A guide to F-automatic sets

by

Christa Danielle Constant Hawthorne

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Pure Mathematics

Waterloo, Ontario, Canada, 2021

© Christa Danielle Constant Hawthorne 2021

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Françoise Point Research director at the F.R.S.-FNRS
Supervisor(s):	Rahim Moosa Professor, Dept. of Pure Mathematics, University of Waterloo
Internal Member:	Jason Bell Professor, Dept. of Pure Mathematics, University of Waterloo
Internal Member:	Ross Willard Professor, Dept. of Pure Mathematics, University of Waterloo
Internal-External Member:	Jeffrey Shallit Professor, School of Computer Science, University of Waterloo

Author’s Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

A self-contained introduction to the theory of F-automatic sets is given. Building on [Bell, Moosa, *F-sets and finite automata*, Journal de théorie des nombres de Bordeaux, 2019], contributions are made to both the foundations of this theory and to questions of a model-theoretic nature.

Suppose Γ is an abelian group and $F\colon \Gamma \to \Gamma$ is an injective endomorphism. If (Γ, F) admits a *spanning set* then the notion of an *F-automatic set* can be defined. It is shown that this notion is independent of the spanning set chosen. A characterization of the existence of a spanning set is given in terms of certain functions on Γ, called *height functions*. It is shown that if Γ is finitely generated then (Γ, F) admits a spanning set if and only if no eigenvalue of the matrix of F lies in the complex unit disk.

A notion of sparsity among F-automatic sets, called *F-sparsity*, is studied. Outstanding questions from [Bell, Moosa, *F-sets and finite automata*, Journal de théorie des nombres de Bordeaux, 2019] are resolved, including independence from the spanning set chosen and closure under set summation. In addition, it is shown that sparsity can be characterized in terms of another natural class of functions introduced here, called *length functions*.

Model-theoretic tameness properties of F-automatic sets are studied. In the case where Γ is finitely generated, a combinatorial description is given of the stable F-sparse sets in terms of the F-sets introduced in [Moosa, Scanlon, *F-structures and integral points on semiabelian varieties over finite fields*, American Journal of Mathematics, 2004]. When $\Gamma = \mathbb{Z}$, this description is extended to a characterization of the stable F-automatic sets. It is shown that if $A \subseteq \Gamma$ is F-sparse then $(\Gamma, +, A)$ is NIP. Automatic methods are used to show that the following structures have NIP theories: $(\mathbb{Z}, +, d^{\mathbb{N}}, \times\restriction d^{\mathbb{N}})$ for $d \geq 2$, $(\mathbb{F}_p[t], +, t^{\mathbb{N}}, \times\restriction t^{\mathbb{N}})$ for prime $p \geq 9$, and $(\mathbb{Z}, +, <, d^{\mathbb{N}})$ for $d \geq 2$. (Here $d^{\mathbb{N}} = \{1, d, d^2, \dots\}$, and likewise with $t^{\mathbb{N}}$.)

Acknowledgements

I thank David Constant, Elfriede Jeller, Louise Philp, and Paul Philp, for their endless love and support.

I dearly thank my friends for being there for me during weather fair and foul.

Warm thanks to my advisory committee, Jason Bell and Ross Willard, for their guidance throughout my PhD. I thank my thesis committee, Jason Bell, Ross Willard, Jeffrey Shallit, and Françoise Point, for their thoroughness and their insight. I am grateful to Barbara Csima, Lis D'Alessio, Jackie Hilts, Nancy Maloney, David McKinnon, Ruxandra Moraru, Pavlina Penk, and Nico Spronk for their support and guidance throughout my university career. I am thankful to Gabriel Conant for his insightful comments on the preprint I sent him.

I cannot overstate my gratitude to my advisor, Rahim Moosa. I thank him for his excellent teaching of my first course (and later courses) in model theory. I thank him for supervising me as an undergraduate research student, as a master's student, and as a PhD student. I thank him for teaching me how to research, and how to give clear and useful mathematical exposition. I thank him for being ever-ready with an encouraging word and an answer to whatever question I came to him with. His mentorship is second to none, and I count myself privileged to have been his student.

Table of Contents

Chapter 1

Introduction

Suppose we fix a finite alphabet Σ. What sets L of strings over Σ can be computed by a computer with a finite amount of memory? That is, when is there a computer program with bounded memory usage that takes in a string and decides whether its input lies in L?

An L for which such a program exists is called a *regular language* over Σ. Regular languages satisfy many nice closure properties, and by their nature are very concrete and easy to reason about compared to general computable languages. Moreover they show up in many applications: compilers, for example, make heavy use of them. So regular languages are an interesting object of study.

One of the most natural examples of a finite alphabet is the set of decimal digits, or more generally the set of digits base d for some $d \geq 2$. In this context, we can view a string as representing a natural number. We are thus lead to the question of when a set of numbers A is such that the set of base-d expansions of elements of A is a regular language. Such A are called *d-automatic sets*. The study of d-automatic sets has nice links to logic and algebra; more details on the theory of d-automatic sets can be found in [3].

The definition of d-automaticity can be extended to subsets of $\mathbb{Z}$, and more generally to subsets of $\mathbb{Z}^m$ (see [20, 1]). More recently, Bell and Moosa show in [6] that in some circumstances we can further generalize this idea to the setting of an arbitrary countably infinite abelian group Γ where d (or rather, multiplication-by-d) is replaced by an arbitrary endomorphism $F\colon \Gamma \to \Gamma$. The idea is that if (Γ, F) admits a finite $\Sigma \subseteq \Gamma$ such that every $a \in \Gamma$ has an "expansion base F over Σ", then we can say that $A \subseteq \Gamma$ is F-automatic if the set of such base-F expansions of elements of A is a regular language.

The original motivation for this generalization came from problems in Diophantine geometry in positive characteristic. In this setting Γ comes from the rational points on some commutative algebraic group over a finite field, and $F\colon \Gamma \to \Gamma$ is induced by the Frobenius endomorphism. In [5], Bell, Ghioca, and Moosa were able to use this generalization to prove an effective isotrivial Mordell-Lang theorem.

Even if one is not interested in Diophantine geometry, this generalization is of intrinsic interest. It applies, for example, in the elementary setting where $\Gamma = \mathbb{Z}^m$ and F is given by some $m \times m$ matrix with integer entries. In this setting, the model theorist is lead naturally to ask which expansions of the group of integers by an F-automatic set result in a tame structure; say one whose first-order theory is stable, or does not satisfy the independence property (i.e., is NIP).

The goal of this thesis is to provide a self-contained account of the theory of F-automatic sets; this includes the basic theory, as well as some model-theoretic tameness results. For the sake of being self-contained, some material from [6] and other sources will be repeated; the bibliographical notes at the end of each chapter will clarify my own contributions.

We begin in Chapter 2 with an introduction to the theory of F-automatic sets. Beyond elementary properties, our main results there are concerned with the Σ chosen above, which we call a *spanning set*. We show that the requirements of Σ can be relaxed slightly: our definition of Σ is less stringent than the one appearing in [6], but we can still recover their main results. We characterize the existence of a spanning set in terms of the existence of a *height function* on (Γ, F) (Theorem 2.43). In the case where Γ is finitely generated, we show that (Γ, F) admits a spanning set if and only if all the eigenvalues of $F \otimes_{\mathbb{Z}} \mathrm{id}_{\mathbb{C}}$ have modulus > 1 (Theorem 2.48). We also show that the notion of F-automaticity is independent of the chosen spanning set (Corollary 2.27). In Chapter 3 we introduce F*-sparse* sets: these are the F-automatic sets A such that the number of elements of A that can be represented by a string of length n grows polynomially in n. We will establish the independence of this notion from the choice of Σ, and demonstrate some closure properties of F-sparsity, including set summation as asked for in [6]. We will show that sparsity can be characterized in terms of certain functions on Γ, which we call *length functions* (Theorem 3.19). Some of these results on F-sparsity have found use in [5]. In Chapter 4, we begin our consideration of model-theoretic tameness properties. For Γ finitely generated, we will characterize stability among the F-sparse sets: we show that an F-sparse set A is stable if and only if it is a Boolean combination of translates of finite sums of sets of the form

$$\{\, a + F^r a + \cdots + F^{r(n-1)} a : n \in \mathbb{N} \,\}$$

(Theorem 4.11). For the special case $\Gamma = (\mathbb{Z}, +)$, in Chapter 5 we extend our characterization to all automatic subsets of $\mathbb{Z}$: we show that the stable F-automatic subsets of $\mathbb{Z}$ are precisely the Boolean combinations of stable F-sparse sets and cosets of subgroups of Γ (Theorem 5.17).

In Chapter 6 we turn our attention from stability to NIP. We introduce a class of subsets of Γ called the F-*EDP* sets, which contains the F-sparse sets. We show that if $A \subseteq \Gamma$ is F-EDP and if $\mathrm{Th}(\Gamma, +)$ is weakly minimal then $\mathrm{Th}(\Gamma, +, A)$ is NIP (Theorem 6.13). We then give some applications of this result: we show that the structures $(\mathbb{Z}, +, d^{\mathbb{N}}, \times\restriction d^{\mathbb{N}})$, $(\mathbb{F}_p[t], +, t^{\mathbb{N}}, \times\restriction t^{\mathbb{N}})$ for prime $p \geq 9$, and $(\mathbb{Z}, +, <, d^{\mathbb{N}})$ have NIP theories. This last was shown in [17], but with a very different proof. Finally, we present some possible avenues of future research in Chapter 7.

The original contributions of this thesis appeared in the publications [14, 13]. Much of the material here is drawn from these papers.

In this thesis, $\mathbb{N}$ denotes the set of non-negative integers.

Chapter 2

F-automatic sets

In this chapter, we develop the basic theory of F-automatic sets. We begin by reviewing the relevant theory of regular languages, as well as the classical theory of d-automatic sets of natural numbers. We give a formal description of spanning sets and F-automatic sets. We establish some elementary properties of F-automaticity, and we show that F-automaticity is independent of the choice of spanning set (Corollary 2.27). We then characterize the existence of spanning sets in Theorems 2.43 and 2.48. Finally, we give a class of examples of F-automatic sets, the *F-sets*, that will come up in Chapters 4 and 5.

Unless otherwise stated, Γ refers to some fixed infinite abelian group, and F refers to some fixed injective endomorphism of Γ.

2.1 Preliminaries from formal languages

We begin by reviewing some basic definitions and results from the study of formal languages. A more detailed treatment can be found in [24].

Fix a finite set Σ, which we will use as an *alphabet*. We let Σ^* denote the set of strings of elements of Σ. We let ε denote the empty string. Given a string σ we let $|\sigma|$ denote its length. A *language* over Σ is any subset of Σ^*.

Definition 2.1. The class of *regular languages* over Σ is the smallest class of languages over Σ that contains all finite languages and is such that if $L_1, L_2 \subseteq \Sigma^*$ are regular then so are:

1. $L_1 \cup L_2$,

2. $L_1L_2 = \{ \sigma\tau : \sigma \in L_1, \tau \in L_2 \}$, and

3. $L_1^* = \{ \sigma_1 \cdots \sigma_n : n \in \mathbb{N}, \sigma_1, \ldots, \sigma_n \in L_1 \}$.

Note that $\varepsilon \in L_1^*$ regardless of L_1.

Regular languages are computationally easy to work with: if L is regular then membership in L can be decided by a restricted computer called a *finite automaton*.

Definition 2.2. A *deterministic finite automaton* (DFA) is a 5-tuple $(\Sigma, Q, q_0, \Omega, \delta)$, where:

- Σ is a finite alphabet;
- Q is a finite set, the set of *states*;
- $q_0 \in Q$ is the *initial state*;
- $\Omega \subseteq Q$ is the set of *accepting states*; and
- $\delta \colon Q \times \Sigma \to Q$ is the *transition function*.

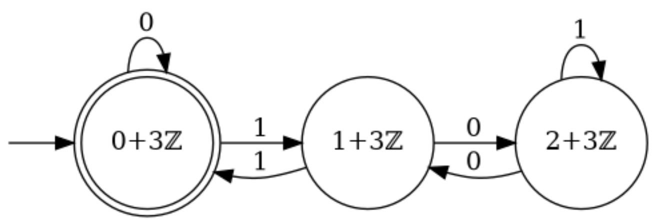
Informally, given a finite automaton $M = (\Sigma, Q, q_0, \Omega, \delta)$ and an input $\sigma \in \Sigma^*$, the machine starts in state q_0. It then reads the first character s_0 of σ, and transitions to state $\delta(q_0, s_0)$. It then reads the second character, and so on. Eventually σ will be exhausted, and the machine will be left in some state q. We say M *accepts* σ if $q \in \Omega$, and otherwise *rejects* σ.

More formally, we extend δ to a map $Q \times \Sigma^* \to Q$ by setting $\delta(q, \varepsilon) = q$ and $\delta(q, \sigma\ell) = \delta(\delta(q, \sigma), \ell)$ for $\ell \in \Sigma$ and $\sigma \in \Sigma^*$. The *language accepted* or *recognized* by M is $\{ \sigma \in \Sigma^* : \delta(q_0, \sigma) \in \Omega \}$.

Example 2.3. Let $\Sigma = \{ 0, 1 \}$, and let $L \subseteq \Sigma^*$ be the set of binary representations of multiples of 3, with leading zeroes allowed. We construct a DFA recognizing L. The idea is as follows: suppose we are given $\sigma \in \Sigma^*$ that is a binary representation of some $n \in \mathbb{N}$. Let $n_0 = 2n$ and $n_1 = 2n + 1$ be the numbers represented in binary by $\sigma 0$ and $\sigma 1$, respectively. Then in order to determine $n_0 + 3\mathbb{Z}$ and $n_1 + 3\mathbb{Z}$, it suffices to know $n + 3\mathbb{Z}$. So we can take the states of our machine to be the congruence classes modulo 3, and the above tells us that we can define a sensible transition map.

More formally, let $Q = \mathbb{Z}/3\mathbb{Z}$, $q_0 = 0 + 3\mathbb{Z}$, and $\Omega = \{ 0 + 3\mathbb{Z} \}$. We then set $\delta(i + 3\mathbb{Z}, j) = 2i + j + 3\mathbb{Z}$. One can verify by induction that if $\sigma \in \Sigma^*$ is a binary representation of n then $\delta(q_0, \sigma) = n + 3\mathbb{Z}$; so this machine accepts the σ that represent multiples of 3, as desired.

Our DFA can be represented in diagram as follows:



where each bubble represents a state; the sourceless arrow indicates the start state; a bubble enclosed by two circles represents an accepting state; and an arrow from state q_1 to state q_2 labelled with $\ell \in \Sigma$ indicates that $\delta(q_1, \ell) = q_2$. We will not make further use of these diagrams, but they are common in the study of automata, and are a useful way to think about DFAs.

Fact 2.4 ([24, Sections 3.2 and 3.3])**.** *A language $L \subseteq \Sigma^*$ is regular if and only if it is recognized by a DFA.*

Corollary 2.5. *Regular languages are closed under Boolean combinations.*

Proof. Since they are closed under union by definition, it suffices to check complementation. Suppose $L \subseteq \Sigma^*$ is regular; say it is recognized by $(\Sigma, Q, q_0, \Omega, \delta)$. Then $\Sigma^* \setminus L$ is recognized by $(\Sigma, Q, q_0, Q \setminus \Omega, \delta)$, and is thus regular. $\square$ Corollary 2.5

A useful fact about DFAs is that we can add non-determinism without changing the class of languages recognized.

Definition 2.6. A *non-deterministic finite automaton* (NFA) is a 5-tuple $(\Sigma, Q, q_0, \Omega, \delta)$ where Σ, Q, q_0, Ω are as in the definition of DFAs and $\delta \colon Q \times \Sigma \to \mathcal{P}(Q)$. The idea is that if the machine is in state q and receives input $\ell \in \Sigma$ then it can transition to any state in $\delta(q, \ell)$; the machine accepts a string if there is some choice of transitions for which the machine ends in an accepting state.

More formally, we extend δ to a map $Q \times \Sigma^* \to \mathcal{P}(Q)$ by setting $\delta(q, \varepsilon) = \{ q \}$ and

$$\delta(q, \sigma\ell) = \bigcup_{q' \in \delta(q, \sigma)} \delta(q', \ell)$$

for $\ell \in \Sigma$ and $\sigma \in \Sigma^*$. The language recognized by the machine is then $\{ \sigma \in \Sigma^* : \delta(q_0, \sigma) \cap \Omega \neq \emptyset \}$.

Fact 2.7 ([24, Lemma 2.2])**.** *$L \subseteq \Sigma^*$ is recognized by an NFA if and only if it is recognized by a DFA.*

We can use NFAs to deduce that regular languages are closed under projections. First, we will need a short discussion of Cartesian powers. We will use the notation Σ^m to denote the alphabet $\underbrace{\Sigma \times \cdots \times \Sigma}_{m \text{ times}}$. Note that this conflicts with the usual notation in the study of formal languages, in which Σ^m denotes $\{\, \sigma \in \Sigma^* : |\sigma| = m \,\}$; we will use $\Sigma^{(m)}$ for this.

