Thus, by (Ax.8), we either have $\mathsf{f}' >_R 0$ or $\mathsf{f}' <_R 0$. Assume, towards a contradiction, that $\mathsf{f}' <_R 0$. Then there exists $n' \in \mathbf{N}$ such that for all $x \in R$, if $x > S^{n'}(r_0)$, then $\mathsf{f}_Q(x) > \mathsf{f}_Q(S(x))$. Let $m > \max\{n', n(Q)\}$. Because $Q \in \mathrm{Pos}$, we have that $d = \mathsf{f}_Q(S^m(r_0)) > 0$. On the other hand, since $\mathsf{f}' <_R 0$, we have that $\mathsf{f}_Q(S^{n'+d+1}(r_0)) < 0$, a contradiction. Thus $\mathsf{f}' >_R 0$, which is enough to conclude. □

The next lemma shows that a positive operator is in between operators of the form $S^k(x)$, $k \in \mathbf{N}$. This is essentially a consequence of (Ax.9) and the triangle inequality.

Lemma 5.2.3. *Let $Q \in \mathrm{Pos}$. Then there exists $k_1 \leq k_2 \in \mathbf{N}$ such that*

$$T_{R,<} \models \forall x \in R(x \leq \mathsf{f}(S^{k_1}(x)) < S^{k_2}(x)).$$

Proof. Assume $Q(X) = \sum_{i=0}^d a_i X^i$. By (Ax.5), the operator $S(x)$ is strictly increasing on R. Thus for all $x \in R$, we have $\mathsf{f}_Q(x) \leq \sum_{i=0}^d |a_i| S^i(x) \leq a S^d(x)$, where $a = \sum_{i=0}^d |a_i|$. Let us show by induction on a that we can find $k_0 \in \mathbf{N}$ such that $S^{k_0}(x) - aS^d(x) >_R 0$. If $a = 1$, then we can choose $k_0 = d + 1$: for all $x \in R$, we have $S^d(x) < S(S^d(x)) = S^{d+1}(x)$. Now, assume that there exists $k_0 \in \mathbf{N}$ such that $S^{k_0}(x) - (a-1)S^d(x) >_R 0$. By (Ax.9), there exists $k \in \mathbf{N}$ such that $S^{k+k_0}(x) > (a-1)S^{d+k_0}(x) + x$ for all $x \in R$. Thus, for all $x \in R$, we have $S^{k+k_0}(x) > (a-1)S^{d+k_0}(x) + x > (a-1)x + x = ax$. Therefore, for all $x \in R$, we have $S^{k+k_0+d}(x) > aS^d(x)$, which is what we wanted. Thus there exists $k_0, n \in \mathbf{N}$ such that $S^{k_0}(x) - \mathsf{f}_Q(x) > 0$ for all $x \in R$. Also, by (Ax.9), there exists $k_1 \in \mathbf{N}$ such that $\mathsf{f}_Q(S^{k_1}(x)) > x$ for all $x \in R$. Therefore, setting $k_2 = k_1 + k_0$, we get what we wanted. □

For $Q \in \mathrm{Pos}$, we let $\min_R(\mathsf{f}_Q) = \min\{\mathsf{f}_Q(S^n(r_0)) \mid n \in \mathbf{N}\}$. Likewise, for $Q \in \mathrm{Neg}$, we let $\max_R(\mathsf{f}_Q) = \max\{\mathsf{f}_Q(S^n(r_0)) \mid n \in \mathbf{N}\}$. In the next lemma, we prove an analogue of (Ax.6) for arbitrary positive operators.

Lemma 5.2.4. *Let $Q \in \mathrm{Pos}$. Then*

$$T_{R,<} \models \forall x \exists y \in R(x \geq \min_R(\mathsf{f}_Q) \to \mathsf{f}_Q(y) \leq x < \mathsf{f}_Q(S(y))).$$

Proof. Let $\ell \in \mathbf{N}$ be given by Lemma 5.2.2 applied to Q and $k_1 \leq k_2 \in \mathbf{N}$ be given by Lemma 5.2.3 applied to Q. Let $y \geq \min_R(\mathsf{f}_Q)$. Let $k = \max\{\ell, k_2\}$ and assume that $y > S^k(r_0)$. By (Ax.6), there exists $x_0 \in R$ such that $x_0 \leq y < S(x_0)$. Because $x_0 \geq S^k(r_0)$, we have that $x_0 > \mathsf{f}_Q(S^{k_1-k_2}(x_0))$ and $S(x_0) \leq \mathsf{f}_Q(S^{k_1+1}(x_0))$. Thus, there exists $m \in [k_1 - k_2, k_1 + 1]$ such that $\mathsf{f}_Q(S^m(x_0)) \leq y < \mathsf{f}_Q(S^{m+1}(x_0))$. Now, for the case where $\min_R(\mathsf{f}_Q) \leq y \leq \mathsf{f}_Q(S^k(r_0))$, since

$$\bigwedge_{i=0}^{\ell-1} \mathsf{f}_Q(S^i(r_0)) < \mathsf{f}_Q(S^\ell(r_0))),$$

we can also find $m \in [0, k]$ such that $\mathsf{f}_Q(S^m(r_0)) \leq y < \mathsf{f}_Q(S^{m+1}(r_0))$. □

In the sequel, for $k \in \mathbf{N}$, we let $\mu_k(x, \bar{y})$ be the formula

$$\bigwedge_{i=1}^{|\bar{y}|} x > S^k(y_i) \text{ if } |\bar{y}| > 0 \text{ and } x > S^k(r_0) \text{ otherwise.}$$

These formulas capture the idea that x dominates the tuple $\bar{y}$. The following lemma gives an example of the usefulness of choosing a dominant variable.

Lemma 5.2.5 ([49, Lemma 2])**.** *Let $Q, Q_1, \ldots, Q_n \in \mathbf{Z}[X]$. Assume that $Q \in \mathrm{Pos}$. Then there exists $k \in \mathbf{N}$ such that*

$$T_{R,<} \models \forall x \in R \forall \bar{y} \in R\Big(\mu_k(x, \bar{y}) \to -\mathsf{f}_Q(x) < \sum_{i=1}^{n} \mathsf{f}_{Q_i}(y_i) < \mathsf{f}_Q(x)\Big).$$

Proof. Assume $Q_i(X) = \sum_{j=0}^{d} a_{ij} X^j$. Let $x, y_1, \ldots, y_n \in R$. Observe first that if $x > y_i$ for all $i \in [n]$, then, using (Ax.5), we have

$$\begin{aligned} \left| \sum_{i=1}^{n} \mathsf{f}_{Q_i}(y_i) \right| &\leq \sum_{i=1}^{n} \sum_{j=0}^{d} |a_{ij}| S^j(y_i) \\ &\leq \sum_{i=1}^{n} \left(\sum_{j=0}^{d} |a_{ij}| \right) S^d(y_i) \\ &< \sum_{i=1}^{n} \left(\sum_{j=0}^{d} |a_{ij}| \right) S^d(x) \\ &= \left(\sum_{i=1}^{n} \sum_{j=0}^{d} |a_{ij}| \right) S^d(x). \end{aligned}$$

Claim 5.2.6. For all $a, d \in \mathbf{N}$, there exists $k' \in \mathbf{N}$ such that $\mathsf{f}_Q(S^{k'}(x)) - aS^d(x) >_R 0$.

Proof of Claim. This is shown by induction on a. For $a = 0$, this follows from the fact that $Q \in \mathrm{Pos}$. Now assume that there exists $k_0 \in \mathbf{N}$ such that $\mathsf{f}_Q(S^{k_0}(x)) - (a-1)S^d(x) >_R 0$. By (Ax.9), there exists $k_1 \in \mathbf{N}$ such that $\mathsf{f}_Q(S^{k_1+k_0}(x)) > (a-1)S^{d+k_1}(x) + x$ for all $x \in R$. As a result, we have for all $x \in R$ that $\mathsf{f}_Q(S^{k_1+k_0}(x)) > (a-1)S^{d+k_1}(x) + x > ax$. In particular $\mathsf{f}_Q(S^{k_1+k_0+d}(x)) - aS^d(x) >_R 0$. □

Now, set

$$a = \sum_{i=1}^{n} \sum_{j=0}^{d} |a_{ij}|.$$

By Claim 5.2.6, there exists $k' \in \mathbf{N}$ such that $\mathsf{f}_Q(S^{k'}(x)) - aS^d(x) >_R 0$. Therefore there exists $k'' \in \mathbf{N}$ such that for all $x > S^{k''}(r_0)$, $\mathsf{f}_Q(S^{k'}(x)) > aS^d(x)$. Put $k = \max\{k', k''\}$. Then, for $x, y_1, \ldots, y_n \in R$ such that $\mu_k(x, \bar{y})$, we have

$$\begin{aligned} \left| \sum_{i=1}^{n} \mathsf{f}_{Q_i}(y_i) \right| &< aS^d(S^{-k}(x)) \\ &< \mathsf{f}_Q(S^k(S^{-k}(x))) \\ &= \mathsf{f}_Q(x), \end{aligned}$$

and this finishes the proof. □

5.3 Inequalities in $T_{R,<}$

In this section, we fix a model $\mathscr{M}$ of $T_{R,<}$ and $\bar{\mathsf{f}}$ a n-tuple of non-trivial operators. Given $a \in M$ and $\bar{b} \in R^n$, we shall identify how $\sum_{i=1}^{n} \mathsf{f}_i(b_i)$ compares to a. We shall see that we have to distinguish two cases: either there is a relation of the form $S^k(b_i) = b_j$ for some $i < j \in [n]$ and a small $k \in \mathbf{Z}$ or there exists $i \in [n]$ such that b_i dominates $\bar{b}^i$ in the sense that $\mu_k(b_i, \bar{b}^i)$ for some k that depends only on f_i. In the latter case, we use the lemmas from the previous section to reduce the comparison of $\sum_{i=1}^n \mathsf{f}_i(b_i)$ and a to a the comparison of $\sum_{i=1}^n \mathsf{f}_i(b_i)$ to a witness of a framing of a by the operator associated the dominant variable.

Let $i \in [n]$. Since we want to compare $\sum_{i=1}^n \mathsf{f}_i(b_i)$ and a, we may assume, at the cost of replacing a by $-a$ and $\bar{\mathsf{f}}$ by $-\bar{\mathsf{f}}$, that $\mathsf{f}_i >_R 0$. By Lemma 5.2.2, there exists $k_{0i} \in \mathbf{N}$ such that

$$T_{R,<} \models \forall x \in R(x \geq S^{k_{0i}}(r_0) \to \mathsf{f}_i(x) < \mathsf{f}_i(S(x)) \wedge \bigwedge_{j=0}^{k_{0i}-1} \mathsf{f}_i(S^j(r_0)) < \mathsf{f}_i(S^{k_{0i}}(r_0))).$$

Thus the operator $\mathsf{f}_i(S(x)) - \mathsf{f}_i(x)$ is positive on R. Thus by Lemma 5.2.5, there exists $k_{1i} \in \mathbf{N}$ such that

$$\begin{aligned} T_{R,<} \models \forall x \in R \forall \bar{y} \in R\Big(&\mu_{k_{1i}}(x, \bar{y}) \to \\ &- (\mathsf{f}_i(S(x)) - \mathsf{f}_i(x)) < \sum_{j \in [n] \setminus \{i\}} \mathsf{f}_j(y_j) < \mathsf{f}_i(S(x)) - \mathsf{f}_i(x)\Big). \end{aligned}$$

As a result, setting $k_i = \max\{k_{0i}, k_{1i}\} + 2$, we have that for all $\bar{b} \in R^n$, if $\mu_{k_i}(b_i, \bar{b}^i)$ holds, then three properties hold:

1. $\mathsf{f}_i(S^{-2}(b_i)) < \mathsf{f}_i(S^{-1}(b_i)) < \mathsf{f}_i(b_i)$;
2. $\mathsf{f}_i(b_i)$ is far from the minimal value of f_i on R:
$$\min_R(\mathsf{f}_i) \leq \mathsf{f}_i(S^{k_{0i}}(r_0)) < \mathsf{f}_i(S^{k_{0i}+2}(r_0)) \leq \mathsf{f}_i(b_i);$$
3. $\mathsf{f}_i(S^{-2}(b_i)) + \sum_{j \in [n] \setminus \{i\}} \mathsf{f}_j(b_j) < \mathsf{f}_i(S^{-1}(b_i)) < \mathsf{f}_i(b_i) + \sum_{j \in [n] \setminus \{i\}} \mathsf{f}_j(b_j)$.