When working over Σ^m, it will be convenient to view $(\Sigma^m)^*$ as a subset of $(\Sigma^*)^m$; we do this as follows. Given a string $\boldsymbol{\sigma} \in (\Sigma^m)^*$, write $\boldsymbol{\sigma} = \begin{pmatrix} s_{11} \\ \vdots \\ s_{m1} \end{pmatrix} \cdots \begin{pmatrix} s_{1n} \\ \vdots \\ s_{mn} \end{pmatrix}$ for $s_{ij} \in \Sigma$. We then identify $\boldsymbol{\sigma}$ with $\begin{pmatrix} s_{11} \cdots s_{1n} \\ \vdots \\ s_{m1} \cdots s_{mn} \end{pmatrix} \in (\Sigma^*)^m$. We thus identify $(\Sigma^m)^*$ with the set of tuples in $(\Sigma^*)^m$ whose constituent strings all have the same length.

Corollary 2.8. *Suppose $L \subseteq (\Sigma^{m+1})^*$ is regular. Then the projection $L_0 \subseteq (\Sigma^m)^*$ of L away from the last coordinate is regular.*

Proof. This follows from the general fact that the image of a regular language under a monoid homomorphism is again regular; for expository purposes, we instead give a direct proof using NFAs. The idea will be to construct an NFA that given $\boldsymbol{\sigma} \in (\Sigma^m)^*$ uses non-determinism to "guess" $\tau \in \Sigma^*$ such that $\begin{pmatrix} \boldsymbol{\sigma} \\ \tau \end{pmatrix} \in L$.

Fix a DFA $M = (\Sigma^{m+1}, Q, q_0, \Omega, \delta)$ recognizing L. Define an NFA $M' = (\Sigma^m, Q, q_0, \Omega, \delta')$ where

$$\delta'(q, \boldsymbol{\ell}) = \left\{ \delta\left(q, \begin{pmatrix} \boldsymbol{\ell} \\ \ell' \end{pmatrix}\right) : \ell' \in \Sigma \right\}.$$

for $\boldsymbol{\ell} \in \Sigma^m$. One can check by induction that $\delta'(q, \boldsymbol{\sigma}) = \left\{ \delta\left(q, \begin{pmatrix} \boldsymbol{\sigma} \\ \tau \end{pmatrix}\right) : \tau \in \Sigma^*, |\tau| = |\boldsymbol{\sigma}| \right\}$ for $\boldsymbol{\sigma} \in (\Sigma^m)^*$. So if $\boldsymbol{\sigma} \in (\Sigma^m)^*$ then M' accepts $\boldsymbol{\sigma}$ if and only if there is $\tau \in \Sigma^*$ such that $|\tau| = |\boldsymbol{\sigma}|$ and $\delta\left(q_0, \begin{pmatrix} \boldsymbol{\sigma} \\ \tau \end{pmatrix}\right) \in \Omega$; i.e., if and only if $\boldsymbol{\sigma} \in L_0$. So M' recognizes L_0, and L_0 is regular. $\square$ Corollary 2.8

A nice fact about regular languages: one can expand the alphabet without changing whether a language is regular.

Proposition 2.9. *Suppose $\Sigma_1 \subseteq \Sigma_2$ are finite alphabets and $L \subseteq \Sigma_1^*$. Then L is a regular language over Σ_1 if and only if it is regular over Σ_2.*

Proof. The left-to-right direction can be done by structural induction on the regular languages over Σ_1. Indeed, finite subsets of Σ_1^* are regular over Σ_2, and if $L_1, L_2 \subseteq \Sigma_1^*$ are regular over Σ_2 then so are $L_1 \cup L_2$, $L_1 L_2$, and L_1^*. Conversely, suppose L is regular over Σ_2; so there is a DFA $M = (\Sigma_2, Q, q_0, \Omega, \delta)$ recognizing L. We can obtain another DFA by simply removing from M any transition labelled by $\Sigma_2 \setminus \Sigma_1$: more formally, we consider $M' = (\Sigma_1, Q, q_0, \Omega, \delta{\restriction}(Q \times \Sigma_1))$. Then M' is a DFA with input alphabet Σ_1 recognizing L, and thus L is regular over Σ_1. □ Proposition 2.9

2.2 d-Automatic sets and generalizations

We would like to use the framework of regular languages to analyze sets of numbers. To do so, we need a way to represent numbers as strings over some finite alphabet. In the case of natural numbers, there is a natural way to do so: base-d representations. Fix an integer $d \geq 2$. We use the alphabet $\Sigma_d := \{0, \ldots, d-1\}$. Given a string $s_0 \cdots s_{n-1} \in \Sigma_d^*$, we define $[s_0 \cdots s_{n-1}]_d = s_0 + ds_1 + \cdots + d^{n-1}s_{n-1}$; that is, $[s_0 \cdots s_{n-1}]_d$ is the number represented by $s_{n-1} \cdots s_0$ in base d.[1] Then $[\cdot]_d \colon \Sigma_d^* \to \mathbb{N}$ is a surjective map, and thus allows us to talk about natural numbers using strings. We will be interested in when a set of natural numbers corresponds to a regular set of strings.

Definition 2.10. We say $A \subseteq \mathbb{N}$ is *d-automatic* if $\{\sigma \in \Sigma_d^* : [\sigma]_d \in A\}$ is a regular language.

A first example: $d^{\mathbb{N}}$ is d-automatic, since $\{\sigma \in \Sigma_d^* : [\sigma]_d \in 2^{\mathbb{N}}\} = 0^*10^*$. (Here we use "$0^*10^*$" as shorthand for $\{0\}^*\{1\}\{0\}^* = \{0^i10^j : i, j \in \mathbb{N}\}$.) Interestingly, a set is rarely both d- and d'-automatic: Cobham's theorem tells us that if d and d' are multiplicatively independent and $A \subseteq \mathbb{N}$ is both d- and d'-automatic, then the characteristic function of A is ultimately periodic. More information on d-automaticity can be found in [3].

So we can analyze subsets of $\mathbb{N}$ using formal languages. What about subsets of $\mathbb{Z}$? Here, too, we have a natural way to represent numbers as strings over a finite alphabet: we simply add a sign to our representations of natural numbers. Our alphabet is now $\Sigma_d \cup \{+, -\}$. Given $\sigma \in \Sigma_d^*$ we set $[+\sigma]_d = [\sigma]_d$, and $[-\sigma]_d = -[\sigma]_d$; all other strings are not considered

[1] Our map $[\cdot]_d$ views the last digit as the most significant; note that this contrasts the usual convention for base-d representations, in which the most significant digit appears first.

valid representations. Again this is surjective, and we say that $A \subseteq \mathbb{Z}$ is d-automatic if the set of valid representations σ such that $[\sigma]_d \in A$ is regular.

What about subsets of $\mathbb{Z}^m$? Let us restrict our attention to non-negative integers for now. We would like to represent a tuple $\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$ of naturals by a tuple $\begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_m \end{pmatrix}$ of base-d representations; but our representation needs to be a single string, not a tuple. As we mentioned at the end of the previous section, if all σ_i have the same length then we can view $\begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_m \end{pmatrix}$ as a single string over Σ_d^m; moreover by adding trailing zeroes to the σ_i we can assume they do all have the same length. This gives us our representations: our alphabet is Σ_d^m, and given $\begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_m \end{pmatrix} \in (\Sigma_d^m)^*$ we let $\left[\begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_m \end{pmatrix}\right]_d = \begin{pmatrix} [\sigma_1]_d \\ \vdots \\ [\sigma_m]_d \end{pmatrix}$. Our approach to negative numbers is analogous to the one-dimensional case; so for example a base-d representation of $\begin{pmatrix} -23 \\ 456 \end{pmatrix}$ is $\begin{pmatrix} - \\ + \end{pmatrix}\begin{pmatrix} 3 \\ 6 \end{pmatrix}\begin{pmatrix} 2 \\ 5 \end{pmatrix}\begin{pmatrix} 0 \\ 4 \end{pmatrix}$. As before, this $[\cdot]_d$ is surjective, and we say that $A \subseteq \mathbb{Z}^m$ is d-automatic if the set of σ such that $[\sigma]_d \in A$ is regular.

What if we replace d with some injective endomorphism F of $(\mathbb{Z}^m, +)$? At this point we may as well replace $(\mathbb{Z}^m, +)$ with a general abelian group Γ as well. If we wish to mimic the constructions we used above, we will need a finite set Σ and a surjective (possibly partial) function $[\cdot]_F \colon \Sigma^* \to \Gamma$. It is not clear in this context what Σ should be, but we can define $[\cdot]_F$ by analogy with the above: given group elements $a_0, \ldots, a_{n-1} \in \Gamma$, we define $[a_0 \cdots a_{n-1}]_F = a_0 + Fa_1 + \cdots + F^{n-1}a_{n-1}$. We are then faced with the question of whether there is a finite $\Sigma \subseteq \Gamma$ such that $[\cdot]_F \colon \Sigma^* \to \Gamma$ is surjective. This question turns out to be complicated; we will explore it in Section 2.5. For now we merely assume Γ and F are such that there is such a Σ, which we call an *F-spanning set*. (The full definition of a spanning set, Definition 2.12, requires that Σ satisfy some other conditions.) Under this assumption, we have found a way to represent group elements as strings over some finite alphabet, and can thus give a definition of automaticity in this context: we say that $A \subseteq \Gamma$ is *F-automatic* if $\{\, \sigma \in \Sigma^* : [\sigma]_F \in A \,\}$ is regular. Remarkably, this definition turns out to be independent of the Σ chosen; this is Corollary 2.27.

Some examples of (Γ, F) to which we might apply our definition:

- We can take $\Gamma = (\mathbb{Z}^m, +)$ equipped with multiplication by some $d \geq 2$, or more generally some $m \times m$ matrix with integer coefficients that is invertible over $\mathbb{Q}$.

- If R is a ring we can take $\Gamma = (R[t], +)$ with $F\colon \Gamma \to \Gamma$ given by $f \mapsto tf$.
- Let G be a commutative algebraic group over a finite field $\mathbb{F}_q$, and let K be a finitely generated field extension of $\mathbb{F}_q$. We can take Γ to be $G(K)$, the K-rational points of G, and let $F\colon \Gamma \to \Gamma$ be the q-power Frobenius mapping. These objects are considered in [18, 6, 5].

We defer for now the question of whether these admit spanning sets.

2.3 Spanning sets

We now give the full definition of spanning sets, which consists of some technical conditions in addition to the property described in the previous section.

Fix an infinite abelian group Γ and an injective endomorphism $F\colon \Gamma \to \Gamma$. Given a string $a_0 \cdots a_{n-1}$ of elements of Γ we let $[a_0 \cdots a_{n-1}]_F = a_0 + Fa_1 + \cdots + F^{n-1}a_{n-1}$.

Remark 2.11. A useful property of $[\cdot]_F$ is that it interacts nicely with string concatenation: given strings $\sigma, \tau \in \Gamma^*$ of elements of Γ, we have that $[\sigma\tau]_F = [\sigma]_F + F^{|\sigma|}[\tau]_F$.

Definition 2.12. An *F-spanning set*[2] (for Γ) is a finite $\Sigma \subseteq \Gamma$ satisfying the following:

(i) For every $a \in \Gamma$ there is $\sigma \in \Sigma^*$ such that $a = [\sigma]_F$.

(ii) $0 \in \Sigma$ and if $a \in \Sigma$ then $-a \in \Sigma$.

(iii) If $a_1, a_2, a_3 \in \Sigma$ then $a_1 + a_2 + a_3 \in \Sigma + F\Sigma$.

(iv) If $a_1, a_2 \in \Sigma$ are such that $a_1 + a_2 = Fb$ for some $b \in \Gamma$, then $b \in \Sigma$.

If Σ is F-spanning we define a map $\lambda_\Sigma\colon \Gamma \to \mathbb{N}$ that sends $a \in \Gamma$ to the length of the shortest $\sigma \in \Sigma^*$ such that $[\sigma]_F = a$.

We saw axiom (i) in Section 2.2, and axiom (ii) is straightforward. Axioms (iii) and (iv) are more technical. Roughly speaking, their purpose is to constrain $\lambda_\Sigma(a+b)$ and $\lambda_\Sigma(Fa)$ in terms of $\lambda(a)$ and $\lambda(b)$; see Proposition 2.14 and Proposition 2.19.

[2]This definition is more general than the one appearing in [6]; see the bibliographical notes for details.

Example 2.13. Some examples of (Γ, F) for which a spanning set exists:

1. Consider $\Gamma = (\mathbb{Z}, +)$ and $Fa = da$ for some $d \geq 2$. Then $\Sigma = \{\, -d+1, -d+2, \ldots, d-1 \,\}$ is an F-spanning set. Indeed, axiom (i) is just the usual base-d representation of integers; axiom (ii) is clear; axiom (iii) is because if $|a_1|, |a_2|, |a_3| \leq d-1$ then $|a_1 + a_2 + a_3| \leq 3(d-1) \leq (d+1)(d-1) = d^2 - 1$ (since $d \geq 2$), so $a_1 + a_2 + a_3$ has a 2-digit representation base d; and axiom (iv) is because if $|a_1|, |a_2| \leq d-1$ and $a_1 + a_2 = da$ then $|a| = d^{-1}|a_1 + a_2| \leq d^{-1}2(d-1) \leq d-1$ (since $d \geq 2$), so $a \in \Sigma$.

   So our setting includes the first generalization we discussed in Section 2.2. (Though we don't yet know that our notion of automaticity coincides with the existing one; we will see this in Corollary 2.30.)

2. Suppose we are given (Γ, F) and (Γ', F'); suppose there is an F-spanning set Σ and an F'-spanning set Σ'. Then $\Sigma \times \Sigma'$ is an $(F \times F')$-spanning set for $\Gamma \times \Gamma'$. Indeed:

   (i) Given $\begin{pmatrix} a \\ a' \end{pmatrix} \in \Gamma \times \Gamma'$ we can write $a = [s_0 \cdots s_{n-1}]_F$ and $a' = [s'_0 \cdots s'_{n'-1}]_{F'}$ for some $s_i \in \Sigma$ and $s'_i \in \Sigma'$. By appending zeroes, we may assume $n = n'$. Then $\begin{pmatrix} a \\ a' \end{pmatrix} = \left[\begin{pmatrix} s_0 \\ s'_0 \end{pmatrix} \cdots \begin{pmatrix} s_{n-1} \\ s'_{n-1} \end{pmatrix}\right]_{F \times F'}$, and each $\begin{pmatrix} s_i \\ s'_i \end{pmatrix} \in \Sigma \times \Sigma'$.

   (ii) Since $0 \in \Sigma$ and $0 \in \Sigma'$ we get that $\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \Sigma \times \Sigma'$. If $\begin{pmatrix} a \\ a' \end{pmatrix} \in \Sigma \times \Sigma'$ then since Σ, Σ' are spanning sets we get that $-a \in \Sigma$ and $-a' \in \Sigma'$; so $-\begin{pmatrix} a \\ a' \end{pmatrix} \in \Sigma \times \Sigma'$.

   (iii) Suppose $\begin{pmatrix} a_1 \\ a'_1 \end{pmatrix}, \begin{pmatrix} a_2 \\ a'_2 \end{pmatrix}, \begin{pmatrix} a_3 \\ a'_3 \end{pmatrix} \in \Sigma \times \Sigma'$. Then since Σ is F-spanning there are $b, c \in \Sigma$ such that $a_1 + a_2 + a_3 = b + Fc$; likewise we can find $b', c' \in \Sigma'$. Then $\begin{pmatrix} a_1 \\ a'_1 \end{pmatrix} + \begin{pmatrix} a_2 \\ a'_2 \end{pmatrix} + \begin{pmatrix} a_3 \\ a'_3 \end{pmatrix} = \begin{pmatrix} b + Fc \\ b' + F'c' \end{pmatrix} = \begin{pmatrix} b \\ b' \end{pmatrix} + (F \times F') \begin{pmatrix} c \\ c' \end{pmatrix}$.

   (iv) Suppose $\begin{pmatrix} a_1 \\ a'_1 \end{pmatrix}, \begin{pmatrix} a_2 \\ a'_2 \end{pmatrix} \in \Sigma \times \Sigma'$ satisfies $\begin{pmatrix} a_1 \\ a'_1 \end{pmatrix} + \begin{pmatrix} a_2 \\ a'_2 \end{pmatrix} = (F \times F') \begin{pmatrix} b \\ b' \end{pmatrix}$ for some $\begin{pmatrix} b \\ b' \end{pmatrix} \in \Gamma \times \Gamma'$. Since Σ is F-spanning and $a_1 + a_2 = Fb$ we get that $b \in \Sigma$; we likewise get that $b' \in \Sigma'$. So $\begin{pmatrix} b \\ b' \end{pmatrix} \in \Sigma \times \Sigma'$.