On the other hand, a is in between two consecutive images of f_i or is below $\min_R(\mathsf{f}_i)$: by Lemma 5.2.4 there exists $b_0 \in R$ such that $\mathsf{f}_i(b_0) \leq a < \mathsf{f}_i(S(b_0))$ or $a < \min_R(\mathsf{f}_i)$. In case $a < \min_R(\mathsf{f}_i)$, we have that

$$\begin{aligned} a &< \min_R(\mathsf{f}_i) \\ &< \mathsf{f}_i(S^{-1}(b_i)) \\ &< \sum_{i=1}^n \mathsf{f}_i(b_i). \end{aligned}$$

Now assume that $\mathsf{f}_i(b_0) \leq a < \mathsf{f}_i(S(b_0))$. We distinguish three cases:

1. $b_i < b_0$. In that case we have

$$\begin{aligned}\sum_{i=1}^{n} \mathsf{f}_i(b_i) &< \mathsf{f}_i(S(b_i)) && \text{since } \mu_k(S^2(b_i), \bar{b}^i)\\ &\leq \mathsf{f}_i(b_0) && \text{since } S^{k_{0i}}(r_0) < S(b_i) \leq b_0\\ &\leq a.\end{aligned}$$

2. $b_i > S(b_0)$. In that case $b_i \geq S^2(b_0)$ and we have

$$\begin{aligned} a &< \mathsf{f}_i(S(b_0))\\ &\leq \mathsf{f}_i(S^{-1}(b_i)) && \text{since } b_i \geq S^2(b_0) \text{ and } S^{-2}(b_i) \geq S^{k_{0i}}(r_0)\\ &< \sum_{i=1}^{n} \mathsf{f}_i(b_i)\end{aligned}$$

3. either $b_i = b_0$ or $b_i = S(b_0)$. In that case we may restart our analysis to determine the relative positions of $\sum_{j\in[n]\setminus\{i\}} \mathsf{f}_j(b_j)$, $a - \mathsf{f}_i(b_0)$ and $\mathsf{f}_i(S(b_0)) - a$.

Now if for all $i \in [n]$, $\mu_{k_i}(b_i, \bar{b}^i)$ does not hold, then there are $i < j \in [n]$ and $k \in [-k_i, k_j]$ such that $b_i = S^k(b_j)$. In this case, we repeat our analysis with the operators f_ℓ for $\ell \in [n] \setminus \{i, j\}$ and the operator $\mathsf{f}_i(S^k(x)) + \mathsf{f}_j(x)$. We summarize this discussion in the following two results.

Lemma 5.3.1. *Let $Q, Q_1, \ldots, Q_n \in \mathbf{Z}[X]$. Assume $Q \in \mathrm{Pos}$. Then there exists $k \in \mathbf{N}$ such that*

$$\begin{aligned} T_{R,<} \models \forall x, \bar{x} \in R\Bigg(\mu_k(x, \bar{x}) \to \Big(&(\mathsf{f}_Q(S^{-2}(x)) < \mathsf{f}_Q(S^{-1}(x)) < \mathsf{f}_Q(x)) \wedge \min_R(\mathsf{f}_Q) < \mathsf{f}_Q(x)\\ &\wedge \mathsf{f}_Q(S^{-2}(x)) + \sum_{i\in[n]} \mathsf{f}_{Q_i}(x_i) < \mathsf{f}_Q(S^{-1}(x))\\ &\wedge \mathsf{f}_Q(S^{-1}(x)) < \mathsf{f}_Q(x) + \sum_{i\in[n]} \mathsf{f}_{Q_i}(x_i)\Big)\Bigg).\end{aligned}$$

Proposition 5.3.2. *Let $Q, Q_1, \ldots, Q_n \in \mathbf{Z}[X]$. Assume $Q \in \mathrm{Pos}$. Let $k \in \mathbf{N}$ be given by*

Lemma 5.3.1. Then

$$T_{R,<} \models \forall x, \bar{x} \in R \forall y \Bigg(\mu_k(x, \bar{x}) \to \Big(\mathsf{f}_Q(x) + \sum_{i=1}^{n} \mathsf{f}_{Q_i}(x_i) > y \leftrightarrow$$
$$\exists w \in R\big((y < \min_R(\mathsf{f}_Q) \vee \mathsf{f}_Q(w) \leq y < \mathsf{f}_Q(S(w)))$$
$$\wedge \big((x = S(w) \wedge \mathsf{f}_Q(S(w)) + \sum_{i=1}^{n} \mathsf{f}_{Q_i}(x_i) > y)$$
$$\vee (x = w \wedge \mathsf{f}_Q(w) + \sum_{i=1}^{n} \mathsf{f}_{Q_i}(x_i) > y)$$
$$\vee S(w) < x)\big)\big)\Big)\Bigg),$$

and

$$T_{R,<} \models \forall x, \bar{x} \in R \forall y \Bigg(\mu_k(x, \bar{x}) \to \Big(\mathsf{f}_Q(x) + \sum_{i=1}^{n} \mathsf{f}_{Q_i}(x_i) < y \leftrightarrow$$
$$\exists w \in R\big((\mathsf{f}_Q(w) \leq y < \mathsf{f}_Q(S(w)))) \wedge \big(x < w$$
$$\vee (x = S(w) \wedge \mathsf{f}_Q(S(w)) + \sum_{i=1}^{n} \mathsf{f}_{Q_i}(x_i) < y)$$
$$\vee (x = w \wedge \mathsf{f}_Q(w) + \sum_{i=1}^{n} \mathsf{f}_{Q_i}(x_i) < y)))\Big)\Bigg)$$

Even though we stated Proposition 5.3.2 under the assumption $Q \in \mathrm{Pos}$, it can be used to analyze the case where $Q \in \mathrm{Neg}$, considering $-Q, -Q_1, \ldots, -Q_k$ instead of $Q, Q_1, \ldots, Q_n$.

5.4 Negation of Im predicates

In this section, we establish that the negation of an Im predicate is equivalent to a disjunction of Im predicates. To this end, we need the following definition.

Definition 5.4.1. Let $\varphi(\bar{x}, \bar{y}, \bar{z})$ be a formula. We call $\varphi(\bar{x}, \bar{y}, \bar{z})$ *simple* if it satisfies one of the following conditions, where $n = |\bar{y}|$, $k = |\bar{x}|$:

1. if $k = 0$, it has the form

$$\bigwedge_{i=1}^{n} 0 > a_i y_i + b_i \wedge \varphi_0^R(\bar{z}),$$

   where $\bar{a}, \bar{b} \in \mathbf{Z}^n$ and $\varphi_0(\bar{z})$ is an $\mathcal{L}_0$-formula;

2. if $k > 0$, it has the form

$$\bigwedge_{i=1}^{n} \sum_{j=1}^{k} \mathsf{f}_{ij}(x_j) > a_i y_i + b_i \wedge \varphi_0^R(\bar{x}, \bar{z}),$$

where $\bar{a}, \bar{b} \in \mathbf{Z}^n$, f_{ij} is an operator for all $i \in [n]$ and $j \in [k]$ and $\varphi_0(\bar{x}, \bar{z})$ is an $\mathcal{L}_0$-formula.

Likewise a formula $\psi(\bar{y}, \bar{z})$ is called an *existential simple* formula if it is of the form $\exists \bar{x} \in R\, \varphi(\bar{x}, \bar{y}, \bar{z})$, where $\varphi(\bar{x}, \bar{y}, \bar{z})$ is a simple formula.

We note that the set of simple formulas and the set of existential simple formulas are closed under conjunction,

Observe that Im predicates are defined by simple existential formulas. We shall show by induction that the negation of simple existential formulas are equivalent to a disjunction of simple existential formulas. The induction step is summarized in the following lemma.

Lemma 5.4.2 ([49, Lemma 3]). *Let $\varphi(\bar{x}, \bar{y}, \bar{z})$ be a simple formula, $n = |\bar{x}|$. Let $\bar{k} \in \mathbf{N}^n$. Then there exists a simple formula $\varphi'(\bar{w}, \bar{y}, \bar{z})$, $|\bar{w}| = n - 1$, such that*

$$T_{R,<} \models \forall \bar{y}\, \forall \bar{z} \in R\Big((\exists \bar{x} \in R\, \varphi(\bar{x}, \bar{y}, \bar{z})) \leftrightarrow \exists \bar{w} \in R\, \varphi'(\bar{w}, \bar{y}, \bar{z}) \vee \bigvee_{i=1}^{n} \exists \bar{x} \in R(\mu_{k_i}(x_i, \bar{x}^i) \wedge \varphi(\bar{x}, \bar{y}, \bar{z}))\Big).$$

Proof. We already observed in Section 5.3 that if none of the relations $\mu_{k_i}(x_i, \bar{x}^i)$ is satisfied, then there exists $i, j \in [n]$ such that $x_i = S^k(x_j)$ for some

$$k \in [-k_i, k_j].$$

Therefore we define $\varphi'(\bar{w}, \bar{y}, \bar{z})$ to be the disjunction of all substitutions in $\varphi(\bar{x}, \bar{y}, \bar{z})$ of x_i by $S^k(w_j)$, with $i \in [n]$, $j \in [n-1]$ and $k \in [-\max\{k_i \mid i \in [n]\}, \max\{k_i \mid i \in [n]\}]$. □

We are now ready to prove the main result of this section.

Theorem 5.4.3. *Let $\varphi(\bar{x}, \bar{y}, \bar{z})$ be a simple formula. Then there exist simple formulas $\varphi_1(\bar{w}, \bar{y}, \bar{z})$, $\ldots, \varphi_\ell(\bar{w}, \bar{y}, \bar{z})$ such that*

$$T_{R,<} \models \forall \bar{y} \forall \bar{z} \in R\Big(\neg(\exists \bar{x} \in R\, \varphi(\bar{x}, \bar{y}, \bar{z})) \leftrightarrow \bigvee_{i=1}^{\ell} \exists \bar{w} \in R\, \varphi_i(\bar{w}, \bar{y}, \bar{z})\Big).$$

Proof. The proof is done by induction on $k = |\bar{x}|$. If $k = 0$, we use the fact that $T_R \models \forall z(\neg(0 > z) \leftrightarrow 0 > -z + 1)$, to show that, given $\bar{a}, \bar{b} \in \mathbf{Z}^n$ and $\varphi_0(\bar{z})$ a $\mathcal{L}_0$-formula,

$$T_R \models \forall \bar{y} \forall \bar{z} \in R\Big(\neg(\bigwedge_{i=1}^{n} 0 > a_i y_i + b_i \wedge \varphi_0^R(\bar{z})) \leftrightarrow \bigvee_{i=1}^{n} 0 > -a_i y_i - b_i + 1 \vee \neg\varphi_0^R(\bar{z})\Big).$$

This concludes the case $k = 0$ since

1. $0 > -a_i y_i - b_i + 1$ is equivalent to the simple formula

$$0 > -a_i y_i - b_i + 1 \wedge \bigwedge_{\substack{j=1 \\ j \neq i}}^{n} 0 > 0 y_j - 1 \wedge r_0 = r_0;$$

2. and the formula $\neg\varphi_0^R(\bar{z})$ is equivalent to the simple formula

$$\bigwedge_{i=1}^{n} 0 > 0 y_i - 1 \wedge \neg\varphi_0^R(\bar{z});$$

Now assume that $k > 0$ and that the theorem holds true for all simple formulas $\varphi(\bar{x}, \bar{y}, \bar{z})$ with $|\bar{x}| < k$. Let $\psi(\bar{x}, \bar{y}, \bar{z})$ be a simple formula with $|\bar{x}| = k$, so that $\psi(\bar{x}, \bar{y}, \bar{z})$ has the form

$$\bigwedge_{i=1}^{n} \sum_{j=1}^{k} \mathsf{f}_{ij}(x_j) > a_i y_i + b_i \wedge \varphi_0^R(\bar{x}, \bar{z}),$$

where $\bar{a}, \bar{b} \in \mathbf{Z}^n$, f_{ij} is an operator for all $i \in [n]$ and $j \in [k]$ and $\varphi_0(\bar{x}, \bar{z})$ is an $\mathcal{L}_0$-formula. We want to show that $\neg\big(\exists \bar{x} \in R\, \psi(\bar{x}, \bar{y}, \bar{z})\big)$ is equivalent to a disjunction of existential simple formulas. We may assume that for all $j \in [k]$ there exists $i \in [n]$ such that $\mathsf{f}_{ij} \neq_R 0$. Indeed, otherwise $\psi(\bar{x}, \bar{y}, \bar{z})$ is equivalent, for some $j_0 \in [k]$, to

$$\bigwedge_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq j_0}}^{k} \mathsf{f}_{ij}(x_j) > a_i y_i + b_i \wedge \exists x_{j_0} \in R\, \varphi_0^R(\bar{x}, \bar{z}),$$

So that by the induction hypothesis, $\neg\big(\exists \bar{x} \in R\, \psi(\bar{x}, \bar{y}, \bar{z})\big)$ is equivalent to a disjunction of existential simple formulas. Let us then assume that for all $j \in [k]$ there exists $i \in [n]$ such that $\mathsf{f}_{ij} \neq_R 0$.