   A particular example: combining this with the previous example, we see that when $\Gamma = (\mathbb{Z}^m, +)$ and F is multiplication by d there is an F-spanning set. So our setting also includes the second generalization discussed in Section 2.2. (Though again we won't see until Corollary 2.30 that F-automaticity in this context agrees with d-automaticity.)

3. Consider $\Gamma = (\mathbb{Z}^2, +)$ with F given in matrix form by $\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$. Let $\Sigma = \{\, v \in \mathbb{Z}^2 : \|v\|_\infty \leq 2 \,\}$. We verify that Σ is F-spanning.

   (i) We show by induction on $\|v\|_\infty$ that if $v \in \mathbb{Z}^2$ then there is $\sigma \in \Sigma^*$ such that $v = [\sigma]_F$. For the base cases $\|v\|_\infty \leq 2$ we simply get that $v \in \Sigma$. For the induction step, suppose $\|v\|_\infty \geq 3$. We compute

$$F\Gamma = \left\{ \begin{pmatrix} 3a+b \\ 3b \end{pmatrix} : a, b \in \mathbb{Z} \right\} = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} : a, b \in \mathbb{Z},\ 3 \mid b,\ a \equiv \frac{b}{3} \pmod{3} \right\}.$$

   In particular, $\{\, 0, 1, 2 \,\}^2 \subseteq \Sigma$ contains a representative of each coset of $F\Gamma$. Indeed, suppose $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{Z}^2$. Pick $b' \in \{\, 0, 1, 2 \,\}$ such that $b' \equiv b \pmod{3}$, and pick $a' \in \{\, 0, 1, 2 \,\}$ such that $a' \equiv a + \frac{b-b'}{3} \pmod{3}$. So $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a' \\ b' \end{pmatrix} + \begin{pmatrix} a-a' \\ b-b' \end{pmatrix}$ and $\begin{pmatrix} a-a' \\ b-b' \end{pmatrix} \in F\Gamma$, as desired.

   So there is $s \in \Sigma$ such that $s \equiv v \pmod{F\Gamma}$. Then by the triangle inequality $\|v - s\|_\infty \leq \|v\|_\infty + \|s\|_\infty \leq \|v\|_\infty + 2$. Moreover given any $w = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{Z}^2$ we have that

$$\begin{aligned} \|Fw\|_\infty &= \left\| \begin{pmatrix} 3a+b \\ 3b \end{pmatrix} \right\|_\infty \\ &= \max\{\, |3a+b|, |3b| \,\} \\ &\geq \max\{\, |3a| - |b|, |3b| - |b| \,\} \\ &= \max\{\, |3a|, |3b| \,\} - |b| \\ &\geq 3\|w\|_\infty - \frac{1}{3}\|Fw\|_\infty \end{aligned}$$

   (since $\|Fw\|_\infty \geq |3b|$). So rearranging we find that

$$\|w\|_\infty \leq \frac{4}{9}\|Fw\|_\infty \tag{2.1}$$

for any $w \in \mathbb{Z}^2$. In particular $\|F^{-1}(v-s)\|_\infty \leq \frac{4}{9}\|v-s\|_\infty \leq \frac{4}{9}\|v\|_\infty + \frac{4}{9}\|s\|_\infty \leq \frac{4}{9}\|v\|_\infty + 1 \leq \frac{7}{9}\|v\|_\infty$ (since $\|s\|_\infty \leq 2$ and $\|v\|_\infty \geq 3$ by hypothesis). So by the induction hypothesis there is $\sigma \in \Sigma^*$ such that $F^{-1}(v-s) = [\sigma]_F$. Then $[s\sigma]_F = s + F[\sigma]_F = v$, as desired.

(ii) It is clear from the definition of Σ that $0 \in \Sigma$ and if $s \in \Sigma$ then so too is $-s$.

(iii) Note that if $s_1, s_2, s_3 \in \Sigma$ then $\|s_1 + s_2 + s_3\|_\infty \leq 6$. Suppose we are given $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{Z}^2$ with $|a|, |b| \leq 6$; we wish to show that $\begin{pmatrix} a \\ b \end{pmatrix} \in \Sigma + F\Sigma$. By negating if necessary we may assume $a \geq 0$. In axiom (i) we argued that $\{0,1,2\}^2 \subseteq \Sigma$ contains a representative for each coset of $F\Gamma$; we can similarly prove that the same holds of $\{0,1,2\} \times \{0,-1,-2\} \subseteq \Sigma$. So we can find $\begin{pmatrix} c \\ d \end{pmatrix} \in \Sigma$ such that

- $\begin{pmatrix} a \\ b \end{pmatrix} \equiv \begin{pmatrix} c \\ d \end{pmatrix} \pmod{F\Gamma}$;
- $a, c \geq 0$; and
- b and d have the same sign.

In particular $|a - c| \leq \max\{\,|a|, |c|\,\} \leq 6$ and $|b - d| \leq \max\{\,|b|, |d|\,\} \leq 6$. So $\left\|\begin{pmatrix} a \\ b \end{pmatrix} - \begin{pmatrix} c \\ d \end{pmatrix}\right\|_\infty \leq 6$, and by Eq. (2.1) we get that $\left\|F^{-1}\left(\begin{pmatrix} a \\ b \end{pmatrix} - \begin{pmatrix} c \\ d \end{pmatrix}\right)\right\|_\infty \leq \frac{4}{9}\cdot 6 < 3$. Thus, recalling the definition of Σ, we find that $F^{-1}\left(\begin{pmatrix} a \\ b \end{pmatrix} - \begin{pmatrix} c \\ d \end{pmatrix}\right) \in \Sigma$, and $\begin{pmatrix} a \\ b \end{pmatrix} \in \Sigma + F\Sigma$.

(iv) Note that if $s_1, s_2 \in \Sigma$ then $\|s_1 + s_2\|_\infty \leq 4$. Hence if $s_1 + s_2 \in F\Gamma$ then by Eq. (2.1) $\|F^{-1}(s_1 + s_2)\|_\infty \leq \frac{4}{9} \cdot 4 \leq 2$. Thus, recalling the definition of Σ, we find that $F^{-1}(s_1 + s_2) \in \Sigma$.

Note that since F isn't diagonalizable this example doesn't arise from combining the previous two examples.

4. Let G be a finite abelian group, and consider $\Gamma = \bigoplus_{i \in \mathbb{N}} G$ equipped with the shift operation $F(a_0, a_1, \ldots) = (0, a_0, a_1, \ldots)$. Then $\Sigma := G$ is an F-spanning set.

Two of the examples above took the form of $\mathbb{Z}^m$ equipped with a matrix with integer coefficients. We will see later on in Theorem 2.48 that a spanning set exists in such contexts if and only if none of the complex eigenvalues of the matrix lie in the unit disk.

It's a bit distasteful that axiom (iii) needs to be about ternary sums, rather than just binary. The main reason for this is the following result, which extends axioms (iii) and (iv) to strings over Σ, at the cost of making axiom (iii) about binary sums.

Proposition 2.14. *Suppose Σ is an F-spanning set and $n \in \mathbb{N}$.*

1. *If $a, b \in \Gamma$ then $\lambda_\Sigma(a+b) \leq \max(\lambda_\Sigma(a), \lambda_\Sigma(b)) + 1$.*
2. *If $a, b \in \Gamma$ with $a+b \in F\Gamma$ then $\lambda_\Sigma(F^{-1}(a+b)) \leq \max(\lambda_\Sigma(a), \lambda_\Sigma(b))$.*

Proof.

1. We show that if $\sigma, \tau \in \Sigma^{(n)}$ then there is $\nu \in \Sigma^{(n+1)}$ such that $[\sigma]_F + [\tau]_F = [\nu]_F$.

   We apply induction on n; the base case $n = 1$ is by axiom (iii). Assume the result holds of $n \geq 1$. Suppose $\sigma, \tau \in \Sigma^{(n+1)}$; say $\sigma = \sigma_0 s$ and $\tau = \tau_0 t$ for $\sigma_0, \tau_0 \in \Sigma^{(n)}$ and $s, t \in \Sigma$. By the induction hypothesis there is $\nu \in \Sigma^{(n+1)}$ such that $[\sigma_0]_F + [\tau_0]_F = [\nu]_F$. Write $\nu = \nu_0 v$ for $\nu_0 \in \Sigma^{(n)}$ and $v \in \Sigma$. Then

   $$[\sigma]_F + [\tau]_F = [\sigma_0]_F + F^n s + [\tau_0]_F + F^n t = [\nu]_F + F^n(s+t) = [\nu_0]_F + F^n(s+t+v).$$

   But by axiom (iii) there are $s', t' \in \Sigma$ such that $s + t + v = s' + Ft'$. So

   $$[\sigma]_F + [\tau]_F = [\nu_0]_F + F^n(s' + Ft') = [\nu_0 s' t']_F,$$

   as desired.

2. We show that if $\sigma, \tau \in \Sigma^{(n)}$ and $[\sigma]_F + [\tau]_F = Fb$ for $b \in \Gamma$ then there is $\nu \in \Sigma^{(n)}$ such that $b = [\nu]_F$.

   We apply induction on n; the base case $n = 1$ is precisely axiom (iv). Assume the result holds of $n \geq 1$. Suppose $\sigma, \tau \in \Sigma^{(n+1)}$ and $b \in \Gamma$ are such that $[\sigma]_F + [\tau]_F = Fb$; as before write $\sigma = \sigma_0 s$ and $\tau = \tau_0 t$. Then $[\sigma_0]_F + [\tau_0]_F = Fb - F^n s - F^n t = F(b - F^{n-1}s - F^{n-1}t)$. So by the inductive hypothesis there is $\nu \in \Sigma^{(n)}$ such that $b - F^{n-1}s - F^{n-1}t = [\nu]_F$. Write $\nu = \nu_0 v$ for $\nu_0 \in \Sigma^{(n-1)}$ and $v \in \Sigma$. Then

   $$b = [\nu]_F + F^{n-1}s + F^{n-1}t = [\nu_0]_F + F^{n-1}(s+t+v).$$

   But by axiom (iii) there are $s', t' \in \Sigma$ such that $s + t + v = s' + Ft'$. So

   $$b = [\nu_0]_F + F^{n-1}(s' + Ft') = [\nu_0 s' t']_F,$$

   as desired. □ Proposition 2.14

Note that in both proofs we needed axiom (iii) to apply to ternary sums, not just binary.

It will sometimes be useful to assume there is a spanning set containing given elements. The following tells us we may do so:

Proposition 2.15. *Suppose there exists an F-spanning set. If $A \subseteq \Gamma$ is finite then there is an F-spanning set containing A.*

Proof. Fix an F-spanning set Σ. By axiom (i) for $a \in A$ there is $\sigma_a \in \Sigma^*$ such that $a = [\sigma_a]_F$. Let $n = \max\{\, |\sigma_a| : a \in A \,\}$; we may assume $n \geq 1$. We will show that $[\Sigma^{(n)}]_F := \{\, [\sigma]_F : \sigma \in \Sigma, |\sigma| = n \,\}$ is F-spanning.

(i) Since $n \geq 1$ we get that $\Sigma = [\Sigma 0^{n-1}]_F \subseteq [\Sigma^{(n)}]_F$. So this follows by axiom (i) of Σ.

(ii) We immediately get that $0 = [0^n]_F \in [\Sigma^{(n)}]_F$. Given $[s_0 \cdots s_{n-1}]_F \in [\Sigma^{(n)}]_F$, note since Σ is F-spanning that each $-s_i \in \Sigma$; so $-[s_0 \cdots s_{n-1}]_F = [(-s_0) \cdots (-s_{n-1})]_F \in [\Sigma^{(n)}]_F$ as well.

(iii) We apply induction on n. The base case $n = 1$ is because Σ is F-spanning. For the induction step, suppose the axiom holds of $[\Sigma^{(n)}]_F$; suppose $\sigma, \tau, \nu \in \Sigma^{(n+1)}$. Write $\sigma = \sigma_0 s$, $\tau = \tau_0 t$, and $\nu = \nu_0 v$ for $s, t, v \in \Sigma$. So by the induction hypothesis there are $\mu_1, \mu_2 \in \Sigma^{(n)}$ such that $[\sigma_0]_F + [\tau_0]_F + [\nu_0]_F = [\mu_1]_F + F[\mu_2]_F$. By axiom (iii) applied to Σ there are $a, b \in \Sigma$ such that $s + t + v = a + Fb$. Then

$$\begin{aligned}
[\sigma]_F + [\tau]_F + [\nu]_F &= [\sigma_0]_F + F^n s + [\tau_0]_F + F^n t + [\nu_0]_F + F^n v \\
&= [\mu_1]_F + F[\mu_2]_F + F^n(a + Fb) \\
&= [\mu_1 a]_F + F[\mu_2 b]_F,
\end{aligned}$$

as desired.

(iv) This is precisely Proposition 2.14 (2). □ Proposition 2.15

It will sometimes transpire that there is an F^r-spanning set for some $r > 0$ but no F-spanning set.

Example 2.16. Consider $\Gamma = (\mathbb{Z}^2, +)$ with F given in matrix form by $\begin{pmatrix} 0 & 3 \\ 1 & 0 \end{pmatrix}$. Then $F^2 = 3I$, so by Example 2.13 (2) we get that $\{\, -2, -1, \ldots, 2 \,\}^2$ is an F^2-spanning set. Suppose for contradiction we had an F-spanning set Σ. Let $M = \sup\left\{\, |w_1| : \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \Sigma \,\right\}$

and $N = \sup\left\{ |w_2| : \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \Sigma \right\}$. The idea will be to use axiom (iii) to obtain contradictory bounds on M in terms of N and vice-versa.

Suppose $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \Sigma$. Then by axiom (iii) there are $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \Sigma$ such that $3\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + F\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} u_1 + 3v_2 \\ u_2 + v_1 \end{pmatrix}$. So

$$\begin{aligned} 3|w_1| &= |u_1 + 3v_2| \\ &\leq |u_1| + 3|v_2| \\ &\leq M + 3N \\ 3|w_2| &= |u_2 + v_1| \\ &\leq N + M. \end{aligned}$$

In particular, if we pick $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \Sigma$ such that $|w_1| = M$ we find that $3M \leq M + 3N$, and thus $M \leq \frac{3}{2}N$. If on the other hand we pick $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \Sigma$ such that $|w_2| = N$ then we obtain $3N \leq N + M$, and thus $N \leq \frac{M}{2} \leq \frac{3}{4}N$. So $N = M = 0$, a contradiction. So no F-spanning set exists.

To deal with this case, we will change our focus from F-spanning sets to F^r-spanning sets for some $r > 0$. The following result will be useful in reconciling F^r- and F^s-spanning sets for different r, s.

Proposition 2.17. *If Σ is an F-spanning set and $r > 0$ then*

$$[\Sigma^{(r)}]_F := \{\, [\sigma]_F : \sigma \in \Sigma^*, |\sigma| = r \,\}$$

is an F^r-spanning set.

Proof. We verify the axioms.

(i) Suppose $a \in \Gamma$. Since Σ is F-spanning there is $\sigma \in \Sigma^*$ such that $[\sigma]_F = a$; by appending zeroes we may assume $r \mid |\sigma|$, say $\sigma = s_0 \cdots s_{kr-1}$ for some k. Then

$$\begin{aligned} [\sigma]_F &= s_0 + Fs_1 + \cdots + F^{kr-1}s_{kr-1} \\ &= (s_0 + Fs_1 + \cdots + F^{r-1}s_{r-1}) + F^r(s_r + Fs_{r+1} + \cdots + F^{r-1}s_{2r-1}) \\ &\quad + \cdots + F^{(k-1)r}(s_{(k-1)r} + Fs_{(k-1)r+1} + \cdots + F^{r-1}s_{kr-1}) \\ &= [s_0 \cdots s_{r-1}]_F + F^r[s_r \cdots s_{2r-1}]_F + \cdots + F^{(k-1)r}[s_{(k-1)r} \cdots s_{kr-1}]_F \\ &= [[s_0 \cdots s_{r-1}]_F[s_r \cdots s_{2r-1}]_F \cdots [s_{(k-1)r} \cdots s_{kr-1}]_F]_{F^r}, \end{aligned}$$

as desired.

(ii) This is precisely as in the proof of the previous proposition.

(iii) The case $r = 1$ is because Σ is F-spanning; suppose then that $r \geq 2$. Suppose $\sigma, \tau, \nu \in \Sigma^{(r)}$. Two applications of Proposition 2.14 (1) yield $\mu \in \Sigma^{(r+2)}$ such that $[\mu]_F = [\sigma]_F + [\tau]_F + [\nu]_F$. Since $r \geq 2$, we get that $|\mu| \leq 2r$. So by the argument in the proof of axiom (i) we get that $[\sigma]_F + [\tau]_F + [\nu]_F = [\mu] \in [\Sigma^{(r)}]_F + F^r[\Sigma^{(r)}]_F$, as desired.

(iv) Suppose $\sigma, \tau \in \Sigma^{(r)}$ and $b \in \Gamma$ is such that $[\sigma]_F + [\tau]_F = F^r b$. It follows from repeated application of Proposition 2.14 (2) that there is $\nu \in \Sigma^{(r)}$ such that $b = [\nu]_F$. $\square$ Proposition 2.17

Remark 2.18. It falls out from the proof of axiom (i) that $\lambda_{[\Sigma^{(r)}]_F}(a) \leq \left\lceil \frac{\lambda_\Sigma(a)}{r} \right\rceil$.

We conclude this section by bounding $\lambda_\Sigma(F^r a)$ in terms of $\lambda_\Sigma(a)$. We immediately get an upper bound: if we fix $\sigma \in \Sigma^*$ such that $[\sigma]_F = a$ and $|\sigma| = \lambda_\Sigma(a)$ then since $[0^r\sigma]_F = F^r a$ we get that $\lambda_\Sigma(F^r a) \leq |0^r\sigma| = \lambda_\Sigma(a) + r$. We can get a weak lower bound from repeated application Proposition 2.14 (2): namely that $\lambda_\Sigma(F^r a) \geq \lambda_\Sigma(a)$. It turns out that we can do much better:

Proposition 2.19. *Suppose Σ is an F-spanning set; suppose $a \in \Gamma \setminus \Sigma$ and $r \in \mathbb{N}$. Then $\lambda_\Sigma(F^r a) \geq \lambda_\Sigma(a) + r - 1$.*

Proof. We first show by induction on r that if $|\sigma| = r$ and $[\sigma]_F \in F^r\Gamma$ then $[\sigma]_F \in F^r\Sigma$. The base case $r = 1$ is just by axiom (iv). For the induction step, suppose the claim holds of r, and suppose we are given $\sigma \in \Sigma^{(r+1)}$ with $[\sigma]_F \in F^{r+1}\Gamma$. Write $\sigma = \sigma_0 s$ for $\sigma_0 \in \Sigma^{(r)}$ and $s \in \Sigma$. Then $[\sigma_0]_F \equiv [\sigma]_F \equiv 0 \pmod{F^r\Gamma}$, so by the induction hypothesis there is $a \in \Sigma$ such that $F^r a = [\sigma_0]_F$. Then $F^r(a + s) = [\sigma]_F \in F^{r+1}\Gamma$, so $a + s \in F\Gamma$. So by axiom (iv) there is $b \in \Sigma$ such that $a + s = Fb$. Then $[\sigma]_F = F^{r+1}b$, as desired.

It follows that $\lambda_\Sigma(F^r a) > r$. Indeed, otherwise there would be $\sigma \in \Sigma^{(r)}$ such that $[\sigma]_F = F^r a$; hence the above would yield that $[\sigma]_F \in F^r\Sigma$ and thus that $a \in \Sigma$, contradicting our hypothesis.

So if $\sigma \in \Sigma^*$ is such that $[\sigma]_F = F^r a$ and $|\sigma| = \lambda_\Sigma(F^r a)$ then we can write $\sigma = \sigma_0\sigma_1$ for some $\sigma_0 \in \Sigma^{(r)}$, $\sigma_1 \in \Sigma^*$. Then $[\sigma_0]_F \equiv [\sigma]_F \equiv 0 \pmod{F^r\Gamma}$, so by the above there is $b \in \Sigma$ such that $[\sigma_0]_F = F^r b$. Then

$$F^r a = [\sigma_0\sigma_1]_F = [\sigma_0]_F + F^r[\sigma_1]_F = F^r(b + [\sigma_1]_F)$$

and thus $a = b + [\sigma_1]_F$. Then by Proposition 2.14 (1)

$$\lambda_\Sigma(a) = \lambda_\Sigma(b + [\sigma_1]_F) \leq \max(\lambda_\Sigma(b), \lambda_\Sigma([\sigma_1]_F)) + 1 \leq \max(1, |\sigma_1|) + 1 = \lambda_\Sigma(F^r a) - r + 1,$$

as desired. □ Proposition 2.19

So for $a \in \Gamma \setminus \Sigma$ and $r \in \mathbb{N}$ we get that $\lambda_\Sigma(a) + r - 1 \leq \lambda_\Sigma(F^r a) \leq \lambda_\Sigma(a) + r$.

2.4 F-automatic sets

Throughout this section we assume there is a spanning set for some power of F.

Definition 2.20. If Σ is an F^r-spanning set for some $r > 0$, we say $A \subseteq \Gamma$ is (F^r, Σ)-*automatic* if $\{ \sigma \in \Sigma^* : [\sigma]_{F^r} \in A \}$ is regular. We say A is F-*automatic* if it is (F^r, Σ)-automatic for some $r > 0$ and some F^r-spanning set Σ.

Some examples:

Example 2.21.

1. Let R be a finite ring, and consider $\Gamma = (R[t], +)$ with $Ff = tf$. Notice that $(\Gamma, +, F)$ is isomorphic to the case we considered in Example 2.13 (4); from this we conclude that $\Sigma := R$ is F-spanning. The following are F-automatic:

   - The set A_1 of monomials, since $\{ \sigma \in \Sigma^* : [\sigma]_F \in A_1 \} = 0^* R 0^*$ is regular.
   - The set A_2 of monic polynomials, since $\{ \sigma \in \Sigma^* : [\sigma]_F \in A_1 \} = R^* 1 0^*$ is regular.
   - The set A_3 of $f \in R[t]$ with some fixed unit $a \in R^\times$ as a root. Indeed, fix N such that $a^N = 1$, and consider the DFA $M = (\Sigma, Q, q_0, \Omega, \delta)$ where

     $$\begin{aligned} Q &= R \times \mathbb{Z}/N\mathbb{Z} \\ q_0 &= (0, 0 + N\mathbb{Z}) \\ \Omega &= \{ 0 \} \times \mathbb{Z}/N\mathbb{Z} \\ \delta((b, i + N\mathbb{Z}), c) &= (b + ca^i, i + 1 + N\mathbb{Z}). \end{aligned}$$

     One can show by a quick induction that if $s_0 \cdots s_{n-1} \in \Sigma^*$ then

     $$\delta(q_0, s_0 \cdots s_{n-1}) = (s_0 + s_1 a + \cdots + s_{n-1} a^{n-1}, n + N\mathbb{Z}) = ([\sigma]_F(a), n + N\mathbb{Z}).$$

     So M accepts σ if and only if a is a root of $[\sigma]_F$, and A_3 is F-automatic.

2. Consider $\Gamma = (\mathbb{Z}^m, +)$ and F is $a \mapsto da$ for some $d \geq 2$. We will see in Corollary 2.30 that every d-automatic set is F-automatic. In particular, $d^{\mathbb{N}}$ is an F-automatic subset of $\mathbb{Z}$.

We will see more examples as we go along. In particular, we will see in Section 2.6 that the F-sets arising in the positive-characteristic Diophantine geometric problem considered in [18] are F-automatic.

We wish to show that F-automaticity can be verified in any spanning set for any power of F. To do this, we will characterize F-automaticity in terms of finiteness of kernels; this generalizes a similar characterization of d-automatic sets.

Definition 2.22. Suppose $S \subseteq \Gamma$ is finite and $A \subseteq \Gamma$. Given $\sigma = s_0 \cdots s_{n-1} \in S^*$ we let

$$A_{F,\sigma} = \{\, a \in \Gamma : [\sigma a]_F \in A \,\} = \{\, a \in \Gamma : s_0 + Fs_1 + \cdots + F^{n-1}s_{n-1} + F^n a \in A \,\}.$$

We define the (F, S)-*kernel of* A to be $\ker_{F,S}(A) = \{\, A_{F,\sigma} : \sigma \in S^* \,\}$.

Remark 2.23. If $\sigma = s_0 \cdots s_{n-1}, \tau = t_0 \cdots t_{n'-1} \in S^*$ then

$$\begin{aligned} & a \in A_{F,\sigma\tau} \\ \iff & s_0 + Fs_1 + \cdots + F^{n-1}s_{n-1} + F^n t_0 + F^{n+1}t_1 + \cdots + F^{n+n'-1}t_{n'-1} + F^{n+n'}a \in A \\ \iff & t_0 + Ft_1 + \cdots + F^{n'-1}t_{n'-1} + F^{n'}a \in A_{F,\sigma} \\ \iff & a \in (A_{F,\sigma})_{F,\tau}. \end{aligned}$$

In particular, given $X \subseteq \mathcal{P}(\Gamma)$, in order to show $\ker_{F,S}(A) \subseteq X$ it suffices to show that $A \in X$, and that if $B \in X$ and $s \in S$ then $B_{F,s} \in X$.

A sample computation:

Example 2.24. Consider $\Gamma = (\mathbb{Z}, +)$ equipped with multiplication by some $d \geq 2$. Let $S = \{\, 0, \ldots, d-1 \,\}$; we compute $\ker_{d,S}(d^{\mathbb{N}})$. Let $X = \{\, d^{\mathbb{N}}, \{\, 0 \,\}, \emptyset \,\}$; I claim that $\ker_{d,S}(d^{\mathbb{N}}) \subseteq X$. Using the previous remark, it suffices to show that if $B \in X$ and $s \in S$ then $B_{d,s} \in X$.

- Consider the case $B = d^{\mathbb{N}}$. If $s \in S$ and $a \in \mathbb{Z}$ then $a \in B_{d,s} \iff s + da \in d^{\mathbb{N}}$. If $s = 0$, then this is equivalent to $a \in d^{\mathbb{N}}$; so $B_{d,0} = d^{\mathbb{N}} \in X$. If $s = 1$, then $s + da \equiv 1 \pmod{d}$. Hence $s + da \in d^{\mathbb{N}}$ if and only if $s + da = 1$; i.e., $a = 0$. So $B_{d,1} = \{\, 0 \,\} \in X$. Finally, if $s > 1$, then since $s + da \equiv s \pmod{d}$ we get that $s + da \notin d^{\mathbb{N}}$ for any $a \in \mathbb{Z}$; so $B_{d,s} = \emptyset \in X$.

- Consider the case $B = \{ 0 \}$. If $s \in S$ and $a \in \mathbb{Z}$ then $a \in B_{d,s} \iff s + da = 0$. It is clear that there is no such a if $s \neq 0$, and otherwise this is equivalent to $a = 0$; so $B_{d,s}$ is either $\{ 0 \}$ or $\emptyset$, and both of these lie in X.
- It is clear that $\emptyset_{d,s} = \emptyset \in X$ for all $s \in S$.

Moreover the above shows that $X = \ker_{d,S}(d^{\mathbb{N}})$, since every $B \in X$ occurs as $(d^{\mathbb{N}})_{d,s}$ for some $s \in S$.

The point of kernels is that automaticity can be characterized in terms of finiteness of kernels. The simplest case of this is when $S = \Sigma$ is an F^r-spanning set:

Proposition 2.25. *Suppose Σ is an F^r-spanning set and $A \subseteq \Gamma$. Then A is (F^r, Σ)-automatic if and only if $\ker_{F^r,\Sigma}(A)$ is finite.*

Proof. The idea is that elements of the kernel will be roughly in correspondence with states in a DFA witnessing (F^r, Σ)-automaticity.

By replacing F with F^r we may assume $r = 1$.

($\implies$) Suppose A is (F, Σ)-automatic; fix a DFA $(\Sigma, Q, q_0, \delta, \Omega)$ recognizing $\{ \sigma \in \Sigma^* : [\sigma]_F \in A \}$.

Suppose $\sigma, \tau \in \Sigma^*$ are such that $\delta(q_0, \sigma) = \delta(q_0, \tau)$. We will show that $A_{F,\sigma} = A_{F,\tau}$. Suppose $a \in \Gamma$; since Σ is F-spanning we can find $\nu \in \Sigma^*$ such that $a = [\nu]_F$. Then

$$\begin{aligned} [\sigma a]_F \in A &\iff [\sigma\nu]_F \in A \\ &\iff \delta(\delta(q_0, \sigma), \nu) \in \Omega \\ &\iff \delta(\delta(q_0, \tau), \nu) \in \Omega \\ &\iff [\tau\nu]_F \in A \\ &\iff [\tau a]_F \in A. \end{aligned}$$

So $A_{F,\sigma} = A_{F,\tau}$. So, since there are finitely many possibilities for $\delta(q_0, \sigma)$, we get that $\ker_{F,S}(A)$ is finite.

($\impliedby$) Suppose $\ker_{F,\Sigma}(A)$ is finite. We describe a DFA recognizing $\{ \sigma \in \Sigma^* : [\sigma]_F \in A \}$. We let

$$\begin{aligned} Q &= \ker_{F,\Sigma}(A) \\ q_0 &= A = A_{F,\varepsilon} \\ \Omega &= \{ A_{F,\sigma} \in \ker_{F,\Sigma}(A) : 0 \in A_{F,\sigma} \} \\ \delta(A_{F,\sigma}, \ell) &= (A_{F,\sigma})_{F,\ell} = A_{F,\sigma\ell} \end{aligned}$$

for $\sigma \in \Sigma^*$ and $\ell \in \Sigma$. One can show by induction that $\delta(q_0, \sigma) = A_{F,\sigma}$. So our machine accepts σ if and only if $0 \in A_{F,\sigma}$; i.e., if and only if $[\sigma]_F \in A$, as desired. □ Proposition 2.25

In fact we can dispense with the requirement that S be a spanning set, as long as it contains a representative of each coset of $F^r\Gamma$. (Note that axiom (i) of spanning sets implies that $\Gamma/F^r\Gamma$ is finite, so we can always find such S finite.)

Lemma 2.26. *Suppose $S, T \subseteq \Gamma$ are finite and contain a representative of each coset of $F^r\Gamma, F^{r'}\Gamma$, respectively. If $A \subseteq \Gamma$ then $\ker_{F^r,S}(A)$ is finite if and only if $\ker_{F^{r'},T}(A)$ is.*

Proof. We first do the case $r = r' = 1$ under the assumption that there is an F-spanning set Σ (as opposed to a spanning set for some power of F).

Note first that $[S^{(n)}]_F = \{ s_0 + Fs_1 + \cdots + F^{n-1}s_{n-1} : s_0, \dots, s_{n-1} \in S \}$ contains a representative of each coset of $F^n\Gamma$. Indeed, given $a + F^n\Gamma$ we can find $s_0 \in S$ such that $a \equiv s_0 \pmod{F\Gamma}$; we then recursively find $s_1, \dots, s_{n-1}$ such that $s_1 + Fs_2 + \cdots + F^{n-2}s_{n-1} \equiv F^{-1}(a - s_0) \pmod{F^{n-1}\Gamma}$. Likewise with $[T^{(n)}]_F$.

Given $A_{F,\sigma} \in \ker_{F,S}(A)$, say with $|\sigma| = n$, we can thus find $\tau \in T^*$ of length n such that $[\sigma]_F \equiv [\tau]_F \pmod{F^n\Gamma}$. Note that

$$\begin{aligned}
A_{F,\tau} &= \{ a \in \Gamma : [\tau]_F + F^n a \in A \} \\
&= \{ a \in \Gamma : [\sigma]_F + F^n(a + F^{-n}([\tau]_F - [\sigma]_F)) \in A \} \\
&= \{ a \in \Gamma : a + F^{-n}([\tau]_F - [\sigma]_F) \in A_{F,\sigma} \} \\
&= A_{F,\sigma} + F^{-n}([\sigma]_F - [\tau]_F).
\end{aligned}$$

We will show that there are finitely many possible values for $F^{-n}([\sigma]_F - [\tau]_F)$. In particular, it will follow that $\ker_{F,T}(A)$ is contained in finitely many translates of $\ker_{F,S}(A)$; so if the latter is finite, then so is the former.

Recall our function $\lambda_\Sigma \colon \Gamma \to \mathbb{N}$ that maps $a \in \Gamma$ to the length of the shortest $\sigma \in \Sigma^*$ such that $[\sigma]_F = a$. We will bound $\lambda_\Sigma(F^{-n}([\sigma]_F - [\tau]_F))$; since there are finitely many $\sigma \in \Sigma^*$ of any given length, this will suffice.