For all $j \in [k]$, let us partition $[n]$ according to whether f_{ij} is positive, negative or trivial. Namely, let $I_{j1}, I_{j2} \subset [n]$ be such that $i \in I_{j1}$ if and only if $\mathsf{f}_{ij} >_R 0$ and $i \in I_{j2}$ if and only if $\mathsf{f}_{ij} <_R 0$. Let $I_{j3} = [k] \setminus (I_{j1} \cup I_{j2})$. Note that by assumption $I_{j1} \cup I_{j2} \neq \emptyset$ for all $j \in [k]$. Now for all $j \in [k]$ and all $i \in I_{j1} \cup I_{j2}$ let k_{ij} be given by Lemma 5.3.1 applied to the polynomials defining $\bar{\mathsf{f}}_i$ if $i \in I_{j1}$ and $-\bar{\mathsf{f}}_i$ if $i \in I_{j2}$. Set

$k = \max\{k_{ij} \mid j \in [k] \text{ and } i \in I_{j1} \cup I_{j2}\}$. By Lemma 5.4.2, we have that $\exists \bar{x} \in R\, \psi(\bar{x}, \bar{y}, \bar{z})$ is equivalent to

$$\exists \bar{w} \in R\, \psi'(\bar{w}, \bar{y}, \bar{z}) \vee \bigvee_{j=1}^{k} \exists \bar{x} \in R\big(\mu_k(x_j, \bar{x}^j) \wedge \psi(\bar{x}, \bar{y}, \bar{z})\big),$$

where $\psi'(\bar{w}, \bar{y}, \bar{z})$ is a simple formula with $|\bar{w}| = k - 1$. Thus, by our induction hypothesis, we need to prove the theorem for the formulas

$$\exists \bar{x} \in R\big(\mu_k(x_j, \bar{x}^j) \wedge \psi(\bar{x}, \bar{y}, \bar{z})\big),$$

for all $j \in [k]$. Let $\ell \in [k]$. For all $i \in I_{\ell 1}$, let $m_{\ell i} = \min_R(\mathsf{f}_{i\ell})$. Recall from Proposition 5.3.2 that the inequalities can be treated as follows, under the assumption $\mu_k(x_\ell, \bar{x}^\ell)$:

1. for $i \in I_{\ell 1}$,

$$\sum_{j=1}^{k} \mathsf{f}_{ij}(x_j) > a_i y_i + b_i$$

is equivalent to the disjunction of $a_i y_i + b_i < m_{\ell i}$ and

$$\begin{aligned} \exists w_i \in R\Big(& \mathsf{f}_{i\ell}(w_i) \le a_i y_i + b_i < \mathsf{f}_{i\ell}(S(w_i)) \\ & \wedge\Big(S(w_i) < x_\ell \vee \big(x_\ell = S(w_i) \wedge \mathsf{f}_{i\ell}(S(w_i)) + \sum_{\substack{j=1 \\ j\neq\ell}}^{k} \mathsf{f}_{ij}(x_j) > a_i y_i + b_i\big) \\ & \vee \big(x_\ell = w_i \wedge \mathsf{f}_{i\ell}(w_i) + \sum_{\substack{j=1 \\ j\neq\ell}}^{k} \mathsf{f}_{ij}(x_j) > a_i y_i + b_i\big)\Big)\Big); \end{aligned}$$

2. and likewise for $i \in I_{j2}$:

$$\sum_{j=1}^{k} \mathsf{f}_{ij}(x_j) > a_i y_i + b_i$$

is equivalent to

$$\begin{aligned} \exists w_i \in R\Big(& -\mathsf{f}_{i\ell}(w_i) \le -a_i y_i - b_i < -\mathsf{f}_{i\ell}(S(w_i)) \\ & \wedge\Big(x_\ell < w_i \vee \big(x_\ell = S(w_i) \wedge \mathsf{f}_{i\ell}(S(w_i)) + \sum_{\substack{j=1 \\ j\neq\ell}}^{k} \mathsf{f}_{ij}(x_j) > a_i y_i + b_i\big) \\ & \vee \big(x_\ell = w_i \wedge \mathsf{f}_{i\ell}(w_i) + \sum_{\substack{j=1 \\ j\neq\ell}}^{k} \mathsf{f}_{ij}(x_j) > a_i y_i + b_i\big)\Big)\Big). \end{aligned}$$

Therefore $\exists \bar{x} \in R\big(\mu_k(x_\ell, \bar{x}^\ell) \wedge \psi(\bar{x}, \bar{y}, \bar{z})\big)$ is equivalent to

$$\exists w_1, \ldots, w_n \in R\Bigg(\bigwedge_{i \in I_{\ell 1}} \Big(a_i y_i + b_i < m_{\ell i} \vee \big(\mathsf{f}_{ij}(w_i) \leq a_i y_i + b_i < \mathsf{f}_{ij}(S(w_i))\big)\Big)$$
$$\wedge \bigwedge_{i \in I_{\ell 2}} -\mathsf{f}_{ij}(w_i) \leq -a_i y_i - b_i < -\mathsf{f}_{ij}(S(w_i))$$
$$\wedge \forall \bar{x} \in R\Big(\bigwedge_{i \in I_{\ell 1}} S(w_i) < x_\ell \wedge \bigwedge_{i \in I_{\ell 2}} x_\ell < w_i \rightarrow \neg \varphi_0^R(\bar{x}, \bar{z})\Big)$$
$$\wedge \bigvee_{\substack{i \in I_{\ell 1} \cup I_{\ell 2} \\ \epsilon \in \{0,1\}}} \neg \exists \bar{x}^\ell \in R\, \psi(x_1, \ldots, x_{\ell-1}, S^\epsilon(w_i), x_{\ell+1}, \ldots, x_k, \bar{y}, \bar{z}) \Bigg).$$

Notice that

$$\forall \bar{x}\Big(\bigwedge_{i \in I_{\ell 1}} S(w_i) < x_\ell \wedge \bigwedge_{i \in I_{\ell 2}} x_\ell < w_i \rightarrow \neg \varphi_0(\bar{x}, \bar{z})\Big)$$

is an $\mathcal{L}_0$-formula. Therefore, what remains to be shown is that for all $(i, \epsilon) \in I_{\ell 1} \cup I_{\ell 2} \times \{0, 1\}$, the formula

$$\neg \exists \bar{x}^\ell \in R\, \psi(x_1, \ldots, x_{\ell-1}, S^\epsilon(w_i), x_{\ell+1}, \ldots, x_k, \bar{y}, \bar{z})$$

is equivalent to a disjunction of existential simple formulas.

Consider the formula $\psi'(\bar{x}^\ell, \bar{y}, z_0, \bar{z})$ defined by

$$\bigwedge_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq \ell}}^{k} \mathsf{f}_{ij}(x_j) > y_i \wedge \varphi_0(x_1, \ldots, x_{\ell-1}, z_0, x_{\ell+1}, \ldots, x_k, \bar{z}).$$

Notice that $\psi(x_1, \ldots, x_{\ell-1}, S^\epsilon(w_i), x_{\ell+1}, \ldots, x_k, \bar{y}, \bar{z})$ is equivalent to

$$\psi'(\bar{x}^\ell, t_i(y_i, S^\epsilon(w_i)), \ldots, t_n(y_n, S^\epsilon(w_n)), S^\epsilon(w_i), \bar{z}),$$

where for all $i \in [n]$, $t_i(y_i, S^\epsilon(w_i)) = a_i y_i + b_i - \mathsf{f}_{i\ell}(S^\epsilon(w_i))$.

By the induction hypothesis, there exists simple formulas $\psi_i'(\bar{u}, \bar{y}, z_0, \bar{z})$, $i \in [m]$, such that

$$T_{R,<} \models \forall \bar{y} \forall z_0, \bar{z} \in R\Big(\neg\big(\exists \bar{x}^\ell \in R\, \psi'(\bar{x}^\ell, \bar{y}, z_0, \bar{z})\big) \leftrightarrow \bigvee_{i=1}^{m} \exists \bar{u} \in R \psi_i'(\bar{u}, \bar{y}, z_0, \bar{z})\Big).$$

Thus for all $(i, \epsilon) \in I_{\ell 1} \cup I_{\ell 2} \times \{0, 1\}$, the formula

$$\neg \exists \bar{x}^\ell \in R\, \psi(x_1, \ldots, x_{\ell-1}, S^\epsilon(w_i), x_{\ell+1}, \ldots, x_k, \bar{y}, \bar{z})$$

is equivalent to a disjunction of existential simple formulas. This concludes the proof. □

As a corollary of the previous theorem, we get that negations of Im predicates are equivalent to a disjunction of Im predicates.

Corollary 5.4.4. *Let* $[Q] \in \mathbf{Z}^{n\times m}$ *and* $\varphi(\bar{x})$ *be an* $\mathcal{L}_0$*-formula,* $|\bar{x}| = n$*. Then there exist* $[Q_1], \dots, [Q_\ell] \in \mathbf{Z}^{n'\times m}$ *and* $\mathcal{L}_0$*-formulas* $\varphi_1(\bar{z}), \dots, \varphi_\ell(\bar{z})$ *and terms* $t_1(y_1), \dots, t_m(y_m)$ *such that*

$$T_{R,<} \models \forall \bar{y} \in R\Big(\neg \mathrm{Im}_{[Q],\varphi}(\bar{y}) \leftrightarrow \bigvee_{i=1}^{\ell} \mathrm{Im}_{[Q_i],\varphi_i}(t_1(y_1), \dots, t_m(y_m))\Big).$$

Proof. This follows from Theorem 5.4.3 by noticing that Im predicates are defined by existential simple formulas and that an existential simple formula with $|\bar{z}| = 0$ defines an Im predicate of the shape

$$\mathrm{Im}_{[Q_i],\varphi_i}(t_1(y_1), \dots, t_m(y_m)).$$

□

We end this section by the observation that a conjunction of Im predicates is an Im predicate.

Lemma 5.4.5. *For all* $[Q_1] \in \mathbf{Z}^{n_1\times m_1}, \dots, [Q_\ell] \in \mathbf{Z}^{n_\ell\times m_\ell}$ *and* $\mathcal{L}_0$*-formulas* $\varphi_{01}, \dots, \varphi_{0\ell}$*, there exists* $[Q] \in \mathbf{Z}^{(n_1\cdots n_\ell)\times(m_1\cdots m_\ell)}$ *and* φ_0 *an* $\mathcal{L}_0$*-formula such that*

$$T_{R,<} \models \forall \bar{y}_1, \dots, \bar{y}_\ell \left(\bigwedge_{i\in[\ell]} \mathrm{Im}_{[Q]_i,\varphi_{0i}}(\bar{y}_i) \leftrightarrow \mathrm{Im}_{[Q],\varphi_0}(\bar{y}_1, \dots, \bar{y}_\ell) \right).$$

Proof. This is similar to the proof of Lemma 4.4.4: we take $[Q] = [Q]_1 \oplus \cdots \oplus [Q]_\ell$ and φ_0 the disjunction of the φ_{0i} after renaming variables when necessary. □

5.5 Quantifier elimination for $T_{R,<}$

Before we prove our quantifier elimination result, we need a lemma on terms in $\mathcal{L}$. This lemma will allow us to keep track of the shape of terms in $\mathcal{L}$ using Im predicates. More precisely, we shall use the fact that for a term $t(x, \bar{y})$, the formula $\neg(S(t(x, \bar{y})) = t(x, \bar{y}))$ is equivalent to $\exists z \in R((z > t(x, \bar{y}) - 1) \wedge (-z > -t(x, \bar{y}) - 1))$, which defines an Im predicate, and likewise $S(t(x, \bar{y})) = t(x, \bar{y})$ is equivalent to $r_0 > t(x, \bar{y}) \vee \exists z \in R((S(z) > t(x, \bar{y})) \wedge (-z > -t(x, \bar{y})))$, which defines a disjunction of Im predicates.

Proposition 5.5.1. *Let $t(x,\bar{y})$ be a term, $\bar{\mathsf{f}}$ a n-tuple and $m \in \mathbf{N}$. Then there exist operators $\bar{\mathsf{f}}_1,\dots,\bar{\mathsf{f}}_k$, $\psi_1(\bar{w},y,\bar{y},\bar{z}_1),\dots,\psi_k(\bar{w},y,\bar{y},\bar{z}_k)$ existential simple formulas, $\bar{n}' \in \mathbf{Z}^k$ and $t'_1(\bar{y}),\dots,t'_k(\bar{y})$ such that for all $\mathscr{M} \models T_{R,<}$, $\mathscr{A} \subset \mathscr{M}$ and $b \in M$, if there exist $\bar{b} \in R(M)^n$ and $a \in A$ such that $mb = \mathsf{f}_1(b_1) + \cdots + \mathsf{f}_n(b_n) + a$, then for all $\bar{a} \in A^{|\bar{y}|}$, there exist a unique $i \in [k]$ and $\bar{d} \in R(M)^{|\bar{\mathsf{f}}_i|}$ such that*

1. $\mathscr{M} \models \psi_i(\bar{b},a,\bar{a},\bar{d})$;
2. $t(b,\bar{a}) = \mathsf{f}_{i1}(d_1) + \cdots + \mathsf{f}_{i\ell}(d_\ell) + n'_i b + t'_i(\bar{a})$.

Proof. The proof is by induction on the number ℓ of occurrences of S and S^{-1} in $t(x,\bar{y})$. If $\ell = 0$, then $t(x,\bar{y})$ is an $\mathcal{L}_g$-term and is of the form $n'x + t'(\bar{y})$, where $t'(\bar{y})$ is an $\mathcal{L}_g$-term. Now if $\ell > 0$, then there are terms $t_0(x,\bar{y}),\dots,t_{k'}(x,\bar{y})$ with $< \ell$ occurrences of the symbols S and S^{-1} and $\epsilon_1,\dots,\epsilon_{k'} \in \{-1,1\}$ such that

$$t(x,\bar{y}) = t_0(x,\bar{y}) + \sum_{i=1}^{k'} S^{\epsilon_i}(t_i(x,\bar{y})).$$

As a result, in order to finish the proof, we only need to consider the case where $t(x,\bar{y}) = S^{\epsilon}(t_1(x,\bar{y}))$, where $\epsilon \in \{-1,1\}$ and $t_1(x,\bar{y})$ is a term with $\ell - 1$ occurrences of the symbols S and S^{-1}. We treat the case $\epsilon = 1$, the other being similar. Now two things can happen: either $S(t_1(x,\bar{y})) \neq t_1(x,\bar{y})$ or not. The first case happens if and only if

$$\exists z \in R(z > t_1(x,\bar{y}) - 1 \wedge -z > -t_1(x,\bar{y}) - 1)$$

and the other if and only if

$$r_0 > t_1(x,\bar{y}) \vee (\exists z \in R(S(z) > t_1(x,\bar{y}) \wedge -z > -t_1(x,\bar{y}))).$$

By the induction hypothesis, there are $\bar{\mathsf{f}}_1,\dots,\bar{\mathsf{f}}_k$ and $\psi_1(\bar{w},y,\bar{y},\bar{z}_1),\dots,\psi_k(\bar{w},y,\bar{y},\bar{z}_k)$ existential simple formulas, $\bar{n}' \in \mathbf{Z}^k$ and $t'_1(\bar{y}),\dots,t'_k(\bar{y})$ such that for all $\mathscr{M} \models T_{R,<}$, $\mathscr{A} \subset \mathscr{M}$ and $b \in M$, if there exist $\bar{b} \in R(M)^n$ and $a \in A$ such that $mb = \mathsf{f}_1(b_1) + \cdots + \mathsf{f}_n(b_n) + a$, then for all $\bar{a} \in A^{|\bar{y}|}$, there exists a unique $i \in [k]$ and $\bar{d} \in R(M)^{|\bar{\mathsf{f}}_i|}$ such that

1. $\mathscr{M} \models \psi_i(\bar{b},a,\bar{a},\bar{d})$;
2. $t_1(b,\bar{a}) = \mathsf{f}_{i1}(d_1) + \cdots + \mathsf{f}_{i\ell}(d_\ell) + n'_i b + t'_i(\bar{a})$.

Let $\chi_{1i}(\bar{w},y,\bar{y},z,\bar{z})$ be the formula

$$mz > \sum_{j=1}^{|\bar{\mathsf{f}}_i|} m\mathsf{f}_{ij}(z_j) + n'_i\Big(\sum_{j=1}^{n} \mathsf{f}_j(w_j) + y\Big) + mt'_i(\bar{y}) - m$$
$$\wedge -mz > -\left(m\sum_{j=1}^{|\bar{\mathsf{f}}_i|} \mathsf{f}_{ij}(z_j) + n'_i\Big(\sum_{j=1}^{n} \mathsf{f}_j(w_j) + y\Big) + mt'_i(\bar{y})\right) - m,$$

$\chi_{2i}(\bar{w},y,\bar{y},z,\bar{z})$ be the formula

$$mr_0 > m\sum_{j=1}^{|\bar{\mathsf{f}}_i|}\mathsf{f}_{ij}(z_j) + n'_i\Big(\sum_{j=1}^{n}\mathsf{f}_j(w_j)+y\Big) + mt'_i(\bar{y}),$$

and $\chi_{3i}(\bar{w},y,\bar{y},z,\bar{z})$ be the formula

$$\begin{aligned} mS(z) >& m\sum_{j=1}^{|\bar{\mathsf{f}}_i|}\mathsf{f}_{ij}(z_j) + n'_i\Big(\sum_{j=1}^{n}\mathsf{f}_j(w_j)+y\Big) + mt'_i(\bar{y}) \\ &\wedge -mz > -\left(m\sum_{j=1}^{|\bar{\mathsf{f}}_i|}\mathsf{f}_{ij}(z_j) + n'_i\Big(\sum_{j=1}^{n}\mathsf{f}_j(w_j)+y\Big) + mt'_i(\bar{y})\right). \end{aligned}$$

Now set $\tau_{ji}(\bar{w},y,\bar{y},z)$ to be the formula $\exists z \in R(\chi_{ji}(\bar{w},y,\bar{y},z,\bar{z}) \wedge \varphi_i(\bar{w},y,\bar{y},\bar{z}))$, for $(j,i) \in [3]\times[k]$. Now one checks that the formulas τ_{ji}, $(j,i)\in[3]\times[k]$, n' and $t'(\bar{y})$ satisfy the proposition's statement for $S(t_1(x,\bar{y}))$. □

In the following lemma, we show that we can eliminate the quantifier $\exists\bar{x}\in R$.

Lemma 5.5.2. *Let $\varphi(\bar{x},\bar{y})$ be a formula of the form*

$$\bigwedge_{i\in I_1} t_i(\bar{x},\bar{y}) > 0 \wedge \bigwedge_{i\in I_2} D_{m_i}(t_i(\bar{x},\bar{y})) \wedge \mathrm{Im}_{[Q],\varphi_0}(t_1(\bar{x},\bar{y}),\dots,t_n(\bar{x},\bar{y})) \wedge \psi(\bar{x},\bar{y}),$$

where $\psi(\bar{x},\bar{y})$ is an existential simple formula and for all $i \in I_1\cup I_2\cup[n]$, $t_i(x,\bar{y})$ is a term of the form $\mathsf{f}_{i1}(x_1)+\dots+\mathsf{f}_{i\ell}(x_\ell)+t'_i(\bar{y})$, $\ell=|\bar{x}|$, $\bar{\mathsf{f}}_i$ a n-tuple of operators and $t'_i(\bar{y})$ is a term, $k=|\bar{y}|$, $[Q]\in\mathbf{Z}^{k\times m}$, $\varphi_0(\bar{z})$ is an $\mathcal{L}_0$-formula with $m=|\bar{z}|$ and $\bar{m}\in(\mathbf{N}^{>1})^{|I_2|}$. Then for all $\mathscr{M},\mathscr{N}\models T_{R,<}$, $\mathscr{A}$ a common substructure and $\bar{a}\in A^m$, if $\mathscr{M}\models\exists\bar{x}\in R\varphi(\bar{x},\bar{a})$ then $\mathscr{N}\models\exists\bar{x}\in R\varphi(\bar{x},\bar{a})$.

Proof. The idea is to show that $\exists\bar{x}\in R\varphi(\bar{x},\bar{y})$ defines a conjunction of an Im predicate and a conjunction of divisibility conditions on terms with variables among $\bar{y}$. The only difficulty is the treatment of the conditions $D_{m_i}(t_i(\bar{x},\bar{y}))$. But our assumption on the terms allow us to say that $D_{m_i}(t_i(\bar{x},\bar{y}))$ is equivalent to a disjunction of conjunctions of formulas of the form $D_{m'}(S^k(x_i)+k')$ and $D_{m'}(t_i(\bar{y})+k')$. This is enough to conclude since $D_{m'}(S^k(x_i)+k')$ is $(D_{m',k'}(S^k(x_i)))^R$. □

Theorem 5.5.3. *Let R be a sparse set. Then $T_{R,<}$ has quantifier elimination.*

Proof. We apply the usual quantifier elimination criteria [33, Corollary 3.1.6]: we shall show that for all $\mathscr{M},\mathscr{N}\models T_{R,<}$, all common substructure $\mathscr{A}$, all quantifier-free formula $\varphi(x,\bar{y})$ and all $\bar{a}\in A^{|\bar{y}|}$, if there exists $b\in M$ such that $\mathscr{M}\models\varphi(b,\bar{a})$, then there exists $b'\in N$ such that $\mathscr{N}\models\varphi(b',\bar{a})$.

Let $\mathscr{M}, \mathscr{N} \models T_{R,<}$ and $\mathscr{A}$ a common substructure. Let $\varphi(x, \bar{y})$ be a quantifier-free formula, $\bar{a} \in A^{|\bar{y}|}$ and $b \in M$ such that $\mathscr{M} \models \varphi(b, \bar{a})$. Using Corollary 5.4 and Lemma 5.4.5, the fact that $x = y$ is equivalent to $(x < y + 1) \wedge (y < x + 1)$ and the fact that a negation of congruence relation is equivalent to a disjunction of such, we may assume that $\varphi(x, \bar{y})$ is of the form

$$\bigwedge_{i \in I_1} t_i(x, \bar{y}) > 0 \wedge \bigwedge_{i \in I_2} D_{m_i}(t_i(x, \bar{y})) \wedge \mathrm{Im}_{[Q], \varphi_0}(t_1(x, \bar{y}), \ldots, t_n(x, \bar{y})),$$

where for all $i \in I_1 \cup I_2 \cup [n]$ $t_i(x, \bar{y})$ is a term, $n = |\bar{y}|$, $[Q] \in \mathbf{Z}^{n \times m}$, $\varphi_0(\bar{z})$ is an $\mathcal{L}_0$-formula with $m = |\bar{z}|$ and $\bar{m} \in (\mathbf{N}^{>1})^{|I_2|}$. We now distinguish two cases: $b \in \operatorname{div}\langle R(M), A\rangle$ or not.

Assume first that $b \in \operatorname{div}\langle R(M), A\rangle$. In that case there exists $m \in \mathbf{Z}$, a n-tuple $\bar{\mathsf{f}}$ of operators, $\bar{b} \in R^n$ and $a \in A$ such that $mb = \mathsf{f}_1(b_1) + \cdots + \mathsf{f}_n(b_n) + a$. Therefore, we may apply Proposition 5.5.1 to all terms involved in φ: for $i \in I_1 \cup I_2 \cup [n]$, let $\psi_i(\bar{w}, y, \bar{y}, \bar{z}_i)$, $\bar{\mathsf{f}}_i$, $t_i(\bar{y})$, $n_i \in \mathbf{N}$ and $\bar{d}_i \in R^{|\bar{\mathsf{f}}_i|}$ such that $\mathscr{M} \models \psi_i(\bar{b}, a, \bar{a}, \bar{d}_i)$ and $mt(b, \bar{a}) = m\mathsf{f}_{i1}(d_1) + \cdots + m\mathsf{f}_{i\ell}(d_\ell) + \mathsf{f}_1(b_1) + \cdots + \mathsf{f}_n(b_n) + a + mt_i(\bar{a})$. Let $t_i(\bar{w}, \bar{z}_i, y, \bar{y})$ be the term

$$m\mathsf{f}_{i1}(z_{i1}) + \cdots + m\mathsf{f}_{i\ell}(z_{i\ell}) + \mathsf{f}_1(w_1) + \cdots + \mathsf{f}_n(w_n) + y + mt_i(\bar{y}).$$

Let $\psi(\bar{w}, y, \bar{y}, \bar{z})$ be

$$\bigwedge_{i \in I_1 \cup I_2 \cup [n]} \psi_i(\bar{w}, y, \bar{y}, \bar{z}_i).$$

Thus instead of $\varphi(x, \bar{y})$, me way consider the formula $\tilde{\varphi}(\bar{w}, \bar{z}, y, \bar{y})$

$$\bigwedge_{i \in I_1} t_i(\bar{w}, \bar{z}_i, y, \bar{y}) > 0 \wedge \bigwedge_{i \in I_2} D_{mm_i}(t_i(\bar{w}, \bar{z}_i, y, \bar{y}))$$
$$\wedge \mathrm{Im}_{m[Q], \varphi_0}(t_1(\bar{w}, \bar{z}_i, y, \bar{y}), \ldots, t_n(\bar{w}, \bar{z}_i, y, \bar{y})) \wedge \psi(\bar{w}, y, \bar{y}, \bar{z}).$$

But by Lemma 5.5.2, we have $\mathscr{N} \models \exists \bar{w}, \bar{z} \in R \tilde{\varphi}(\bar{w}, \bar{z}, a, \bar{a})$. But then, if $\bar{b}' \in R^n$ and $\bar{d}' \in R^{|\bar{z}|}$ witness this fact, we get, by Proposition 5.5.1 that $b' = \mathsf{f}_1(b'_1) + \cdots + \mathsf{f}_n(b'_n) + a$ satisfies $\varphi(x, \bar{a})$ in $\mathscr{N}$.