Let $M = \max\{ \lambda_\Sigma(s - t) : s \in S, t \in T \}$. We first bound $\lambda_\Sigma([\sigma]_F - [\tau]_F)$: we show by induction on $n = |\sigma| = |\tau|$ that $\lambda_\Sigma([\sigma]_F - [\tau]_F) \le M + n$. The case $n = 0$ is vacuous, and the case $n = 1$ is simply the fact that $\lambda_\Sigma(s_0 - t_0) \le M$. For the induction step, suppose the result holds of n, and suppose $|\sigma| = |\tau| = n + 1$. Write $\sigma = \sigma_0 s$ for some $s \in S$ and $\tau = \tau_0 t$ for some $t \in T$. Note that

$$\lambda_\Sigma(F^n s - F^n t) = \lambda_\Sigma(F^n(s - t)) \le \lambda_\Sigma(s - t) + n \le M + n,$$

where the first inequality is because if $[\nu]_F = s - t$ then $[0^n\nu]_F = F^n(s-t)$. Moreover by the induction hypothesis $\lambda_\Sigma([\sigma_0]_F - [\tau_0]_F) \le M + n$. Thus

$$\begin{aligned}\lambda_\Sigma([\sigma]_F - [\tau]_F) &= \lambda_\Sigma(([\sigma_0]_F - [\tau_0]_F) + (F^n s - F^n t)) \\ &\le \max\{\, \lambda_\Sigma([\sigma_0]_F - [\tau_0]_F), \lambda_\Sigma(F^n s - F^n t) \,\} + 1 \text{ (Proposition 2.14 (1))} \\ &\le M + n + 1,\end{aligned}$$

as desired.

Recall from Proposition 2.19 that if $a \notin \Sigma$ then $\lambda_\Sigma(F^n a) \ge \lambda_\Sigma(a) + n - 1$. From this, and the above bound on $\lambda_\Sigma([\sigma]_F - [\tau]_F)$, it follows that $\lambda_\Sigma(F^{-n}([\sigma]_F - [\tau]_F)) \le M + 1$. Indeed, otherwise we would have $\lambda_\Sigma([\sigma]_F - [\tau]_F) = \lambda_\Sigma(F^n(F^{-n}([\sigma]_F - [\tau]_F))) > M+1+n-1 = M+n$, contradicting our bound.

So we always have that $F^{-n}([\sigma]_F - [\tau]_F) \in [\Sigma^{(M+1)}]_F$. So

$$\ker_{F,T}(A) \subseteq \{\, B + \delta : B \in \ker_{F,S}(A), \delta \in [\Sigma^{(M+1)}]_F \,\}.$$

So if $\ker_{F,S}(A)$ is finite, then so too is $\ker_{F,T}(A)$. By symmetry, we have proven the case $r = s = 1$ under the assumption that there is an F-spanning set.

We now move on to the general case. By our standing assumption, there is an F^s-spanning set Σ for some $s > 0$. By Proposition 2.17 we may assume $r, r' \mid s$; say $s = ir = jr'$. As noted earlier, $S' := [S^{(i)}]_{F^r}$ contains a representative of each coset of $F^s\Gamma$; likewise with $T' := [T^{(j)}]_{F^{r'}}$. By the above special case of the lemma (applied to F^s), we have that $\ker_{F^s,S'}(A)$ is finite if and only if $\ker_{F^s,T'}(A)$ is. So it suffices to show that $\ker_{F^r,S}(A)$ is finite if and only if $\ker_{F^s,S'}(A)$ is (at which point the analogous result about T will follow by symmetry). Replacing F with F^r, we are left to show that if S contains a representative of each coset of F and $n \in \mathbb{N}$ then $\ker_{F,S}(A)$ is finite if and only if $\ker_{F^n,S'}(A)$ is, where $S' = [S^{(n)}]_F$.

$(\implies)$ I claim that $\ker_{F^n,S'}(A) = \{\, A_{F,\sigma} : n \mid |\sigma| \,\}$. Indeed, if

$$\sigma = [s_0 \cdots s_{n-1}]_F \cdots [s_{(k-1)n} \cdots s_{kn-1}]_F \in (S')^*$$

and $a \in \Gamma$ then

$$\begin{aligned} & a \in A_{F^n,\sigma} \\ \iff & [s_0 \cdots s_{n-1}]_F + F^n[s_n \cdots s_{2n-1}]_F + \cdots + F^{(k-1)n}[s_{(k-1)n} \cdots s_{kn-1}] + F^{kn}a \in A \\ \iff & [s_0 s_1 \cdots s_{kn-1}]_F + F^{kn}a \in A \\ \iff & a \in A_{F,s_0 s_1 \cdots s_{kn-1}}. \end{aligned}$$

So $A_{F^n,\sigma} = A_{F,s_0 s_1 \cdots s_{kn-1}}$. Conversely if $s_0 s_1 \cdots s_{kn-1} \in S^*$ has length divisible by n, then the above equivalence shows that $A_{F,s_0 \cdots s_{kn-1}} = A_{F^n,\sigma} \in \ker_{F^n,S'}(A)$.

In particular, we get that $\ker_{F^n,S'}(A) \subseteq \ker_{F,S}(A)$; so if the latter is finite, then so is the former.

($\Longleftarrow$) Suppose we are given $\sigma \in S^*$; pick $k < n$ such that $n \mid |\sigma| + k$. Note that we can recover $A_{F,\sigma}$ from the set of $A_{F,\sigma\tau}$ as τ ranges over S^k. Indeed, fix $a \in \Gamma$. Since S' contains a representative of each coset of $F^n\Gamma$, it must also contain a representative of each coset of $F^k\Gamma$; so we can find some $\tau \in S^{(k)}$ such that $a \equiv [\tau]_F \pmod{F^k\Gamma}$. Then

$$\begin{aligned} a \in A_{F,\sigma} &\iff [\tau]_F + F^k(F^{-k}(a - [\tau]_F)) \in A_{F,\sigma} \\ &\iff F^{-k}(a - [\tau]_F) \in (A_{F,\sigma})_{F,\tau} \\ &\iff F^{-k}(a - [\tau]_F) \in A_{F,\sigma\tau}. \end{aligned}$$

So we can indeed recover $A_{F,\sigma}$ from the $A_{F,\sigma\tau}$:

$$A_{F,\sigma} = \bigcup_{\tau \in S^{(k)}} ([\tau]_F + F^k(A_{F,\sigma\tau})).$$

In particular, if $\ker_{F^n,S'}(A)$ is finite, then as noted in the left-to-right direction there are only finitely many $A_{F,\sigma\tau}$; hence there are only finitely many $A_{F,\sigma}$, and $\ker_{F,S}(A)$ is finite. $\square$ Lemma 2.26

Corollary 2.27. *Suppose $A \subseteq \Gamma$, $r_1, r_2 > 0$, Σ is an F^{r_1}-spanning set, and $S \subseteq \Gamma$ is finite and contains a representative of each coset of $F^{r_2}\Gamma$. The following are equivalent:*

1. *A is F-automatic.*

2. *A is (F^{r_1}, Σ)-automatic.*

3. *$\ker_{F^{r_2},S}(A)$ is finite.*

Proof. (3) $\Longrightarrow$ (2) is Lemma 2.26 and Proposition 2.25, and (2) $\Longrightarrow$ (1) is by definition. Suppose (1) holds; say there is an $F^{r'}$-spanning set Σ' such that A is $(F^{r'}, \Sigma')$-automatic. Then again by Lemma 2.26 and Proposition 2.25 we get that (3) holds. $\square$ Corollary 2.27

In particular, as we claimed much earlier, F-automaticity is independent of the spanning set chosen (and indeed, of the power of F for which a spanning set is chosen). Using this, we can deduce some elementary properties of F-automaticity:

Corollary 2.28. *Suppose $A \subseteq \Gamma$ and $r > 0$. Then A is F-automatic if and only if A is F^r-automatic.*

Proof. The right-to-left direction is by definition. Suppose for the left-to-right that A is F-automatic. By Proposition 2.17 we may assume there is an $F^{r'}$-spanning set Σ for some $r' \in r\mathbb{N}$. Then by Corollary 2.27 A is $(F^{r'}, \Sigma)$-automatic, and hence is F^r-automatic. $\square$ Corollary 2.28

Corollary 2.29. *F-automatic sets are closed under Boolean combinations.*

Proof. Suppose $A, B \subseteq \Gamma$ are F-automatic; we will show that $A \cup B$ and $\Gamma \setminus A$ are as well.

Fix $r > 0$ for which there is an F^r-spanning set Σ. By Corollary 2.27 both $\{ \sigma \in \Sigma^* : [\sigma]_{F^r} \in A \}$ and $\{ \sigma \in \Sigma^* : [\sigma]_{F^r} \in B \}$ are regular. So their union $\{ \sigma \in \Sigma^* : [\sigma]_{F^r} \in A \cup B \}$ is regular, and $A \cup B$ is F-automatic. The case of complements is similar, except we use Corollary 2.5 for the fact that the complement of a regular language is regular. $\square$ Corollary 2.29

Using the finite kernel characterization of automaticity, we can now see that our notion of F-automaticity includes the notion of d-automaticity on $\mathbb{Z}^m$:

Corollary 2.30. *Suppose $d \in \mathbb{N}$ with $d \geq 2$; let $F \colon \mathbb{Z}^m \to \mathbb{Z}^m$ be multiplication by d. Then $A \subseteq \mathbb{Z}^m$ is d-automatic if and only if it is F-automatic.*

Proof. Recall from Example 2.13 (2) that there is an F-spanning set for $(\mathbb{Z}^m, F)$; so F-automaticity is indeed defined. We mentioned previously that our kernel characterization of automaticity generalizes a similar characterization of d-automatic sets; in this context, the known characterization tells us that $A \subseteq \mathbb{Z}^m$ is d-automatic if and only if

$$\{ \{ \mathbf{a} \in \mathbb{Z}^m : \mathbf{b} + d^n \mathbf{a} \} : n \in \mathbb{N}, \mathbf{b} \in \{ 0, \ldots, d^n - 1 \}^m \}$$

is finite (see [1, Proposition 5.1]). But if we let $S = \{ 0, \ldots, d-1 \}^m$ then this is just $\ker_{F,S}(A)$, which by Corollary 2.27 is finite if and only if A is F-automatic. $\square$ Corollary 2.30

We can extend the notion of F-automaticity to subsets of Γ^m:

Definition 2.31. We say $A \subseteq \Gamma^m$ is F-automatic if it is $(F \times \cdots \times F)$-automatic in $(\Gamma^m, F \times \cdots \times F)$. Note that $(\Gamma^m, F \times \cdots \times F)$ does admit a spanning set: we saw in Example 2.13 (2) that if Σ is F-spanning in Γ then Σ^m is $(F \times \cdots \times F)$-spanning in Γ^m.

In general when working with subsets of Γ^m we will conflate F with $F \times \cdots \times F$.

Proposition 2.32. *Suppose Γ admits an F^r-spanning set for some $r > 0$. The following are F-automatic:*

1. *The diagonal* $\left\{ \begin{pmatrix} a \\ a \end{pmatrix} : a \in \Gamma \right\}$ *in* Γ^2.
2. *The graph of addition* $\left\{ \begin{pmatrix} a \\ b \\ a+b \end{pmatrix} : a, b \in \Gamma \right\}$ *in* Γ^3.
3. *The graph of F in* Γ^2.

This result is a bit surprising. We assume only that elements of Γ can be nicely described using strings over some finite alphabet (i.e., that there is a spanning set for some power of F). The conclusion is a very strong computability condition on Γ: that there is a finite automaton that decides whether two strings represent the same element, and a finite automaton that computes the sum in Γ of two strings.

All of the above are examples of F-invariant subgroups of powers of Γ; in fact we will see in Theorem 2.54 that any F-invariant subgroup of any Γ^m is F-automatic.

Proof.

1. Let Δ denote the diagonal. Fix $S \subseteq \Gamma$ containing exactly one representative of each coset of $F\Gamma$; note that S^2 contains exactly one representative of each coset of $F\Gamma^2$. If $\mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \in S^2$ with s_1, s_2 distinct then $\Delta_{F,\mathbf{s}} = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \Gamma^2 : Fa + s_1 = Fb + s_2 \right\} = \emptyset$ since
$$Fa + s_1 \equiv s_1 \not\equiv s_2 \equiv Fb + s_2 \pmod{F\Gamma}.$$
If on the other hand $s_1 = s_2$ then $\Delta_{F,\mathbf{s}} = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \Gamma^2 : Fa + s_1 = Fb + s_1 \right\} = \Delta$ since F is injective. Proceeding inductively we see that for all $\sigma \in (S^2)^*$ we have $\Delta_{F,\sigma} \in \{ \Delta, \emptyset \}$. So $\ker_{F,S^2}(\Delta)$ is finite, and Δ is F-automatic.

2. Let A denote the graph of addition. Fix $S \subseteq \Gamma$ containing exactly one representative of each coset of $F\Gamma$; let
$$S' = \left\{ \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} : s_1, s_2, s_3 \in S, s_1 + s_2 \not\equiv s_3 \pmod{F\Gamma} \right\} \cup \left\{ \begin{pmatrix} s_1 \\ s_2 \\ s_1 + s_2 \end{pmatrix} : s_1, s_2 \in S \right\}.$$

So S' contains a representative of each coset of $F\Gamma^3$, and is such that if $\begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} \in S'$ and $s_3 \neq s_1 + s_2$ then $s_3 \not\equiv s_1 + s_2 \pmod{F\Gamma}$. Then given $\mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} \in S'$ with $s_1 + s_2 \neq s_3$, we have

$$A_{F,\mathbf{s}} = \left\{ \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \Gamma^3 : Fa + s_1 + Fb + s_2 = Fc + s_3 \right\} = \emptyset,$$

since if we had such a, b, c then we would have $s_1 + s_2 \equiv s_3 \pmod{F\Gamma}$, a contradiction. On the other hand, if $s_1 + s_2 = s_3$ then

$$A_{F,\mathbf{s}} = \left\{ \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \Gamma^3 : Fa + s_1 + Fb + s_2 = Fc + s_3 \right\} = A,$$

since F is injective. So again we inductively get that $\ker_{F,S}(A) = \{ A, \emptyset \}$, and thus that A is F-automatic.

3. Let B denote the graph of F. Fix $S \subseteq \Gamma$ containing exactly one representative of each coset of $F\Gamma$; so as before S^2 contains a representative of each coset of $F\Gamma^2$. We may assume that $0 \in S$.

   I claim that $\ker_{F,S^2}(B)$ contains only $\emptyset$ and sets of the form $C = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} : b = s + Fa \right\}$ for $s \in S$. Indeed, since $0 \in S$ we get that B takes the given form. Suppose we are given such a C and $\mathbf{t} = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \in S^2$. Then

   $$\begin{pmatrix} a \\ b \end{pmatrix} \in C_{F,\mathbf{t}} \iff \begin{pmatrix} t_1 + Fa \\ t_2 + Fb \end{pmatrix} \in C \iff t_2 + Fb = s + Ft_1 + F^2 a.$$

   If $t_2 \neq s$, so $t_2 \not\equiv s \pmod{F\Gamma}$, then $t_2 + Fb \equiv t_2 \not\equiv s \equiv s + Ft_1 + F^2 a \pmod{F\Gamma}$, and there are no such $\begin{pmatrix} a \\ b \end{pmatrix}$. If on the other hand $t_2 = s$, then by injectivity of F we get that $C_{F,\mathbf{t}} = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} : b = t_1 + Fa \right\}$ takes the desired form.

   So by an inductive argument, our claim about $\ker_{F,S^2}(B)$ holds. In particular, $\ker_{F,S^2}(B)$ is finite, and thus B is F-automatic. $\square$ Proposition 2.32

The following proposition tells us that to show a set is automatic, it suffices to find some set of representations of its elements that is regular; this is sometimes considerably simpler than showing that the set of all representations of its elements is.

Proposition 2.33. *If Σ is an F^r-spanning set for some $r > 0$ and $L \subseteq \Sigma^*$ is regular then $[L]_{F^r}$ is F-automatic.*

Proof. Note that given $\sigma \in \Sigma^*$, the question of whether $[\sigma]_{F^r} \in [L]_{F^r}$ is an existential one: we are asking whether there exists $\tau \in L$ such that $[\sigma]_{F^r} = [\tau]_{F^r}$. We thus look to apply the fact that regular languages are closed under projections (Corollary 2.8).