Let us now assume that $b \notin \operatorname{div}\langle R(M), A\rangle$. Then, given $n \in \mathbf{Z}$ and $a \in A$, we have $S(nb + a) = nb + a$, unless $n = 0$ and $a \in R(M)$. Indeed, if $S(nb + a) \neq nb + a$, then there exists $d \in R(M)$ such that $nb = d - a$, hence $n = 0$. As a result, for all term $t(x, \bar{y})$ and $\bar{a} \in A^{|\bar{y}|}$, there exists $n \in \mathbf{Z}$ and $a \in A$ such that $t(b, \bar{a}) = nb + a$, and we can keep track of this information as in the first case. Therefore, we may assume that $\varphi(b, \bar{a})$ is of the form

$$\bigwedge_{i \in I_1} n_i b + a'_i > 0 \wedge \bigwedge_{i \in I_2} D_{m_i}(n_i b + k_i) \wedge \mathrm{Im}_{[Q], \varphi_0}(\ell_1 b + a'_1, \ldots, \ell_n b + a'_n).$$

We may also assume that $|I_2| = 1$ and up to multiplying each inequalities that appear in φ and the Im predicates, me may further assume that φ is of the form

$$\bigwedge_{i \in I_1} \ell b + a'_i > 0 \wedge \bigwedge_{i \in I'_1} \ell b + a'_i < 0 \wedge D_m(\ell b + k) \wedge \mathrm{Im}_{[Q],\varphi_0}(\epsilon_1 \ell b + a'_1, \dots, \epsilon_n \ell b + a'_n),$$

where $\epsilon_i \in \{-1, 1\}$ for all $i \in [n]$.

Let $\bar{b} \in R^{|\bar{Q}_1|}$ be a witness of $\mathrm{Im}_{[Q],\varphi_0}(\epsilon_1 \ell b + a'_1, \dots, \epsilon_n \ell b + a'_n)$, that is

$$\bigwedge_{j=1}^{n} \sum_{i=1}^{|\bar{Q}_1|} \mathsf{f}_{Q_{ij}}(b_i) > \epsilon_j \ell b + a'_j \wedge \varphi_0(\bar{b}).$$

Let g_1 be the maximum of the set

$$\{-a'_i \mid i \in I_1\} \cup \{a_j - \sum_{i=1}^{|\bar{Q}_1|} \mathsf{f}_{Q_{ij}}(b_i) \mid j \in [n], \epsilon_j = -1\}$$

and g_2 be the minimum of the set

$$\{-a'_i \mid i \in I_2\} \cup \{\sum_{i=1}^{|\bar{Q}_1|} \mathsf{f}_{Q_{ij}}(b_i) - a_j \mid j \in [n], \epsilon_j = 1\}.$$

Since we assumed that $b \notin \mathrm{div}\langle R(M), A\rangle$ we have that for all $s \in \mathbf{N}$, $g_1 + s < g_2 - s$. All this can be captured by a conjunction of inequalities $\tau_s(\bar{x}, \bar{a}')$ so that for all $s \in \mathbf{N}$, $g_1 + s < g_2 - s$ if and only if $\tau_s(\bar{b}, \bar{a})$. Now we consider the formula

$$\exists \bar{x} \in R(\tau_s(\bar{x}, \bar{y}) \wedge \varphi_0(\bar{b})).$$

This formula defines an Im predicate $\mathrm{Im}_{[Q'],\varphi_0,s}(\bar{y})$ and we have $\mathscr{M} \models \mathrm{Im}_{[Q'],\varphi_0,s}(\bar{a})$. Therefore, we have that for all $s \in \mathbf{N}$, $\mathscr{N} \models \mathrm{Im}_{[Q'],\varphi_0,s}(\bar{a})$. Using a witness of $\mathrm{Im}_{[Q'],\varphi_0,s}(\bar{a})$ in N, we have that $g_1 + s < g_2 - s$, where g_1 and g_2 are defined as above, but working in $\mathscr{N}$. Thus, by taking $s \in \mathbf{N}$ large enough, we can find $b' \in N$ such that $\ell b' \in]g_1, g_2[$ and $D_m(\ell b' + k)$. Hence, we have $\mathscr{N} \models \varphi(b', \bar{a})$. □

Corollary 5.5.4. *Let R be a sparse set. Then $T_{R,<}$ is complete.*

Proof. This is a consequence that, by Theorem 5.5.3, $\mathscr{Z}_{R,\mathcal{L}_<}$ is a prime model of $T_{R,<}$. □

5.6 Dependency of $T_{R,<}$

We show in this section that for a sparse set R, the theory $T_{R,<}$ is dependent. (Recall that $\mathscr{R}$ is the natural $\mathcal{L}_0$-structure on R defined after Example 1.2.8.) The strategy is to apply Theorem 5.5.3 and Lemma 1.2.2 to reduce the work to atomic formulas. The first step towards the proof of the dependency of $T_{R,<}$ is to show a separation of variables in terms $t(x, \bar{y})$, see Lemmas 5.6.2 and 5.6.3. This allows us to further reduce the dependency of $T_{R,<}$ to formulas of the form $\mathrm{Im}_{[Q],\varphi_0}(x_1 + y_1, \ldots, x_n + y_n)$. Then using preliminary material from Sections 1.2 and 2.1, we show in Corollary 5.6.6, that these formulas are dependent.

We now begin the proofs of Lemmas 5.6.2 and 5.6.3 that deal with a separation of variables in $\mathcal{L}_<$-terms.

We will use Hahn's representation theorem for ordered abelian groups [24, Section 4.5]. This theorem states that an ordered abelian group embeds in the group $\mathbf{R}^X$ for some set X, generalizing Hölder's theorem that an archimedean abelian group embeds in $\mathbf{R}$. Let us briefly recall how Hahn's representation theorem works. Let G be an abelian totally ordered group and let $\bar{G}$ be its divisible closure. Given an element $g \in G \setminus \{0\}$ there is a unique convex subgroup V maximal for the property of not containing g; it is called a *value* for g. There is also a smallest convex subgroup V^+ containing g and the quotient V^+/V is an archimedean ordered group (which by Hölder's theorem, embeds in $(\mathbf{R}, +, <, 0)$). The set of all values in G forms a chain denoted by $\Gamma(G)$; we set $\Gamma(G) = \{V_\gamma \mid \gamma \in \Gamma\}$ and for $V = V_\gamma$, we denote V^+ by V^γ. Note that $\Gamma(\bar{G}) = \Gamma(G)$. Denote by $R_\gamma = V^\gamma/V_\gamma$ and let $\bar{R}_\gamma$ be the $\mathbf{Q}$-vector-subspace generated by R_γ in $\mathbf{R}$. One can decompose $\bar{G}$ as a direct sum $\bar{G} = \bar{V}^\gamma \oplus D_\gamma$, where $\bar{V}^\gamma$ is the divisible closure of V^γ in $\bar{G}$ and D_γ is some direct summand. Denote by π_γ the projection of $\bar{G}$ to $\bar{V}^\gamma$ and let $\rho_\gamma : \bar{V}^\gamma \to \bar{R}_\gamma$. Then one sends g to the function $\hat{g} : \Gamma(G) \to \mathbf{R} : \gamma \mapsto \rho_\gamma \pi_\gamma(g) = \hat{g}(\gamma)$. One verifies that $\mathrm{supp}(g) = \{\gamma \in \Gamma(G) \mid \hat{g}(\gamma) \neq 0\}$ is an anti-well ordered subset of $\Gamma(G)$. Denote by $V(\Gamma(G), \bar{R}_\gamma)$ the lexicographically ordered group of functions f from $\Gamma(G)$ to $\mathbf{R}$ with anti-well-ordered support, such that $f(\gamma) \in \bar{R}_\gamma$, for any $\gamma \in \Gamma(G)$. Then G embeds in $V(\Gamma(G), \bar{R}_\gamma)$ by the map $g \mapsto \hat{g}$ [24, Theorem 4C]. Define a map $v : G \to \Gamma(G)$ which sends $g \in G \setminus \{0\}$ to $\max(\mathrm{supp}(\hat{g}))$. This is a valuation map on G as defined in [23, Chapter 4, section 4] except that there one takes the opposite order on $\Gamma(G)$. It is constant on an archimedean class, namely if satisfies the following: for all $g, h \in G^{>0}$, $v(g) = v(h)$ if and only if $g \leq nh \leq mg$ for some $n, m \in \mathbf{N}^{>0}$ (in other words, g and h are in the same archimedean class).

Let us introduce a special function, definable in $T_{R,<}$, that has been used in [41].

This function which we call λ is defined as follows.

$$\lambda(x) = \begin{cases} 0 & \text{if } x < r_0; \\ y & \text{if } r_0 \leq x,\ y \in R \text{ and } y \leq x < S(y). \end{cases}$$

By (Ax.5), this function is well defined. Also λ is a definable function: $\lambda(x) = y$ if and only if $(y \geq r_0 \wedge \mathrm{Im}_{(-1,X)}(-y-1,y)) \vee (x < r_0 \wedge y = 0)$.

Remark 5.6.1. Let $M \models T_{<,R}$. We observe that if the sequence (r_{n+1}/r_n) is unbounded, then $r_{n+1}/r_n \to \infty$. Indeed, if (r_{n+1}/r_n) is unbounded, then for all $m \in \mathbf{N}^{>0}$, for all but finitely many elements x of R, by (Ax.8), we have $mx < S(x)$. So we get that either there exists $n \in \mathbf{N}^{>0}$ such that for all positive y but finitely many $\lambda(y) \leq y \leq n\lambda(y)$ or for all $m \in \mathbf{N}^{>0}$ for all but finitely many y, $m\lambda(y) \leq S(\lambda(y))$. In other words for elements y bigger than $\mathbf{Z}$, either y and $\lambda(y)$ are in the same archimedean class or never.

We now show how to evaluate $\lambda(x \pm y)$ for $x, y > 0$ along an indiscernible sequence. This will later on be useful to understand terms of the form $S(t(x, \bar{y}))$.

Lemma 5.6.2. *Let $\mathcal{M} \models T_{<,R}$, $d \in M$ and $(c_i \mid i \in \omega_1)$ be a non-constant indiscernible sequence in M such that $d > 0$ and $c_i > 0$ for all $i \in \omega_1$. Then there exist $i_0 \in \omega_1$, $\ell \in \mathbf{Z}$ such that one of the following holds for all $i \geq i_0$:*

- $\lambda(c_i \pm d) = S^\ell(\lambda(c_i))$;
- $\lambda(d \pm c_i) = S^\ell(\lambda(d))$;
- $\lambda(|d - c_i|) = S^\ell(\lambda(|d - c_{i_0}|))$;
- $\lambda(|d - c_i|) = S^\ell(\lambda(|c_{i+1} - c_i|))$;
- $\lambda(|d - c_{i+1}|) = S^\ell(\lambda(|c_{i+1} - c_i|))$.

Proof. The proof has two main ingredients. The first one is that in a model of a dependent theory, an indiscernible sequence indexed by ω_1 remains eventually indiscernible over a parameter (see [50, Claim in the proof of Proposition 2.11]). The second one is the following consequence Lemma 5.2.3: given $n, m \in \mathbf{N}^{>0}$, there exists $\ell, \ell' \in \mathbf{Z}$ such that whenever $x \leq ny \leq mx$, $\lambda(y) \in \{S^k(\lambda(x)) \mid k \in [\ell, \ell']\}$.

Let $(c_i \mid i \in \omega_1)$ be a non-constant indiscernible sequence. We will apply the first ingredient to certain sequences of the form $(t(c_i) \mid i \in \omega_1)$, where $t(x)$ is a $\mathcal{L}_{<,R}$-terms, and to the dependent theory $\mathrm{Th}(\mathbf{Z},+,-,0,<)$: we may assume that $(c_i \mid i \in \omega_1)$ is indiscernible over d in $\{+,-,0,<\}$. In particular we have that either $d > c_i$ for all $i \in \omega_1$ or $d < c_i$ for all $i \in \omega_1$. By comparing $v(d)$ to $v(c_0)$, we obtain that the sequence $(c_i \mid i \in \omega_1)$ falls into the following cases for some $m, n \in \mathbf{N}^{>0}$:

1. $c_i \leq n(c_i \pm d) \leq mc_i$ for all $i \in \omega_1$;

2. $d \leq n(d \pm c_i) \leq md$ for all $i \in \omega_1$;
3. $|c_{i+1} - c_i| \leq n|d - c_{i+1}| \leq m|c_{i+1} - c_i|$ for all $i \in \omega_1$;
4. $|d - c_0| \leq n|d - c_i| \leq m|d - c_0|$ for all $i \in \omega_1$;
5. $|c_{i+1} - c_i| \leq n|d - c_i| \leq m|c_{i+1} - c_i|$ for all $i \in \omega_1$.

Cases 1 and 2 occur when $v(d) \neq v(c_0)$ and case 2 together with the remaining cases when $v(d) = v(c_0)$. In case $v(d) = v(c_0)$, we further compare $v(c_1 - c_0)$, $v(d - c_1)$ and $v(d - c_0)$. Let us check all this in details.

1. $v(d) < v(c_0)$. Then $v(c_0 \pm d) = v(c_0)$, so that $c_0 \leq n(c_0 \pm d) \leq mc_0$ for some $n, m \in \mathbf{N}^{>0}$. By indiscernibility over d, we get that $c_i \leq n(c_i \pm d) \leq mc_i$ for all $i \in \omega_1$;
2. $v(d) > v(c_0)$. Then $v(d \pm c_0) = v(d)$, so that $d \leq n(d \pm c_0) \leq md$ for some $n, m \in \mathbf{N}^{>0}$. By indiscernibility over d, we get that $d \leq n(d \pm c_i) \leq md$ for all $i \in \omega_1$;
3. $v(d) = v(c_0)$. Then, as $v(d + c_0) = v(d)$, we have that $d \leq n(d + c_0) \leq md$ for some $n, m \in \mathbf{N}^{>0}$. Hence $d \leq n(d + c_i) \leq md$ for all $i \in \omega_1$. Furthermore, one of the following holds, using indiscernibility as before:
   a) $v(c_1 - c_0) = v(d - c_1)$. Then there are $n, m \in \mathbf{N}^{>0}$ such that $|c_1 - c_0| \leq n|d - c_1| \leq m|c_1 - c_0|$ and so for all $i \in \omega_1$, $|c_{i+1} - c_i| \leq n|d - c_{i+1}| \leq m|c_{i+1} - c_i|$;
   b) $v(c_1 - c_0) \neq v(d - c_1)$. Then $v(d - c_0) = \max\{v(d - c_1), v(c_1 - c_0)\}$. So,
      i. either $v(d - c_0) = v(d - c_1)$, in which case there are $n, m \in \mathbf{N}^{>0}$ such that $|d - c_0| \leq n|d - c_1| \leq m|d - c_0|$ and so for all $i \in \omega_1$, $|d - c_0| \leq n|d - c_i| \leq m|d - c_0|$;
      ii. or $v(d - c_0) = v(c_1 - c_0)$, in which case there are $n, m \in \mathbf{N}^{>0}$ such that $|c_1 - c_0| \leq n|d - c_0| \leq m|c_1 - c_0|$ and so for all $i \in \omega_1$, $|c_{i+1} - c_i| \leq n|d - c_i| \leq m|c_{i+1} - c_i|$.

Now given $g, h \in M^{>0}$ in the same archimedean class, more precisely such that $h \leq ng \leq mh$, for some $n, m \in \mathbf{N}^{>0}$, let us show that there are $\ell, \ell' \in \mathbf{Z}$ such that $\lambda(g) \in \{S^k(\lambda(h)) \mid \ell \leq k \leq \ell'\}$. We may assume that $g, h \notin \mathbf{Z}$. By Lemma 5.2.4 applied to the polynomial mX, there exists $\ell_1 \in \mathbf{N}$ such that $mS(\lambda(h)) \leq S^{\ell_1}(\lambda(h))$. This shows that $\lambda(g) \leq S^{\ell_1}(\lambda(h))$ since $n\lambda(g) \leq ng$. On the other hand, again by Lemma 5.2.4, there exists $\ell_2 \in \mathbf{N}$ such that $n\lambda(g) \leq S^{\ell_2}(\lambda(g))$. Therefore, we get that $\lambda(h) \leq S^{\ell_2}(\lambda(g))$, since $h \leq ng$. Altogether, we have $S^{-\ell_2}(\lambda(h)) \leq \lambda(g) \leq S^{\ell_1}(\lambda(h))$, as we wanted.

Applying the discussion above, with g of the form $|d \pm c_i|$ and h equal to either c_i, d, $d - c_0$, $|c_{i+1} - c_i|$, or $|c_i - c_{i-1}|$, we get:

- $\lambda(c_i \pm d) \in \{S^k(\lambda(c_i)) \mid \ell \leq k \leq \ell'\}$;

– $\lambda(d \pm c_i) \in \{S^k(\lambda(d)) \mid \ell \leq k \leq \ell'\}$;
– $\lambda(|d - c_i|) \in \{S^k(\lambda(|d - c_0|)) \mid \ell \leq k \leq \ell'\}$;
– $\lambda(|d - c_i|) \in \{S^k(\lambda(|c_{i+1} - c_i|)) \mid \ell \leq k \leq \ell'\}$;
– $\lambda(|d - c_{i+1}|) \in \{S^k(\lambda(|c_{i+1} - c_i|)) \mid \ell \leq k \leq \ell'\}$.

Thus, in all cases, $\lambda(|d - c_i|)$ and $\lambda(d + c_i)$ belong to a finite set, so for i sufficiently big, we get a constant value since these sequences are monotone. Let us prove this on the case where $\lambda(|d - c_i|) \in \{S^k(\lambda(|c_{i+1} - c_i|)) \mid \ell \leq k \leq \ell'\}$. The other cases are similar.

First let us assume that $|d - c_i| = d - c_i$ for all $i \in \omega_1$. Recall that $\lambda(d - c_i) = S^k(\lambda(|c_{i+1} - c_i|))$ means $S^k(\lambda(|c_{i+1} - c_i|)) \leq d - c_i < S^{k+1}(\lambda(|c_{i+1} - c_i|))$. So, for $k \in [\ell, \ell']$, let $(t_k(c_i)) \mid i \in \omega_1)$ be the sequence defined by $t_k(c_i) = S^k(\lambda(|c_{i+1} - c_i|)) + c_i$. For all $k \in [\ell, \ell']$, the sequence $(t_k(c_i) \mid i \in \omega_1)$ is indiscernible, because λ is definable. Hence for each $k \in [\ell, \ell']$, we may assume that $(t_k(c_i) \mid i \in \omega_1)$ is $\{<\}$-indiscernible over d, for $i > i_k$. Let $k_0 \in [\ell, \ell']$ maximal such that $t_{k_0}(c_i) \leq d$ for all $i > i_0 = \max\{i_k \mid \ell \leq k \leq \ell'\}$. Then we have that for all $i > i_0$ that $d < t_{k_0+1}(c_i)$. Thus, $\lambda(d - c_i) = S^{k_0}(\lambda(|c_{i+1} - c_i|))$ for all $i > i_0$.

Second, if $|d - c_i| = c_i - d$ for all $i \in \omega_1$, we can repeat the argument using the sequences $(t'_k(c_i) \mid i \in \omega_1)$ defined by $t'_k(c_i) = c_i - S^k(\lambda(|c_{i+1} - c_i|))$, for all $i \in \omega_1$ and $k \in [\ell, \ell']$. □

In the following lemma, for a set X, $\langle X \rangle^\lambda$ denotes the structure generated by X in the language $\mathcal{L}_< \cup \{\lambda\}$.

Lemma 5.6.3. *Let $t(x, \bar{y})$ be a $\mathcal{L}_<$-term. Let $(a_i \mid i \in \omega_1)$ be an indiscernible sequence and $b \in M$. Then there exists $i_0, i_1, \ldots, i_n \in \omega_1$ and a $(2n+1)$-ary definable function f and $b' \in \langle b, \{a_i \mid i \in \omega_1\} \rangle^\lambda$ such that for all $i \geq i_0$*

$$t(a_i, b) = f(a_i, a_{i+1}, \ldots, a_{i+n}, a_{i-1}, \ldots, a_{i-n}) + b'.$$

Proof. We show the lemma by induction on the number of occurrences of S and S^{-1} in $t(x, \bar{y})$. If $t(x, \bar{y})$ is an $\mathcal{L}_g$-term, it is well known that it can be written in the form $t_1(x) + t_2(\bar{y})$ for some $\mathcal{L}_g$-terms $t_i(x)$ and $t_2(\bar{y})$ (see [41, Lemma 4]). Therefore, what remains to be shown is that the lemma holds for $S^\epsilon(f(\bar{x}) + b')$, where $\epsilon \in \{-1, 1\}$, f is a $(2n+1)$-ary definable function and $b' \in \langle b, \{a_i \mid i \in \omega_1\} \rangle^\lambda$. Since f is definable, we have that $(f(a_i, a_{i+1}, \ldots, a_{i+n}, a_{i-1}, \ldots, a_{i-n}) \mid i \in \omega_1)$ is indiscernible. Put $c_i = f(a_i, a_{i+1}, \ldots, a_{i+n}, a_{i-1}, \ldots, a_{i-n})$. Without loss of generality, we may assume that $b' \neq 0$ and $(c_i \mid i \in \omega_1)$ non-constant. Then one of the following holds:

1. $c_i + b' < r_0$ for all $i \in \omega_1$ sufficiently large. In that case, we have $S^\epsilon(c_i + b') = c_i + b'$ for all $i \in \omega_1$ sufficiently large;

2. $c_i + b' \geq r_0$ for all $i \in \omega_1$ sufficiently large. In that case, we apply Lemma 5.6.2 to conclude. We treat the case where $\lambda(c_i + b') = S^\ell(\lambda(c_{i+1} - c_i))$ for all $i \in \omega_1$ sufficiently large, the other cases being similar. By definition of λ, we have $S^\ell(\lambda(c_{i+1} - c_i)) \leq c_i + b' < S^{\ell+1}(\lambda(c_{i+1} - c_i))$ for all $i \in \omega_1$ sufficiently large. As a result, two cases are possible, since $(S^\ell(\lambda(c_{i+1} - c_i)) - c_i \mid i \in \omega_1)$ is indiscernible:
   - either $S^\ell(\lambda(c_{i+1} - c_i)) - c_i = b'$ for all $i \in \omega_1$. In that case, we have $S^\epsilon(c_i + b') = S^{\ell+\epsilon}(\lambda(c_{i+1} - c_i))$ for all $i \in \omega_1$ sufficiently large;
   - or $S^\ell(\lambda(c_{i+1} - c_i)) - c_i < b'$ for all $i \in \omega_1$ sufficiently large, in which case $S^\epsilon(c_i + b') = c_i + b'$ for all $i \in \omega_1$ sufficiently large.

   Therefore, in both cases, $S^\epsilon(c_i + b')$ is a definable function of c_i and c_{i-1}. □

Now we move on to the analysis of the dependency of Im predicates. Our first task is to prove that they have honest definitions over R.

Lemma 5.6.4. *Let $\bar{\mathsf{f}}$ be an n-tuple of operators and let $\varphi(\bar{x}, y)$ be the formula $\mathsf{f}_1(x_1) + \cdots + \mathsf{f}_n(x_n) > y$. Then for all $\mathscr{M} \models T_{R,<}$ and $a \in M$, the formula $\varphi(\bar{x}, a)$ has an honest definition over $R(M)$, which is the relativization of an $\mathcal{L}_0$-formula.*

Proof. This is done by induction on n. Let $M \models T_{R,<}$ and $a \in M$. Assume that $n = 1$. Then by (Ax.9), either $\mathsf{f} >_R 0$ or $\mathsf{f} = 0$ or $\mathsf{f} <_R 0$. In case $\mathsf{f} >_R 0$, by Lemma 5.2.2, we know that f is ultimately strictly increasing: there exists $k \in \mathbf{N}$ such that for all $x \in R(M)$ if $x > S^k(r_0)$ then $\mathsf{f}(x) < \mathsf{f}(S(x))$.

We distinguish two cases:

1. either $a < \min_R \mathsf{f}$, in which case $r_0 = r_0$ is an honest definition of $\varphi(x, y)$;
2. or $a \geq \min_R \mathsf{f}$, in which case, by Lemma 5.2.4, there exists a maximal $b \in R(M)$ such that $\mathsf{f}(b) \leq a < \mathsf{f}(S(b))$. Then $\varphi(d, a)$ holds for $d \in R(M)$ if and only if $\mathsf{f}(S(b)) \leq \mathsf{f}(d)$. Let $k' = \min\{n \in \mathbf{N} \mid n \geq k \text{ and } \mathsf{f}(S^n(r_0)) > \mathsf{f}(S^i(r_0)) \text{ for all } i \leq k\}$. We then consider the following two cases:
   a) $b \leq S^{k'}(r_0)$. In that case, there exists $I \subset [k']$ such that
   $$x > b \vee \bigvee_{i \in I} x = S^i(r_0)$$
   is an honest definition of $\varphi(x, a)$;
   b) $b > S^{k'}(r_0)$. In that case, $x > b$ is an honest definition of $\varphi(x, a)$.

This finishes the case $\mathsf{f} >_R 0$. Now assume that $\mathsf{f} = 0$. In this case an honest definition of $\varphi(x, a)$ is $r_0 = r_0$ if $0 < a$ or $r_0 < r_0$ if $\neg(0 < a)$.