It suffices to check the case $r = 1$. Since L is regular so too is $L0^*$. One can use an automaton recognizing $L0^*$ to construct an automaton recognizing $\left\{ \begin{pmatrix} \sigma \\ \tau \end{pmatrix} \in (\Sigma^2)^* : \tau \in L0^* \right\}$; hence the latter is also regular. By Corollary 2.27 and Proposition 2.32

$$\left\{ \begin{pmatrix} \sigma \\ \tau \end{pmatrix} \in (\Sigma^2)^* : [\sigma]_F = [\tau]_F \right\}$$

is regular. So since regular languages are closed under Boolean combinations (Corollary 2.5) their intersection

$$\left\{ \begin{pmatrix} \sigma \\ \tau \end{pmatrix} \in (\Sigma^2)^* : [\sigma]_F = [\tau]_F, \tau \in L0^* \right\}$$

is regular; hence by closure of regular languages under projection (Corollary 2.8) its projection

$$L' = \{ \sigma \in \Sigma^* : \exists \tau \in L0^* \text{ such that } |\sigma| = |\tau|, [\sigma]_F = [\tau]_F \}$$

is regular. Finally, note that $L'' := \{ \sigma \in \Sigma^* : \sigma 0^* \cap L' \neq \emptyset \}$ is regular: given a DFA $(\Sigma, Q, q_0, \Omega, \delta)$ recognizing L', the DFA

$$(\Sigma, Q, q_0, \{ q \in Q : \delta(q, 0^i) \in \Omega \text{ for some } i \in \mathbb{N} \}, \delta)$$

recognizes L''. But if $\sigma \in \Sigma^*$ then

$$\begin{aligned} \sigma \in L'' &\iff \text{there is } i \in \mathbb{N} \text{ such that } \sigma 0^i \in L' \\ &\iff \text{there is } i \in \mathbb{N}, \tau \in L0^* \text{ such that } |\tau| = |\sigma| + i \text{ and } [\sigma 0^i]_F = [\tau]_F \\ &\iff \text{there is } \tau \in L0^* \text{ such that } |\tau| \geq |\sigma| \text{ and } [\sigma]_F = [\tau]_F \\ &\iff [\sigma]_F \in [L]_F. \end{aligned}$$

So $\{ \sigma \in \Sigma^* : [\sigma]_F \in [L]_F \}$ is regular, and $[L]_F$ is F-automatic. □ Proposition 2.33

Another closure property of automatic sets:

Proposition 2.34. *Suppose $A \subseteq \Gamma^{m+1}$ is F-automatic. Then the projection A_0 of A away from the last coordinate is F-automatic in Γ^m.*

Proof. By Corollary 2.28 it suffices to check the case where there is an F-spanning set Σ. Recall that in Example 2.13 we showed that Σ^{m+1} is F-spanning for Γ^{m+1}. So by Corollary 2.27 $\left\{ \begin{pmatrix} \boldsymbol{\sigma} \\ \tau \end{pmatrix} \in ((\Sigma^m) \times \Sigma)^* : \begin{pmatrix} [\boldsymbol{\sigma}]_F \\ [\tau]_F \end{pmatrix} \in A \right\}$ is regular. Then by closure of regular languages under projection (Corollary 2.8) we get that

$$L := \left\{ \boldsymbol{\sigma} \in (\Sigma^m)^* : \exists \tau \in \Sigma^* \text{ such that } |\boldsymbol{\sigma}| = |\tau|, \begin{pmatrix} [\boldsymbol{\sigma}]_F \\ [\tau]_F \end{pmatrix} \in A \right\}$$

is regular.

It's clear that $[L]_F \subseteq A_0$. Conversely if $\mathbf{a} \in A_0$, say with $b \in \Gamma$ such that $\begin{pmatrix} \mathbf{a} \\ b \end{pmatrix} \in A$, then we can find $\boldsymbol{\sigma} \in (\Sigma^m)^*, \tau \in \Sigma^*$ such that $[\boldsymbol{\sigma}]_F = \mathbf{a}$ and $[\tau]_F = b$. Pick i, j such that $|\boldsymbol{\sigma}| + i = |\tau| + j$. Then $\begin{pmatrix} [\boldsymbol{\sigma}\mathbf{0}^i]_F \\ [\tau 0^j]_F \end{pmatrix} = \begin{pmatrix} \mathbf{a} \\ b \end{pmatrix} \in A$ and $|\boldsymbol{\sigma}\mathbf{0}^i| = |\tau 0^j|$. So $\boldsymbol{\sigma}\mathbf{0}^i \in L$; thus $\mathbf{a} = [\boldsymbol{\sigma}]_F = [\boldsymbol{\sigma}\mathbf{0}^i]_F \in [L]_F$.

So $A_0 = [L]_F$, and $L \subseteq (\Sigma^m)^*$ is regular. So Proposition 2.33 yields that A_0 is F-automatic. □ Proposition 2.34

Corollary 2.35. *Every definable set in the following structures is F-automatic:*

1. $(\Gamma, 0, +, -, F)$.
2. *The structure $\mathcal{G}$ with domain Γ and a predicate for every F-automatic subset of every Γ^m.*

Here we use "structure" in the sense of first-order logic, and we use "definable" to mean definable with parameters.

Proof. The definable sets of $(\Gamma, 0, +, -, F)$ coincide with those of $(\Gamma, +, F)$, which by Proposition 2.32 is a reduct of $\mathcal{G}$; so it suffices to check that the definable sets of $\mathcal{G}$ are F-automatic. We first prove the claim for 0-definable sets. The atomic formulas define F-automatic sets: this is by Proposition 2.32 (for equality) and the definition of $\mathcal{G}$ (for all other atomic formulas). The induction then follows from closure under Boolean combinations (Corollary 2.29) and projections (Proposition 2.34).

To prove the result for all definable sets, note from Proposition 2.33 that singletons are F-automatic, and hence 0-definable in $\mathcal{G}$. So every definable set in $\mathcal{G}$ is 0-definable, and hence is F-automatic by the above. □ Corollary 2.35

A particular case of the above is that if $A, B \subseteq \Gamma$ are F-automatic then so is $A + B$.

2.5 When spanning sets exist

We turn our attention to the question of when Γ admits an F^r-spanning set for some $r > 0$. Unfortunately, this isn't always the case, even when Γ is finitely generated:

Example 2.36. Let $\Gamma = \mathbb{Z}^2$ and $F\colon \Gamma \to \Gamma$ be given by some invertible $T \in M_2(\mathbb{Z})$ that has two complex eigenvalues μ, ν with $|\mu| > 1$ and $|\nu| < 1$. (One could, for example, take

$$T = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

which has eigenvalues $\varphi, -\varphi^{-1}$ where $\varphi = \frac{1+\sqrt{5}}{2}$.)

We first show there is no F-spanning set; suppose for contradiction that Σ were F-spanning. Pick eigenvectors u for μ and v for ν; given $w \in \mathbb{C}^2$ write $w = f_u(w)u + f_v(w)v$. Note that $f_v\colon \mathbb{C}^2 \to \mathbb{C}$ is linear, and if $w \in \mathbb{Z}^2$ then $f_v(Fw) = \nu f_v(w)$. Let $M = \max\{\, |f_v(a)| : a \in \Sigma \,\}$. Then given $s_0 \cdots s_{n-1} \in \Sigma^*$ we have

$$\begin{aligned}
|f_v([s_0 \cdots s_{n-1}]_F)| &= |f_v(s_0 + Fs_1 + \cdots + F^{n-1}s_{n-1})| \\
&= |f_v(s_0) + \nu f_v(s_1) + \cdots + \nu^{n-1} f_v(s_{n-1})| \\
&\leq |f_v(s_0)| + |\nu f_v(s_1)| + \cdots + |\nu^{n-1} f_v(s_{n-1})| \\
&\leq M(1 + |\nu| + \cdots + |\nu|^{n-1}) \\
&\leq M \sum_{i=0}^{\infty} |\nu|^i \\
&= M \frac{1}{1 - |\nu|}
\end{aligned}$$

since $|\nu| < 1$. So by axiom (i) we get that $|f_v(w)| \leq M\frac{1}{1-|\nu|}$ for all $w \in \mathbb{Z}^2$. Since f_v is linear and $2w \in \mathbb{Z}^2$ if $w \in \mathbb{Z}^2$, this implies that $f_v(w) = 0$ for all $w \in \mathbb{Z}^2$. So $\mathbb{Z}^2 \subseteq \mathbb{C}u$, and $\mathbb{Z}^2$ spans a one-dimensional subspace of $\mathbb{C}^2$, a contradiction.

So there is no F-spanning set. But if $r > 0$ then F^r also satisfies the hypotheses; so there is no F^r-spanning set for any $r > 0$.

In the above example we exhibited a (Γ, F) for which no finite set could satisfy axiom (i) of spanning sets. It is also possible for there to exist finite sets satisfying axiom (i), but nonetheless for there to be no spanning set.

Example 2.37. Consider any infinite finitely generated abelian group Γ; let $F = \mathrm{id}_\Gamma$. If $\{ \gamma_1, \ldots, \gamma_n \}$ is any generating set for Γ then $\Sigma = \{ \pm\gamma_1, \ldots, \pm\gamma_n \}$ satisfies axiom (i) of spanning sets: if $a = k_1\gamma_1 + \cdots + k_n\gamma_n \in \Gamma$ for $k_1, \ldots, k_n \in \mathbb{Z}$, then

$$a = [(\mathrm{sgn}(k_1)\gamma_1)^{|k_1|} \cdots (\mathrm{sgn}(k_n)\gamma_n)^{|k_n|}]_F.$$

But when $F = \mathrm{id}_\Gamma$ axiom (iv) simply says that Σ is closed under addition, which cannot be true of any finite Σ that satisfies axiom (i) (since that would imply that Σ contains an element of infinite order). So there is no F-spanning set (and hence no F^r-spanning set, since $F = F^r$), despite the existence of finite sets that satisfy axiom (i).

We will characterize the existence of spanning sets in terms of the existence of certain functions on Γ.

Definition 2.38. A *length function* for (Γ, F) is a map $\lambda \colon \Gamma \to \mathbb{R}_{\geq 0}$ that satisfies:

(i) (Northcott property). For all $N \in \mathbb{N}$ there are only finitely many $a \in \Gamma$ such that $\lambda(a) \leq N$.

(ii) (Ultrametric inequality). There is $D \in \mathbb{R}_{\geq 0}$ such that $\lambda(a+b) \leq \max\{ \lambda(a), \lambda(b) \} + D$ for all $a, b \in \Gamma$.

(iii) (Symmetry). $\lambda(-a) = \lambda(a)$ for all $a \in \Gamma$.

(iv) (Logarithmic property). There are $C \in \mathbb{R}_{>0}$, $E \in \mathbb{R}_{\geq 0}$, and a finite exceptional set $X \subseteq \Gamma$ such that

- $\lambda(F^n a) \geq \lambda(a) + nC - E$ for $a \in \Gamma \setminus X$, and
- $\lambda(Fa) \leq \lambda(a) + C$ for all $a \in \Gamma$.

Recall that if Σ is F-spanning we define $\lambda_\Sigma \colon \Gamma \to \mathbb{N}$ to be

$$a \mapsto \min\{ |\sigma| : \sigma \in \Sigma^*, [\sigma]_F = a \}.$$

As the terminology suggests, these are the prototypical length functions.

Proposition 2.39. *If Σ is an F-spanning set then λ_Σ is a length function with respect to the constants $C = D = E = 1$ and exceptional set $X = \Sigma$.*

Proof. We verify the axioms.

(i) This follows from the fact that there are only finitely many strings of a given length.

(ii) This is precisely Proposition 2.14 (1).

(iii) If $s_0 \cdots s_{n-1} \in \Sigma^*$ is such that $[s_0 \cdots s_{n-1}]_F = a$ then by axiom (ii) of spanning sets $(-s_0) \cdots (-s_{n-1}) \in \Sigma^*$ as well, and $[(-s_0) \cdots (-s_{n-1})]_F = -a$.

(iv) Proposition 2.19 says that if $a \notin \Sigma$ then $\lambda_\Sigma(F^n a) \geq \lambda_\Sigma(a) + n - 1$. That $\lambda_\Sigma(Fa) \leq \lambda_\Sigma(a) + 1$ is because if $\sigma \in \Sigma^*$ then $[0\sigma]_F = F[\sigma]_F$. □ Proposition 2.39

Here is a related notion that will also be of use:

Definition 2.40. A *height function*[3] for (Γ, F) is a map $h\colon \Gamma \to \mathbb{R}_{\geq 0}$ that satisfies:

(i) (Northcott property). For all $N \in \mathbb{N}$ there are only finitely many $a \in \Gamma$ such that $h(a) \leq N$.

(ii) (Weak ultrametric inequality). There are $\alpha, \kappa \in \mathbb{R}$ with $\alpha \geq 1$ and $\kappa \geq 0$ such that if $a, b \in \Gamma$ then $h(a + b) \leq \alpha \max\{ h(a), h(b) \} + \kappa$.

(iii) (Canonicity). There is $\beta \in \mathbb{R}$ with $\beta > 1$ and a finite exceptional set X such that $h(Fa) \geq \beta h(a)$ for $a \in \Gamma \setminus X$.

In practice we will use height functions (which have less stringent requirements) to deduce the existence of spanning sets, whereas we will use length functions to understand the structure of (Γ, F) that we already know have spanning sets, as we did with λ_Σ in the proof of Lemma 2.26.

The following proposition tells us that the existence of a length function is stronger than the existence of a height function. In fact we will see in Theorem 2.43 that they are both equivalent to the existence of a spanning set.

[3]Our definition of height function differs slightly from the definition appearing in [6], but the two are equivalent; see the bibliographical notes.

Proposition 2.41. *Suppose λ is a length function for (Γ, F) with associated constants C, D, E and exceptional set X. Pick $r \in \mathbb{N}$ such that $rC > E$. Then $h(a) := 2^{\lambda(a)}$ is a height function for (Γ, F^r) with associated constants $\alpha = 2^D, \beta = 2^{rC-E}, \kappa = 0$ and exceptional set X.*

Proof. The weak ultrametric inequality and Northcott property of h follow directly from the ultrametric inequality and Northcott property of λ. For canonicity, we note that if $a \in \Gamma \setminus X$ then the logarithmic property of λ yields that $h(F^r a) \geq 2^{\lambda(a)+rC-E} = 2^{rC-E}h(a)$, as desired. □ Proposition 2.41

Remark 2.42. If λ is a length function for (Γ, F) with associated constants C, D, E and exceptional set X, then λ is also a length function for (Γ, F^r) for any $r > 0$, with associated constants rC, D, E, and exceptional set X. So if we are willing to work with powers of F we can assume C is arbitrarily large, and in particular is large compared to the other constants. A similar property holds of height functions.

We now give our first characterization of the existence of spanning sets. As we remarked earlier, if $\Gamma/F\Gamma$ is infinite there is no hope for a spanning set. The setting for our characterization therefore includes the assumption that $\Gamma/F\Gamma$ is finite.

Theorem 2.43. *Suppose $\Gamma/F\Gamma$ is finite. The following are equivalent:*

1. *Γ admits an F^r-spanning set for some $r > 0$.*
2. *There is a length function for (Γ, F^r) for some $r > 0$.*
3. *There is a height function for (Γ, F^r) for some $r > 0$.*

A particular case where $\Gamma/F\Gamma$ is finite is when Γ is finitely generated. Indeed, since F is injective, Γ and $F\Gamma$ have the same rank; so $\Gamma/F\Gamma$ is torsion, and thus since Γ is finitely generated we get that $\Gamma/F\Gamma$ is finite.

Proof. That (1) $\implies$ (2) is Proposition 2.39, and that (2) $\implies$ (3) is Proposition 2.41. We check that (3) $\implies$ (1). By replacing F with F^r, it suffices to check the case where h is a height function for (Γ, F), say with associated constants α, β, κ and exceptional set X.

Claim 2.44. *We may assume $\kappa = 0$.*

Proof. We produce a height function h' for (Γ, F) with the associated $\kappa = 0$. Let $K = \min\{\, h(a) : a \in \Gamma, h(a) \neq 0 \,\}$; note that the minimum exists by the Northcott property. Define

$$h'(a) = \begin{cases} h(a) & \text{if } h(a) \neq 0 \\ K & \text{else.} \end{cases}$$

So $h'(a) \geq h(a)$ and $h'(a) \geq K$ for all $a \in \Gamma$. It is clear that h' satisfies Northcott. If $X' = X \cup h^{-1}(0)$ and $\beta' = \beta$ then h' satisfies canonicity with respect to β' and X'. Indeed, if $a \in \Gamma \setminus X'$ then since $a \notin X$ and $h(a) \neq 0$ we get that $h(Fa) \geq \beta h(a) > 0$; hence $h'(Fa) = h(Fa) \geq \beta h(a) = \beta h'(a)$.

It remains to verify the weak ultrametric inequality. Let $\alpha' = \alpha + \frac{\kappa}{K}$, and suppose $a, b \in \Gamma$. If $h(a+b) \neq 0$, then

$$\begin{aligned} h'(a+b) &= h(a+b) \\ &\leq \alpha \max\{\, h(a), h(b) \,\} + \kappa \\ &\leq \alpha \max\{\, h'(a), h'(b) \,\} + \kappa \\ &\leq \left(\alpha + \frac{\kappa}{K}\right) \max\{\, h'(a), h'(b) \,\} \\ &\leq \alpha' \max\{\, h'(a), h'(b) \,\}. \end{aligned}$$

If on the other hand $h(a+b) = 0$ then $h'(a+b) = K \leq h'(a) \leq \alpha' \max\{\, h'(a), h'(b) \,\}$ (since $\alpha' \geq \alpha \geq 1$). $\square$ Claim 2.44

Claim 2.45. *We may further assume $h(a) = h(-a)$ for all $a \in \Gamma$.*

Proof. Let $h'(a) = \max(h(a), h(-a))$. We verify that this is a height function for (Γ, F) with the same associated constants (and in particular with $\kappa = 0$) and with exceptional set $\pm X$. The Northcott property follows from the fact that if $h'(a) \leq N$ then $h(a) \leq N$. For the weak ultrametric inequality, note that if $a, b \in \Gamma$ then

$$\begin{aligned} h'(a+b) &= \max(h(a+b), h(-a-b)) \\ &\leq \max(\alpha \max(h(a), h(b)), \alpha \max(h(-a), h(-b))) \\ &= \alpha \max(\max(h(a), h(-a)), \max(h(b), h(-b))) \\ &= \alpha \max(h'(a), h'(b)). \end{aligned}$$

For canonicity, we note that if $a \in \Gamma \setminus \pm X$ then

$$h'(Fa) = \max(h(Fa), h(F(-a))) \geq \max(\beta h(a), \beta h(-a)) = \beta h'(a),$$

as desired. $\square$ Claim 2.45

Claim 2.46. *We may further assume that $\beta > \alpha^3$.*

Proof. If we pick $r > \log_\beta(\alpha^3)$ then h is also a height function for (Γ, F^r) with respect to the constants $\alpha, \beta^r, \kappa = 0$ and exceptional set X. (See Remark 2.42.) So we can replace F with F^r and assume that $\beta > \alpha^3$. □ Claim 2.46

Since $\Gamma/F\Gamma$ is finite, there is $N > 0$ such that $\Sigma := \{ a \in \Gamma : h(a) \leq N \} \supseteq X \cup \{ 0 \}$ and Σ contains a representative of each coset of $F\Gamma$. Let us check that Σ is F-spanning:

(i) We mimic the usual argument that a base-d expansion exists for any natural number. Suppose $a \in \Gamma$, and we wish to find $\sigma \in \Sigma^*$ such that $a = [\sigma]_F$. If $h(a) \leq N$ then $a \in \Sigma$, and we can take $\sigma = a$; suppose then that $h(a) > N$. Since Σ contains a representative of each coset of $F\Gamma$, there is $\ell \in \Sigma$ such that $a = \ell + Fb$ for some $b \in \Gamma$; we then recursively find $\sigma_0 \in \Sigma^*$ such that $[\sigma_0]_F = b$, at which point $\sigma := \ell\sigma_0$ is our desired string.

It remains to show that this recursion terminates. Let us consider $h(b)$. By the weak ultrametric inequality and symmetry, we find that

$$h(Fb) = h(a - \ell) \leq \alpha \max\{ h(a), h(-\ell) \} \leq \alpha \max\{ h(a), N \} = \alpha h(a)$$

(since $h(a) > N$). Now, if $b \in X$ then $b \in \Sigma$, and the recursion terminates; assume then that $b \notin X$. Then canonicity yields that $h(b) \leq \frac{1}{\beta} h(Fb) \leq \frac{\alpha}{\beta} h(a)$; so, since $\frac{\alpha}{\beta} \leq \frac{\alpha^3}{\beta} < 1$ by assumption, this recursion terminates.

(ii) That $0 \in \Sigma$ is by assumption; that Σ is closed under negation is because we chose h such that $h(a) = h(-a)$.

(iii) Suppose $a, b, c \in \Sigma$. Then $h(a), h(b), h(c) \leq N$; so by the weak ultrametric inequality

$$h(a + b + c) \leq \alpha \max\{ h(a), h(b + c) \} \leq \alpha \max\{ h(a), \alpha \max\{ h(b), h(c) \} \} \leq \alpha^2 N.$$

If $a + b + c \in \Sigma$ then we're done. If not, then by our computation for axiom (i), if we write $a + b + c = \ell + Fd$ for $\ell \in \Sigma$ and $d \in \Gamma$, then $h(d) \leq \frac{\alpha}{\beta} h(a + b + c) \leq \frac{\alpha^3}{\beta} N \leq N$ (since by assumption $\beta > \alpha^3$). So $d \in \Sigma$, and $a + b + c \in \Sigma + F\Sigma$.

(iv) Suppose $a, b \in \Sigma$ are such that $a + b = Fc$ for some $c \in \Gamma$. If $c \in X$ then $c \in \Sigma$, as desired; suppose then that $c \notin X$. Then by canonicity and the weak ultrametric inequality we get that

$$h(c) \leq \frac{h(Fc)}{\beta} = \frac{h(a + b)}{\beta} \leq \frac{\alpha \max\{ h(a), h(b) \}}{\beta} \leq \frac{\alpha}{\beta} N \leq \frac{\alpha^3}{\beta} N < N.$$

So $c \in \Sigma$, as desired. □ Theorem 2.43

Remark 2.47. The proof that Σ satisfies axiom (i) further shows that if h is a height function and Σ is the associated spanning set, then $\lambda_\Sigma \in O(\log(h))$.

As an application of this characterization, we get a much more concrete description when Γ is finitely generated:

Theorem 2.48. *Suppose Γ is finitely generated. Then there is an F^r-spanning set for some $r > 0$ if and only if every complex eigenvalue μ of $F \otimes_{\mathbb{Z}} \mathrm{id}_{\mathbb{C}}$ satisfies $|\mu| > 1$.*

Before proving this, we will need to know more about length functions. The following lemma tells us that by increasing E, we can assume the exceptional set only contains $a \in \Gamma$ such that the F-orbit $F^{\mathbb{N}}a$ of a is finite.

Lemma 2.49. *Suppose λ is a length function for (Γ, F) with associated constants C, D, E and exceptional set X. Then there is $E' \geq E$ such that λ is a length function with respect to constants C, D, E' and exceptional set $\{\, a \in X : F^{\mathbb{N}}a$ is finite $\}$.*

The idea is to observe that if a has infinite F-orbit, then since the exceptional set is finite it follows that the logarithmic property applies to $F^i a$ for i sufficiently large.

Proof. Let $X_0 = \{\, a \in X : F^{\mathbb{N}}a \text{ is finite} \,\}$. Since X is finite, and every element of $X \setminus X_0$ has infinite F-orbit, there is $N \in \mathbb{N}$ such that $F^N a \notin X$ for all $a \in X \setminus X_0$.

Suppose $a \in X \setminus X_0$. We will show that by increasing E we can make the logarithmic property apply to a. The upper bound of the logarithmic property already applies to a, since no exceptions to it are permitted; we will focus on the lower bound. If $n \geq N$ then since the logarithmic property applies to $F^N a$ we get

$$\begin{aligned} \lambda(F^n a) &= \lambda(F^{n-N}(F^N a)) \\ &\geq \lambda(F^N a) + (n-N)C - E \\ &= \lambda(a) + nC - (\underbrace{\lambda(a) - \lambda(F^N a) + NC + E}_{E_{a,N}}). \end{aligned}$$

If on the other hand $n < N$ then

$$\lambda(F^n a) = \lambda(a) + nC - (\underbrace{\lambda(a) - \lambda(F^n a) + nC}_{E_{a,n}}).$$

Hence if we pick $E_a \geq E_{a,n}$ for all $n \leq N$ then $\lambda(F^n a) \geq \lambda(a) + nC - E_a$ for all $n \in \mathbb{N}$; so the logarithmic property applies to a.

So if we pick $E' \geq E$ such that $E' \geq E_a$ for all $a \in X \setminus X_0$, then the logarithmic property applies to all such a with respect to the constants C, D, E'. Moreover since $E' \geq E$ the logarithmic property still applies to all $a \notin X$. So λ is a length function for (Γ, F) with constants C, D, E' and exceptional set X_0. □ Lemma 2.49

As a consequence, we can assume that every exceptional element is torsion:

Lemma 2.50. *Suppose there is a length function for (Γ, F). If $a \in \Gamma$ has finite F-orbit then a is torsion. In particular, if λ is a length function for (Γ, F) with associated constants C, D, E and exceptional set X, then there is $E' \geq E$ such that λ is a length function for (Γ, F) with respect to the constants C, D, E' and with exceptional set $\{ a \in X : a \text{ is torsion} \}$.*

Proof. Fix a length function λ for (Γ, F). Suppose $a \in \Gamma$ has finite F-orbit. If $na \in \mathbb{Z}a$ then $F^{\mathbb{N}}(na) = n(F^{\mathbb{N}}a)$ is also finite. But if $b \in \Gamma$ has finite F-orbit then $\lambda(F^n b)$ is bounded as n ranges; so b cannot satisfy the logarithmic property with respect to any C, D, E, and thus b is exceptional. So $\mathbb{Z}a$ is finite, and a is torsion.

The "in particular" then follows from Lemma 2.49. □ Lemma 2.50

In fact when Γ is finitely generated then all torsion elements also have finite F-orbit; this is because endomorphisms preserve the torsion subgroup, which is finite. We will not need this, however.

The consequence that is of interest to us in proving Theorem 2.48 is that if Γ is finitely generated and there is non-torsion $a \in \Gamma$ such that $F^{\mathbb{N}}a$ is finite, then there cannot be a length function for (Γ, F).

To prove Theorem 2.48, we will want to reduce to the torsion-free case. The following lemma allows us to do so:

Lemma 2.51. *Suppose Γ is finitely generated. Write $\Gamma = \Gamma_0 \times H$ for $\Gamma_0 \leq \Gamma$ torsion-free and $H \leq \Gamma$ finite. Let $\pi \colon \Gamma \to \Gamma_0$ be the projection, and let $F_0 \colon \Gamma_0 \to \Gamma_0$ be $(\pi \circ F) \restriction \Gamma_0$. Then:*

1. *F_0 is an injective endomorphism of Γ_0, and satisfies $\pi \circ F = F_0 \circ \pi$.*
2. *If $s_0 \cdots s_{n-1} \in \Gamma^*$ then $\pi([s_0 \cdots s_{n-1}]_F) = [\pi(s_0) \cdots \pi(s_{n-1})]_{F_0}$.*
3. *If Σ is an F^r-spanning set for Γ then $\pi(\Sigma)$ is an F_0^r-spanning set for Γ_0.*
4. *If Σ_0 is an F_0^r-spanning set for Γ_0 then $\pi^{-1}(\Sigma_0)$ is an F^r-spanning set for Γ.*

Proof.

1. For injectivity of F_0, suppose $a, b \in \Gamma_0$ satisfy $\pi(Fa) = \pi(Fb)$. Then $\pi(F(a-b)) = 0$, so $F(a-b) \in H$; so, since H is the torsion subgroup of Γ and F is injective, we get that $a - b \in H$. But $a - b \in \Gamma_0$; so $a = b$, and F_0 is injective.

   To see that $\pi \circ F = F_0 \circ \pi$, note for $a \in \Gamma$ that

   $$\pi(Fa) - F_0(\pi(a)) = \pi(Fa) - \pi(F(\pi(a))) = \pi(F(a - \pi(a))) \in \pi(H) = \{0\}.$$

2. Suppose $s_0, \ldots, s_{n-1} \in \Gamma$. Then part (1) yields that

   $$\begin{aligned}[(\pi(s_0)) \cdots (\pi(s_{n-1}))]_{F_0} &= \pi(s_0) + F_0(\pi(s_1))) + \cdots + F_0^{n-1}(\pi(s_{n-1}))) \\ &= \pi(s_0) + \pi(Fs_1) + \cdots + \pi(F^{n-1}s_{n-1}) \\ &= \pi([s_0 \cdots s_{n-1}]_F),\end{aligned}$$

   as desired.

3. Note by part (1) we get that $(\pi \circ F^r)\restriction\Gamma_0 = (F_0^r \circ \pi)\restriction\Gamma_0 = F_0^r$. So we can replace F with F^r and F_0 with F_0^r, and thus assume $r = 1$. Suppose $\Sigma \subseteq \Gamma$ is F-spanning; we show that $\pi(\Sigma)$ is F_0-spanning.

   (i) Suppose $a \in \Gamma_0$. Since Σ is F-spanning there is $s_0 \cdots s_{n-1} \in \Sigma^*$ such that $[s_0 \cdots s_{n-1}]_F = a$. Then $\pi(s_0), \ldots, \pi(s_{n-1}) \in \pi(\Sigma)$, and by part (2) we get that

   $$[(\pi(s_0)) \cdots (\pi(s_{n-1}))]_{F_0} = \pi([s_0 \cdots s_{n-1}]_F) = \pi(a) = a.$$

   (ii) Since Σ is F-spanning we get that $0 = \pi(0) \in \pi(\Sigma)$, and that if $\pi(a) \in \pi(\Sigma)$ then $-a \in \Sigma$, and hence $\pi(-a) \in \pi(\Sigma)$.

   (iii) Suppose $\pi(a_1), \pi(a_2), \pi(a_3) \in \pi(\Sigma)$. Since Σ is F-spanning there are $b, c \in \Sigma$ such that $a_1 + a_2 + a_3 = b + Fc$; then

   $$\pi(a_1) + \pi(a_2) + \pi(a_3) = \pi(b) + \pi(Fc) = \pi(b) + F_0(\pi(c)).$$

   (iv) Suppose $\pi(a_1), \pi(a_2) \in \pi(\Sigma)$ and $\pi(a_1) + \pi(a_2) = F_0 b$ for some $b \in \Gamma_0$. Then $\pi(a_1 + a_2) = F_0(\pi(b)) = \pi(Fb)$; so there is $h \in H$ such that $a_1 + a_2 = Fb + h$. Since F is injective and H is F-invariant and finite, we get that F is bijective on H; so there is $h_0 \in H$ such that $Fh_0 = h$. So $a_1 + a_2 = F(b + h_0)$; so since Σ is F-spanning we get that $b + h_0 \in \Sigma$. So $a = \pi(b + h_0) \in \pi(\Sigma)$.

4. As in part (3), it suffices to check the case $r = 1$. Suppose $\Sigma_0 \subseteq \Gamma_0$ is F_0-spanning; we show that $\pi^{-1}(\Sigma_0)$ is F-spanning.

   (i) Suppose $a \in \Gamma$. Since Σ_0 is F_0-spanning there is $s_0 \cdots s_{n-1} \in \Sigma_0^*$ such that $[s_0 \cdots s_{n-1}]_{F_0} = \pi(a)$. Then as in axiom (i) of the converse, we get
   $$\pi([s_0 \cdots s_{n-1}]_F) = [\pi(s_0) \cdots \pi(s_{n-1})]_{F_0} = [s_0 \cdots s_{n-1}]_{F_0} = \pi(a);$$
   so $[s_0 \cdots s_{n-1}]_F = a + h$ for some $h \in H$. Then $[(s_0 - h)s_1 \cdots s_{n-1}]_F = a$, and $s_0 - h, s_1, \ldots, s_{n-1} \in \pi^{-1}(\Sigma_0)$, as desired.

   (ii) Since $0 \in \Sigma_0$ we get that $0 \in \pi^{-1}(\Sigma_0)$. If $a \in \pi^{-1}(\Sigma_0)$ then since Σ_0 is F_0-spanning we get that $\pi(-a) \in \Sigma_0$; hence $-a \in \pi^{-1}(\Sigma_0)$ as well.

   (iii) Suppose $a_1, a_2, a_3 \in \pi^{-1}(\Sigma_0)$. Since Σ_0 is F-spanning there are $b, c \in \Sigma_0$ such that
   $$\pi(a_1) + \pi(a_2) + \pi(a_3) = b + F_0 c = \pi(b) + F_0(\pi(c)) = \pi(b) + \pi(Fc).$$
   So there is $h \in H$ such that $a_1 + a_2 + a_3 = b + Fc + h = (b + h) + Fc$; so $a_1 + a_2 + a_3 \in \pi^{-1}(\Sigma_0) + F(\pi^{-1}(\Sigma_0))$, as desired.

   (iv) Suppose $a_1, a_2 \in \pi^{-1}(\Sigma_0)$ and $a_1 + a_2 = Fb$ for some $b \in \Gamma$. Then $\pi(a_1) + \pi(a_2) = \pi(Fb) = F_0(\pi(b))$. So since Σ_0 is F_0-spanning we get that $\pi(b) \in \Sigma_0$, and thus that $b \in \pi^{-1}(\Sigma_0)$, as desired. $\square$ Lemma 2.51

Proof of Theorem 2.48. We first reduce to the case $\Gamma = (\mathbb{Z}^m, +)$ for some m. By the fundamental theorem of finitely generated abelian groups, we may assume $\Gamma = \mathbb{Z}^m \times H$ for some finite group H. As in Lemma 2.51, we let $F_0 = (\pi \circ F){\restriction}\mathbb{Z}^m$, where $\pi \colon \Gamma \to \mathbb{Z}^m$ is the projection; so by Lemma 2.51 there is an F^r-spanning set for Γ if and only if there is an F_0^r-spanning set for $\mathbb{Z}^m$. Moreover $\Gamma \otimes_{\mathbb{Z}} \mathbb{C} \cong \mathbb{C}^m$, and under this identification $F \otimes \mathrm{id}_{\mathbb{C}}$ acts as $F_0 \otimes \mathrm{id}_{\mathbb{C}}$; so $F \otimes \mathrm{id}_{\mathbb{C}}$ has an eigenvalue in the unit disk if and only if $F_0 \otimes \mathrm{id}_{\mathbb{C}}$ does.