Finally, assume that $\mathsf{f} <_R 0$. There are two cases to consider:

1. either $a \geq \max_R \mathsf{f}$. In that case $r_0 < r_0$ is an honest definition of $\varphi(x, a)$;
2. or $a < \max_R \mathsf{f}$. Then there exists $b \in R(M)$ maximal such that $\mathsf{f}(S(b)) < a \leq \mathsf{f}(b)$. We may assume, at the cost of considering $S^{-1}(b)$ instead of b, that $a < \mathsf{f}(b)$. Let $k' = \min\{n \in \mathbf{N} \mid n \geq k \text{ and } \mathsf{f}(S^n(r_0)) < \mathsf{f}(S^i(r_0)) \text{ for all } i \leq k\}$ and let us consider the following cases
   a) $b \leq S^{k'}(r_0)$. In that case, There exists $I \subset [k']$ such that
   $$\bigvee_{i \in I} x = S^i(r_0)$$
   is an honest definition of $\varphi(x, a)$;
   b) $b > S^{k'}(r_0)$. In that case, $x \leq b$ is an honest definition of $\varphi(x, a)$.

This ends the case $n = 1$.

Assume that the lemma holds for all tuple $\bar{\mathsf{f}}$ of operators of length $< n$ and all a. Let $\bar{\mathsf{f}}$ be an n-tuple of operators and $a \in M$. By induction we may assume that $\mathsf{f}_i \neq 0$ for all $i \in [n]$. Let $k = \max\{k_1 \mid i \in [n]\}$, where k_i is given by Proposition 5.3.2 applied to the polynomials defining $\bar{\mathsf{f}}$ if f_i is positive and to $-\bar{\mathsf{f}}$ otherwise. By Lemmas 2.1.16 and 5.4.2, and induction, we only need to show that the formulas $\varphi(\bar{x}, a) \wedge \mu_k(x_i, \bar{x}^i)$ have an honest definition for all $i \in [n]$. Let $i \in [n]$. We first treat the case $\mathsf{f}_i > 0$. As before, we consider the following cases:

1. either $a < \min_R \mathsf{f}_i$, in which case $\mu_k(x_i, \bar{x}^i)$ is an honest definition of $\varphi(\bar{x}, a) \wedge \mu_k(x_i, \bar{x}^i)$;
2. or $a \geq \min_R \mathsf{f}_i$. In that case, there exists $b \in R(M)$ such that $\mathsf{f}_i(b) \leq a < \mathsf{f}_i(S(b))$. As a result, $\varphi(\bar{x}, a) \wedge \mu_k(x_i, \bar{x}^i)$ holds if and only if
   $$\mu_k(x_i, \bar{x}^i) \wedge (x_i > S(b) \vee (x_i = S(b) \wedge \sum_{j \in [n] \setminus \{i\}} \mathsf{f}_j(x_j) > a - \mathsf{f}_i(S(b))))$$
   holds. But by induction, there is an $\mathcal{L}_0$-formula $\theta(\bar{x}^i, \bar{z})$ and $\bar{b} \in R(M)$ such that $\theta(\bar{x}^i, \bar{b})$ is an honest definition for
   $$\sum_{j \in [n] \setminus \{i\}} \mathsf{f}_j(x_j) > a - \mathsf{f}_i(S(b)).$$
   Therefore, the formula
   $$\mu_k(x_i, \bar{x}^i) \wedge (x_i > S(b) \vee (x_i = S(b) \wedge \theta(\bar{x}^i, \bar{b}))$$
   is an honest definition of $\varphi(\bar{x}, a) \wedge \mu_k(x_i, \bar{x}^i)$.

The case where $\mathsf{f}_i <_R 0$ is done similarly. □

Corollary 5.6.5. *Let* $[Q] \in (\mathbf{Z}[X])^{n\times m}$ *and* $\varphi_0(\bar{x})$ *an* $\mathcal{L}_0$*-formula. Let* $\varphi(\bar{y},\bar{z})$ *be defined by*

$$\mathrm{Im}_{[Q],\varphi_0}(y_1+z_1,\dots,y_n+z_n).$$

Then for all $\mathscr{M} \models T_{R,<}$ *and* $\bar{a} \in M$, $\varphi(\bar{y},\bar{a})$ *has on honest definition over* $R(M)$ *which is the relativization of an* $\mathcal{L}_0$*-formula.*

Proof. By Lemma 2.1.15, it is enough to show that, given $\bar{b} \in M^n$, $[Q] \in (\mathbf{Z}[X])^{n\times m}$ and $\varphi_0(\bar{x})$ and $\mathcal{L}_0$-formula, the formula

$$\bigwedge_{i\in[n]} \sum_{j\in[m]} \mathsf{f}_{Q_{ij}}(x_j) > y_i + b_i \wedge \varphi_0^R(\bar{x})$$

has an honest definition over R. But, by Lemma 5.6.4, for all $i \in [n]$, the formula

$$\sum_{j\in[m]} \mathsf{f}_{Q_{ij}}(x_j) - y_i > b_i$$

has an honest definition (over R) $\theta_i(\bar{x},y,d_i)$. Therefore, by Lemma 2.1.16, the formula

$$\bigwedge_{i\in[n]} \theta_i(\bar{x},y,d_i) \wedge \varphi_0^R(\bar{x})$$

is an honest definition over R for

$$\bigwedge_{i\in[n]} \sum_{j\in[m]} \mathsf{f}_{Q_{ij}}(x_j) > y_i + b_i \wedge \varphi_0^R(\bar{x}). \qquad \square$$

Corollary 5.6.6. *Let* $[Q] \in (\mathbf{Z}[X])^{n\times m}$ *and* $\varphi_0(\bar{x})$ *and* $\mathcal{L}_0$*-formula. Let* $\varphi(\bar{y},\bar{z})$ *be defined by*

$$\mathrm{Im}_{[Q],\varphi_0}(y_1+z_1,\dots,y_n+z_n).$$

Then $\varphi(\bar{y},\bar{z})$ *is dependent.*

Proof. The idea is to apply Lemma 2.1.17 to the formula which expresses indiscernibility with respect to the simple formula that defines $\varphi(\bar{y},\bar{z})$, assuming towards a contradiction that $\varphi(\bar{y},\bar{z})$ is independent. So let us assume that $\varphi(\bar{y},\bar{z})$ is independent. Then by Proposition 1.2.3, there are $(\bar{a}_i \mid i \in \omega)$ indiscernible and $\bar{b} \in M$ such that $\varphi(\bar{a}_i,\bar{b})$ holds if and only if i is even.

Now let $\psi(\bar{x},\bar{y},\bar{z})$ be the formula

$$\bigwedge_{i\in[n]} \sum_{j\in[m]} \mathsf{f}_{Q_{ij}}(x_j) > y_i + z_i \wedge \varphi_0^R(\bar{x}).$$

Of course, $\varphi(\bar{a}_i,\bar{b})$ is equivalent to $\exists \bar{x} \in R\psi(\bar{x},\bar{y},\bar{z})$. Also, $\psi(\bar{x},\bar{y},\bar{z})$ is dependent, because the formulas $x > y$ and $\varphi_0^R(\bar{x})$ are dependent. For each $i \in \omega$, there exists, by assumption, $\bar{d}_{2i} \in M$ such that $\psi(\bar{d}_{2i},\bar{a}_{2i},\bar{b})$ holds.

For $k \in \mathbf{N}$, consider the formula $\delta(\bar{x}_1,\ldots,\bar{x}_k,\bar{y}_1,\ldots,\bar{y}_k)$, which expresses the fact that the tuple $(\bar{x}_1,\ldots,\bar{x}_k,\bar{y}_1,\ldots,\bar{y}_k)$ is indiscernible with respect to the formulas

$$\exists \bar{z}\Big(\bigwedge_{i\in I}\psi(\bar{x}_i,\bar{y}_i,\bar{z}) \wedge \bigwedge_{i\notin I}\neg\psi(\bar{x}_i,\bar{y}_i,\bar{z})\Big), \text{ for } I \subset [k].$$

Note that the formulas above are equivalent to a disjunction of simple formulas. Therefore $\delta(\bar{x}_1,\ldots,\bar{x}_k,\bar{y}_1,\ldots,\bar{y}_k)$ itself is a disjunction of simple formulas.

We wish to apply Lemma 2.1.17. Thus we have to show that $\delta(\bar{x}_1,\ldots,\bar{x}_k,\bar{a}_1,\ldots,\bar{a}_k)$ has an honest definition over R that satisfy the requirements of Lemma 2.1.17. By Lemma 5.6.4 and Lemma 2.1.16, we have that $\delta(\bar{x}_1,\ldots,\bar{x}_k,\bar{a}_1,\ldots,\bar{a}_k)$ has an honest definition over R, $\theta(\bar{x},\bar{c})$, which is the relativization of an $\mathcal{L}_0$-formula. Also $\exists x_1x_3\ldots x_{k-1} \in R\theta(\bar{x},\bar{z})$ is dependent over R. Therefore, we may invoke Lemma 2.1.17 and find $i_0,\ldots,i_k \in \omega$ such that $i_j \equiv_2 j$ and $(d_{i_j})_{j\equiv_2 1} \in R$ such that

$$\delta(\bar{d}_{i_1},\ldots,\bar{d}_{i_k},\bar{a}_{i_1},\ldots,\bar{a}_{i_k}).$$

Then taking k large enough, we contradict the fact that ψ is dependent. □

Theorem 5.6.7. *The theory $T_{<,R}$ is dependent.*

Proof. Since $T_{<,R}$ has quantifier elimination by Theorem 5.5.3, we may apply Lemma 1.2.2 and consider atomic formulas. Let $\varphi(x,\bar{y})$ be an atomic formula. As $x = y$ is equivalent to $x < y+1 \wedge y < x+1$, we may assume that $\varphi(x,\bar{y})$ is of the form $t(x,\bar{y}) > 0$ or

$$\mathrm{Im}_{[Q],\varphi_0}(t_1(x,\bar{y}),\ldots,t_n(x,\bar{y}))$$

where $t(x,\bar{y}),t_1(x,\bar{y}),\ldots,t_n(x,\bar{y})$ are terms, $[Q] \in \mathbf{Z}[X]^{n\times m}$ and φ_0 is a $\mathcal{L}_0$-formula.

Now let $(a_i \mid i \in \omega_1)$ be indiscernible and $\bar{b} \in M$. Let us show that the truth value of $\varphi(a_i,\bar{b})$ is eventually constant.

Assume first that $\varphi(x,\bar{y})$ is the formula $t(x,\bar{y}) > 0$. By Lemma 5.6.3, there exists $(c_i \mid i \in \omega_1)$ and $b' \in M$ such that $t(a_i,\bar{b}) = c_i + b'$ for all $i \in \omega_1$ sufficiently large. Therefore, as the formula $x > y$ is dependent, the truth value of $\varphi(a_i,\bar{b})$ is eventually constant.

Now, assume that $\varphi(x,\bar{y})$ is the formula

$$\mathrm{Im}_{[Q],\varphi_0}(t_1(x,\bar{y}),\ldots,t_n(x,\bar{y})).$$

As in the previous case, by Lemma 5.6.3, there exist $(c_{1i} \mid i \in \omega_1),\ldots,(c_{ni} \mid i \in \omega_1)$ and $b'_1,\ldots,b'_n \in M$ such that $t_j(a_i,\bar{b}) = c_{ji} + b'_j$ for all $j \in [n]$ and all $i \in \omega_1$ sufficiently large. Thus, by Corollary 5.6.6, the truth value of $\varphi(a_i,\bar{b})$ is eventually constant. □

Corollary 5.6.8. *Let R be a regular set. Then the theory $\mathrm{Th}(\mathscr{Z}_{<,R})$ is dependent.*

Proof. By Lemma 5.1.2 R is enumerated by a sparse sequence, so that R is sparse. □

5.7 Expansions of divisible ordered abelian groups by a sparse set

This section contains results similar to those obtained in Section 3.8 for expansions of divisible torsion-free abelian groups by a regular set.

Theorem 5.7.1. *Let R be a sparse set. Then*

1. *if $(\mathbf{Q}, +, 0, <, R)$ is bounded, then $(\mathbf{Q}, +, 0, <, R)$ is dependent;*
2. *if $(\mathbf{R}, +, 0, <, R)$ is bounded, then $(\mathbf{R}, +, 0, <, R)$ is dependent;*
3. *if $(\mathbf{R}, +, 0, \lfloor\cdot\rfloor, <, R)$ is bounded, then $(\mathbf{R}, +, 0, \lfloor\cdot\rfloor, <, R)$ is dependent, where $\lfloor\cdot\rfloor$ is the integer part function.*

Proof. We treat the case of $(\mathbf{Q}, +, 0, <, R)$ and $(\mathbf{R}, +, 0, \lfloor\cdot\rfloor, <, R)$, the case of the pair $(\mathbf{R}, +, 0, <, R)$ being identical to the case of $(\mathbf{Q}, +, 0, <, R)$.

It is well known that $\mathrm{Th}(\mathbf{Q}, +, 0, <)$ has quantifier elimination (see [33, Corollary 3.1.17]). As a consequence of this quantifier elimination, we get that $\mathrm{Th}(\mathbf{Q}, +, 0, <)$ is dependent. As a result, by Theorem 2.1.4 and our assumption that $(\mathbf{Q}, +, 0, <, R)$ is bounded, we only need to show that the induced structure on R is dependent. But by quantifier elimination we only need to look at the trace of formulas of the form $n_1x_1 + \cdots + n_kx_k = 0$ and $n_1x_1 + \cdots + n_kx_k > 0$. The formulas $n_1x_1 + \cdots + n_kx_k = 0$ can be replaced by the conjunction

$$n_1x_1 + \cdots + n_kx_k > -1 \wedge n_1x_1 + \cdots + n_kx_k < 1.$$

But then, using the proof of Lemma 5.6.4, we get that the induced structure is definable in $(R, S, S^{-1}, r_0, <)$, which has a dependent theory (this last statement follows from quantifier elimination). Therefore, $(\mathbf{Q}, +, 0, <, R)$ is dependent.

V. Weispfenning showed in [54] that $(\mathbf{R}, +, 0, \lfloor\cdot\rfloor, <, D_n \mid n \in \mathbf{N})$ has quantifier elimination, where D_n is interpreted as $D_n(x)$ if and only if $\lfloor x \rfloor$ is divisible by n. This quantifier elimination can be used to show that $(\mathbf{R}, +, 0, \lfloor\cdot\rfloor, <)$ is dependent, but it is also a consequence of [18, Proposition 3.1]. Therefore, as in the first case, by Theorem 2.1.4, $(\mathbf{R}, +, 0, \lfloor\cdot\rfloor, <, R)$ is dependent if and only if R_{ind} is dependent. But, quantifier elimination in $(\mathbf{R}, +, 0, \lfloor\cdot\rfloor, <, R, D_n \mid n \in \mathbf{N})$ states that R_{ind} is determined by the trace of equations, inequations and divisibility conditions, that is formulas of the form $n_1x_1 + \cdots + n_kx_k = 0$, $n_1x_1 + \cdots + n_kx_k > 0$ and $D_m(n_1x_1 + \cdots + n_kx_k)$ (here, we use the fact that, by [54, Lemma 3.2], terms evaluated at integers are equivalent to $\{+, -, 0\}$-terms). Therefore, we may use again Lemma 5.6.4 and the fact that in $D_m(n_1x_1 + \cdots + n_kx_k)$ we can separate the variables to show that R_{ind} is in this case definable in $\mathscr{R}$. Hence, $(\mathbf{R}, +, 0, \lfloor\cdot\rfloor, <, R)$ is dependent. □

The boundedness hypothesis in the previous theorem is a bit disappointing. However, in each case we could adapt the language we used for $\mathscr{L}_{<,R}$ to establish quantifier elimination for the pairs in Theorem 5.7.1. As we haven't checked this in details, we leave the boundedness of the pairs in Theorem 5.7.1 open.

The theory $\mathrm{Th}(\mathbf{Q}, +, 0, <)$ is known to be dp-minimal, by [50, Theorem A.6], as it is an o-minimal theory. As dp-minimal theories are considered as the dependent analogue of strongly minimal theories, it would be interesting to know if boundedness of a pair $\mathscr{M}_A$, where $\mathscr{M}$ is dp-minimal and $A \subset M$, is automatic, as in the strongly minimal case.

Bibliography

[1] Shigeki Akiyama, Jan-Hendrik Evertse, and Attila Pethő. "On nearly linear recurrence sequences". In: *Number theory—Diophantine problems, uniform distribution and applications*. Springer, Cham, 2017, pp. 1–24.

[2] Matthias Aschenbrenner et al. "Vapnik-Chervonenkis density in some theories without the independence property, I". In: *Transactions of the American Mathematical Society* 368 (Sept. 2011).

[3] Matthias Aschenbrenner et al. "Vapnik-Chervonenkis Density in Some Theories without the Independence Property, II". In: *Notre Dame J. Formal Logic* 54.3-4 (2013), pp. 311–363.

[4] P. T. Bateman, C. G. Jockusch, and A. R. Woods. "Decidability and undecidability of theories with a predicate for the primes". In: *Journal of Symbolic Logic* 58.2 (June 1993), pp. 672–687.

[5] Oleg Belegradek, Ya'acov Peterzil, and Frank Wagner. "Quasi-o-minimal structures". In: *J. Symbolic Logic* 65.3 (2000), pp. 1115–1132.

[6] Alexis Bès. "A survey of arithmetical definability". In: suppl. A tribute to Maurice Boffa. 2001, pp. 1–54.

[7] M. Boffa. "More on an undecidability result of Bateman, Jockusch and Woods". In: *Journal of Symbolic Logic* 63.1 (1998), p. 50.

[8] Véronique Bruyère. "Entiers et automates finis". MA thesis. Université de Mons, 1985.

[9] Véronique Bruyère et al. "Logic and p-recognizable sets of integers". In: vol. 1. 2. Journées Montoises (Mons, 1992). 1994, pp. 191–238.

[10] J. R. Büchi. *Weak second-order arithmetic and finite automata*. Russian. Kibern. Sb. 8, 42-77 (1964); translation from Z. Math. Logik Grundlagen Math. 6, 66-92 (1960). 1960.

[11] Enrique Casanovas and Martin Ziegler. "Stable theories with a new predicate". In: *The Journal of Symbolic Logic* 66.3 (03 2001), pp. 1127–1140.

[12] Artem Chernikov and Pierre Simon. "Externally definable sets and dependent pairs". In: *Israel J. Math.* 194.1 (2013), pp. 409–425.

[13] Artem Chernikov and Pierre Simon. "Externally definable sets and dependent pairs II". In: *Trans. Amer. Math. Soc.* 367.7 (2015), pp. 5217–5235.

[14] Alan Cobham. "On the base-dependence of sets of numbers recognizable by finite automata". In: *Math. Systems Theory* 3 (1969), pp. 186–192.

[15] Gabriel Conant. "Multiplicative structure in stable expansions of the group of integers". In: *Illinois J. Math.* 62.1-4 (2018), pp. 341–364.

[16] Gabriel Conant. "Stability and sparsity in sets of natural numbers". In: *Israel J. Math.* 230.1 (2019), pp. 471–508.

[17] Gabriel Conant and Michael C. Laskowski. "Weakly minimal groups with a new predicate". In: *J. Math. Log.* 20.2 (2020), pp. 2050011, 27.

[18] Alfred Dolich and John Goodrick. "Strong theories of ordered Abelian groups". In: *Fund. Math.* 236.3 (2017), pp. 269–296.

[19] Lou van den Dries and Ayhan Günaydın. "Mann pairs". In: *Trans. Amer. Math. Soc.* 362.5 (2010), pp. 2393–2414.

[20] Lou van den Dries and Ayhan Günaydın. "The fields of real and complex numbers with a small multiplicative group". In: *Proc. London Math. Soc. (3)* 93.1 (2006), pp. 43–81.

[21] Graham Everest et al. *Recurrence sequences*. Vol. 104. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2003, pp. xiv+318.

[22] A. Fiorenza and G. Vincenzi. "Limit of ratio of consecutive terms for general order-k linear homogeneous recurrences with constant coefficients". In: *Chaos Solitons Fractals* 25 (01 2011), pp. 145–152.

[23] Laszlo Fuchs. *Partially Ordered Algebraic Systems*. Dover, 2011.

[24] A. M. W. Glass. *Partially ordered groups*. Vol. 7. Series in Algebra. World Scientific Publishing Co., Inc., River Edge, NJ, 1999, pp. xiv+307.

[25] W. Hodges. *Model Theory*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1993.

[26] N. Jacobson. *Basic Algebra I: Second Edition*. Dover Books on Mathematics. Dover Publications, 2012.

[27] G. J. O. Jameson. *The prime number theorem*. Vol. 53. London Mathematical Society Student Texts. Cambridge University Press, Cambridge, 2003, pp. x+252.

[28] Itay Kaplan and Saharon Shelah. "Decidability and classification of the theory of integers with primes". In: *The Journal of Symbolic Logic* 82.3 (2017), pp. 1041–1050.

[29] Quentin Lambotte and Françoise Point. "On expansions of $(\mathbf{Z}, +, 0)$". In: *Ann. Pure Appl. Logic* 171.8 (2020), pp. 1–36.

[30] B.M. Landman and A. Robertson. *Ramsey Theory on the Integers: Second Edition*. Student Mathematical Library. American Mathematical Society, 2014.

[31] Michel Laurent. "Équations diophantiennes exponentielles". In: *Invent. Math.* 78.2 (1984), pp. 299–327.

[32] Henry B. Mann. "On linear relations between roots of unity". In: *Mathematika* 12 (1965), pp. 107–117.

[33] D. Marker. *Model Theory : An Introduction*. Graduate Texts in Mathematics. Springer New York, 2002.

[34] Robert McNaughton. "J. Richard Büchi. Weak second-order arithmetic and finite automata. Zeitschrift für mathematische Logik und Grundlagen der Mathematik, vol. 6 (1960), pp. 66–92. - J. Richard Büchi. On a decision method in restricted second order arithmetic. Logic, methodology and philosophy of science, Proceedings of the 1960 International Congress, edited by Ernest Nagel, Patrick Suppes, and Alfred Tarski, Stanford University Press, Stanford, Calif., 1962, pp. 1–11." In: *Journal of Symbolic Logic* 28.1 (1963), pp. 100–102.

[35] Christian Michaux and Roger Villemaire. "Presburger arithmetic and recognizability of sets of natural numbers by automata: new proofs of Cobham's and Semenov's theorems". In: *Ann. Pure Appl. Logic* 77.3 (1996), pp. 251–277.

[36] Rahim Moosa and Thomas Scanlon. "F-structures and integral points on semiabelian varieties over finite fields". In: *Amer. J. Math.* 126.3 (2004), pp. 473–522.

[37] Daniel Palacín and Rizos Sklinos. "On superstable expansions of free Abelian groups". In: *Notre Dame J. Form. Log.* 59.2 (2018), pp. 157–169.

[38] Hector Pasten and Xavier Vidaux. "Positive existential definability of multiplication from addition and the range of a polynomial". In: *Israel J. Math.* 216.1 (2016), pp. 273–306.

[39] Anand Pillay. *An introduction to stability theory*. English. Mineola, NY: Dover Publications, 2008, pp. xi + 146.

[40] Anand Pillay. "The model-theoretic content of Lang's conjecture". In: *Model theory and algebraic geometry*. Vol. 1696. Lecture Notes in Math. Springer, Berlin, 1998, pp. 101–106.

[41] Françoise Point. "On decidable extensions of Presburger arithmetic: from A. Bertrand numeration systems to Pisot numbers". In: *Journal of Symbolic Logic* 65 (03 2000), pp. 1347–1374.

[42] Bruno Poizat. *Cours de théorie des modèles*. Une introduction à la logique mathématique contemporaine. [An introduction to contemporary mathematical logic]. Bruno Poizat, Lyon, 1985, pp. vi+584.

[43] Bruno Poizat. *Groupes stables*. Vol. 2. Bruno Poizat, Lyon, 1987, pp. vi+218.

[44] Bruno Poizat. "Supergénérix". In: *J. Algebra* 404 (2014). À la mémoire d'Éric Jaligot. [In memoriam Éric Jaligot], pp. 240–270.

[45] M. Presburger. *Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt*. German. C. R. Congrès Math. Pays slaves 92-101, Zusatz ebenda, 395 (1930). 1930.

[46] P. Rothmaler. *Introduction to Model Theory*. Algebra, logic and applications. Taylor & Francis, 2000.

[47] Raphaël Salem. "Algebraic numbers and Fourier analysis". In: *Selected reprints*. Wadsworth Math. Ser. Wadsworth, Belmont, CA, 1983, pp. iii+68.

[48] A. L. Semenov. "Presburgerness of predicates regular in two number systems". English. In: *Sib. Math. J.* 18 (1977), pp. 289–300.

[49] A. L. Semenov. "Some extensions of the arithmetic of natural number addition". In: *Izv. Akad. Nauk SSSR Ser. Mat.* 43.5 (1979), pp. 1175–1195, 1199.

[50] Pierre Simon. *A Guide to NIP Theories*. Lecture Notes in Logic. Cambridge University Press, 2015.

[51] Richard P. Stanley. *Enumerative combinatorics. Volume 1*. Second. Vol. 49. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2012, pp. xiv+626.

[52] Terence Tao. "Every odd number greater than 1 is the sum of at most five primes". In: *Math. Comp.* 83.286 (2014), pp. 997–1038.

[53] Katrin Tent and Martin Ziegler. *A course in model theory*. Vol. 40. Lecture Notes in Logic. Association for Symbolic Logic, La Jolla, CA; Cambridge University Press, Cambridge, 2012, pp. x+248.

[54] Volker Weispfenning. "Mixed real-integer linear quantifier elimination". In: *Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation (Vancouver, BC)*. ACM, New York, 1999, pp. 129–136.

Index