It thus suffices to check the case $\Gamma = \mathbb{Z}^m$. So F is the restriction of the $\mathbb{C}$-linear map $F_{\mathbb{C}} := F \otimes_{\mathbb{Z}} \mathrm{id}_{\mathbb{C}}$ on $\mathbb{C}^m$; moreover $F_{\mathbb{C}}$ can be represented by a matrix over $\mathbb{Z}$.

($\Longleftarrow$) By replacing F with a power thereof, we may assume that $|\mu| \geq 3$ for all eigenvalues μ of $F_{\mathbb{C}}$. Fix a basis $\{ w_1, \ldots, w_m \}$ for $\mathbb{C}^m$ that puts $F_{\mathbb{C}}$ into Jordan canonical form, and let h be the associated infinity norm; so if
$$v = \sum_{i=1}^{m} f_i(v) w_i$$

then $h(v) = \max\{\, |f_1(v)|, \ldots, |f_m(v)| \,\}$. We show that h is a height function for $(\mathbb{Z}^m, F)$. If we take $\alpha = 2$ and $\kappa = 0$, then the triangle inequality for norms implies that h satisfies the weak ultrametric inequality. The Northcott property follows from the equivalence of norms on finite-dimensional spaces. Indeed, recall that given two norms $\|\cdot\|_1, \|\cdot\|_2$ on a finite-dimensional space there are $C_1, C_2 > 0$ such that $C_1\|v\|_1 \le \|v\|_2 \le C_2\|v\|_1$ for all vectors v. So if $\|\cdot\|_\infty$ denotes the usual infinity norm on $\mathbb{C}^m$, then there is $C > 0$ such that $\|v\|_\infty \le Ch(v)$ for all $v \in \mathbb{C}^m$. Thus if $h(v) \le N$ then $\|v\|_\infty \le CN$. The Northcott property of h follows from the fact that there are only finitely many $v \in \mathbb{Z}^m$ such that $\|v\|_\infty \le CN$.

For canonicity, let $\beta = 2$ and $X = \emptyset$. Suppose $v \in \mathbb{Z}^m$; fix k such that $h(v) = |f_k(v)|$. Then since $\{\, w_1, \ldots, w_m \,\}$ puts $F_\mathbb{C}$ into Jordan canonical form we get that either $f_k(F_\mathbb{C} v) = \mu f_k(v)$ or $f_k(F_\mathbb{C} v) = \mu f_k(v) + f_{k+1}(v)$, where μ is the eigenvalue of the Jordan block associated to w_k. In the former case, we get

$$h(F_\mathbb{C} v) \ge |f_k(F_\mathbb{C} v)| = |\mu||f_k(v)| = |\mu| h(v) \ge 2h(v).$$

In the latter case, we get by the reverse triangle inequality that

$$h(F_\mathbb{C} v) \ge |f_k(F_\mathbb{C} v)| \ge |\mu||f_k(v)| - |f_{k+1}(v)| \ge 3h(v) - h(v) = 2h(v).$$

So $h(Fv) \ge \beta h(v)$ for all $v \in \mathbb{Z}^m \setminus X$; so h satisfies canonicity, and is thus a height function. Theorem 2.43 then yields an F^r-spanning set for some $r > 0$.

($\Longrightarrow$) There are two cases. Suppose first that there is an eigenvalue μ of $F_\mathbb{C}$ with $|\mu| < 1$. In this case, our approach generalizes that of Example 2.36. Fix a basis $\{\, w_1, \ldots, w_m \,\}$ that puts $F_\mathbb{C}$ in Jordan canonical form; as before for $v \in \mathbb{C}^m$ write

$$v = \sum_{i=1}^{m} f_i(v) w_i.$$

So each $f_i \colon \mathbb{C}^m \to \mathbb{C}$ is linear. Since $\{\, w_1, \ldots, w_m \,\}$ puts $F_\mathbb{C}$ in Jordan canonical form, there is k such that $f_k(F_\mathbb{C} v) = \mu f_k(v)$ for all $v \in \mathbb{C}^m$. Suppose for contradiction that there is an F-spanning set $\Sigma \subseteq \mathbb{Z}^m$; let $M = \max\{\, |f_k(a)| : a \in \Sigma \,\}$. Then if $s_0 \cdots s_{n-1} \in \Sigma^*$ then

$$\begin{aligned} |f_k([s_0 \cdots s_{n-1}]_F)| &\le |f_k(s_0)| + \cdots + |f_k(F_\mathbb{C}^{n-1} s_{n-1})| \\ &= |f_k(s_0)| + \cdots + |\mu^{n-1} f_k(s_{n-1})| \\ &\le M + \cdots + |\mu|^{n-1} M \\ &\le \frac{M}{1 - |\mu|} \end{aligned}$$

since $|\mu| < 1$. So since Σ is F-spanning for $\mathbb{Z}^m$ we get that $|f_k(v)| < \frac{M}{1-|\mu|}$ for all $v \in \mathbb{Z}^m$. Since $\mathbb{Z}^m$ is closed under doubling and $f_k(2v) = 2f_k(v)$, this implies that $f_k(v) = 0$ for all $v \in \mathbb{Z}^m$. So $\mathbb{Z}^m \subseteq \operatorname{span}_{\mathbb{C}}(\{ w_1, \ldots, w_m \} \setminus \{ w_k \})$, contradicting the fact that $\mathbb{Z}^m$ spans $\mathbb{C}^m$. So no F-spanning set can exist. But given any $r > 0$ we get that μ^r is an eigenvalue of $F^r_{\mathbb{C}} = F^r \otimes_{\mathbb{C}} \mathrm{id}_{\mathbb{C}}$, and $|\mu^r| < 1$. So we can apply the above argument to F^r, and conclude that there is no F^r-spanning set for any $r > 0$.

For the second case, suppose there is no eigenvalue μ of $F_{\mathbb{C}}$ with $|\mu| < 1$. Then by hypothesis there is some eigenvalue μ of $F_{\mathbb{C}}$ with $|\mu| = 1$. Let $p_\mu \in \mathbb{Q}[t]$ be the minimal polynomial of μ over $\mathbb{Q}$. Suppose $\nu \in \mathbb{C}$ is another root of p_μ; so $|\nu| \geq 1$. There is an automorphism $\Phi \colon \mathbb{C} \to \mathbb{C}$ fixing $\mathbb{Q}$ that sends $\mu \mapsto \nu$. So since $\mu^{-1} = \overline{\mu}$ we get that $p_\mu(\nu^{-1}) = p_\mu(\Phi(\mu^{-1})) = \Phi(p_\mu(\overline{\mu})) = 0$; thus ν^{-1} is also a root of p_μ. Let p_F be the characteristic polynomial of $F_{\mathbb{C}}$; note that p_F has coefficients in $\mathbb{Z}$, since $F_{\mathbb{C}}$ can be represented by a matrix over $\mathbb{Z}$. So, since $p_F(\mu) = 0$, we get that $p_\mu \mid p_F$. So ν^{-1} is also an eigenvalue of $F_{\mathbb{C}}$, and hence by assumption satisfies $|\nu^{-1}| \geq 1$. So $|\nu| \leq 1$, and thus $|\nu| = 1$. So all roots of p_μ lie on the unit circle.

Now, let $V = \ker(p_\mu(F_{\mathbb{C}}))$, and let $F_V = F_{\mathbb{C}} \restriction V$; note that V is $F_{\mathbb{C}}$-invariant, so we may view F_V as a map $V \to V$. Then $p_\mu(F_V) = 0$, so the minimal polynomial of F_V divides p_μ, and thus splits into distinct linear factors over $\mathbb{C}$. So F_V is diagonalizable, and moreover all eigenvalues of F_V are on the unit circle; so if $\langle \cdot, \cdot \rangle$ is the inner product induced by any eigenbasis for F_V, then F_V is unitary with respect to $\langle \cdot, \cdot \rangle$. So if $\|\cdot\|$ is the 2-norm induced by $\langle \cdot, \cdot \rangle$ then F_V preserves $\|\cdot\|$.

Since $p_\mu(F_{\mathbb{C}})$ has a non-trivial zero in $\mathbb{C}^m$ (for instance $p_\mu(F_{\mathbb{C}})v = p_\mu(\mu)v = 0$ whenever v is an eigenvector associated to μ), we get that $\det(p_\mu(F_{\mathbb{C}})) = 0$; so, since $F_{\mathbb{C}}$ can be represented by a matrix over $\mathbb{Z}$ and $p_\mu \in \mathbb{Q}[t]$, we get that $p_\mu(F_{\mathbb{C}})$ has a non-trivial zero in $\mathbb{Q}^m$. So there is some non-zero $v \in \mathbb{Q}^m \cap V$; since V is closed under scalar multiplication, we may assume $v \in \mathbb{Z}^m$. Then since F_V is a $\|\cdot\|$-isometry, we get that $\|F^i v\| = \|F^i_V v\| = \|v\|$ for all i. As argued in the right-to-left direction above, the equivalence of norms on a finite-dimensional space implies there can only be finitely many $w \in \mathbb{Z}^m$ with $\|w\| = \|v\|$. So by injectivity of F there is i such that $v = F^i v$. Hence by Lemma 2.50 no length function for (Γ, F) can exist. Again any power of F also satisfies our hypothesis; so we can apply the same argument to F^r, and again conclude that there is no length function for (Γ, F^r) for any $r > 0$. Theorem 2.43 then yields that there is no F^r-spanning set for any $r > 0$. □ Theorem 2.48

2.6 F-sets are F-automatic

In this section we introduce *F-sets* and show that they are F-automatic. We continue to assume that Γ is an infinite abelian group, that $F\colon \Gamma \to \Gamma$ is an injective endomorphism, and that there is a spanning set for some power of F.

The class of F-sets arose in the study of the isotrivial Mordell-Lang problem: given a commutative algebraic group G over a finite field $\mathbb{F}_q$, characterize the sets of the form $X \cap \Gamma$ where $\Gamma \leq G$ is finitely generated and $X \subseteq G$ is a closed subvariety. Moosa and Scanlon showed in [18] that if G is a semiabelian variety and Γ is F-invariant, where $F\colon G \to G$ is the q-power Frobenius endomorphism, then all such intersections are F-sets. (The full problem was subsequently solved by Bell, Ghioca, and Moosa in [5].)

Definition 2.52. Suppose Γ is an abelian group and $F\colon \Gamma \to \Gamma$ is an endomorphism of Γ (not necessarily injective). For $a \in \Gamma$ we let

$$K(a; F) := \{\, a + Fa + \cdots + F^{n-1}a : n \in \mathbb{N} \,\} = [a^*]_F.$$

An *elementary F-set* is a set of the form $\alpha + K(a_1; F^{r_1}) + \cdots + K(a_n; F^{r_n})$ for some $\alpha, a_1, \ldots, a_n \in \Gamma$ and $r_1, \ldots, r_n > 0$. A *groupless F-set*[4] is a finite union of elementary F-sets. An *F-set* is a finite union of sets of the form $A + H$ where A is an elementary F-set and $H \leq \Gamma$ is F-invariant. The *F-structure* $(\Gamma, \mathcal{F})$ on Γ has domain Γ and a predicate for every F-set in every Γ^m (where F is viewed as an endomorphism of Γ^m coordinatewise).

The class of F-sets also shows up indirectly in Derksen's positive-characteristic Skolem-Mahler-Lech theorem of [11]: Derksen shows that the solutions to a linear recurrence over a field of characteristic p are what he calls a *p-normal* set. Up to finite symmetric differences, these turn out to be precisely the F-sets of $\mathbb{Z}$ where $F\colon \mathbb{Z} \to \mathbb{Z}$ is multiplication by p; this was shown in [6].

Remark 2.53. F-orbits are examples of elementary F-sets: if $a \in \Gamma$ then

$$F^{\mathbb{N}}a = a + K(Fa - a; F).$$

A converse to this is shown in [18] under some additional assumptions on Γ and $\mathbb{Z}[F]$: they show that given any $a \in \Gamma$ there is an extension $\Gamma' \geq \Gamma$ and $F'\colon \Gamma' \to \Gamma'$ extending F such that $K(a; F)$ takes the form $\alpha + (F')^{\mathbb{N}}b$ for some $\alpha, b \in \Gamma'$. It follows that any elementary F-set in Γ can be written as a translate of a finite sum of orbits in Γ'.

[4] The definitions of F-sets and groupless F-sets here differ slightly from the definitions in [18], though they turn out to be equivalent; see the bibliographical notes.

We will later make use of F-sets when studying stability of automatic sets; for now they are, for us, simply an interesting class of examples of F-automatic sets:

Theorem 2.54. *F-sets are F-automatic.*

To show this, we will need to show in particular that F-invariant subgroups of Γ are always F-automatic. This will be an instance of a more general consequence of our characterization of the existence of spanning sets in terms of height functions:

Proposition 2.55. *Suppose $H \leq \Gamma$ is F-invariant. Then there is a spanning set for $(H, (F{\restriction}H)^r)$ for some $r > 0$. Moreover if $A \subseteq H$ then A is $(F{\restriction}H)$-automatic in H if and only if A is F-automatic in Γ.*

Proof. By Theorem 2.43 there is a height function h for (Γ, F^r) for some $r > 0$. Then $h{\restriction}H$ is a height function for $(H, (F{\restriction}H)^r)$: the Northcott property, the weak ultrametric inequality, and canonicity of $h{\restriction}H$ all follow directly from the corresponding property of h. So again by Theorem 2.43 there is an $(F{\restriction}H)^{r'}$-spanning set for some $r' > 0$.

We now check the "moreover" clause. By Proposition 2.17 there is $r > 0$ such that there is both an F^r-spanning set Σ for Γ and an $(F{\restriction}H)^r$-spanning set Σ_0 for H. By Proposition 2.15 we may assume $\Sigma \supseteq \Sigma_0$.

($\implies$) Suppose $A \subseteq H$ is $(F{\restriction}H)$-automatic in H. By Corollary 2.27 we get that

$$L := \{\, \sigma \in \Sigma_0^* : [\sigma]_{(F{\restriction}H)^r} \in A \,\} = \{\, \sigma \in \Sigma_0^* : [\sigma]_{F^r} \in A \,\}$$

is regular over Σ_0. Then since regularity isn't dependent on the alphabet chosen (see Proposition 2.9) L is regular over Σ. So, since $[L]_{F^r} = [L]_{(F{\restriction}H)^r} = A$, we get by Proposition 2.33 that A is F-automatic in Γ.

($\impliedby$) Suppose $A \subseteq H$ is F-automatic in Γ; then by Corollary 2.27, and since regular languages are closed under Boolean combinations (Corollary 2.5), we get that

$$L := \{\, \sigma \in \Sigma_0^* : [\sigma]_{(F{\restriction}H)^r} \in A \,\} = \{\, \sigma \in \Sigma^* : [\sigma]_{F^r} \in A \,\} \cap \Sigma_0^*$$

is regular over Σ. Then L is also regular over Σ_0, since $L \subseteq \Sigma_0^*$; so A is $(F{\restriction}H)$-automatic in H. □ Proposition 2.55

We now complete the proof that F-sets are F-automatic.

Proof of Theorem 2.54. Recall from Corollary 2.35 that the finite union or sum of F-automatic sets is again F-automatic. So it suffices to check that singletons, F-invariant subgroups of Γ, and sets of the form $K(a; F^r)$ are F-automatic. That singletons are F-automatic follows from Proposition 2.33 and the fact that finite languages are regular. If $H \leq \Gamma$ is F-invariant then by Proposition 2.55 there is an $(F{\restriction}H)^r$-spanning set for H for some $r > 0$. Then H is $(F{\restriction}H)$-automatic in H, since Σ^* is regular over Σ for any alphabet Σ; so the "moreover" clause of Proposition 2.55 yields that H is F-automatic in Γ.

Suppose then that $a \in \Gamma$ and $r > 0$; we show that $K(a; F^r)$ is F-automatic. Replacing F with F^r (which is harmless by Corollary 2.28) we may assume $r = 1$. Fix $s > 0$ for which there is an F^s-spanning set Σ. Then

$$\begin{aligned} K(a; F) &= \{ [a^n]_F : n \in \mathbb{N} \} \\ &= \bigcup_{i=0}^{s-1} \{ [a^i a^{sn}]_F : n \in \mathbb{N} \} \\ &= \bigcup_{i=0}^{s-1} \{ [a^i]_F + F^i [a^{sn}]_F : n \in \mathbb{N} \} \\ &= \bigcup_{i=0}^{s-1} ([a^i]_F + K(F^i [a^s]_F; F^s)). \end{aligned}$$

In particular by closure under unions and translations (Corollary 2.35) it suffices to show that $K(b; F^s)$ is F-automatic for all $b \in \Gamma$. By Proposition 2.15 we may assume Σ contains b. But then $K(b; F^s) = [b^*]_{F^s}$ and $b^* \subseteq \Sigma^*$ is regular. So Proposition 2.33 yields that $K(b; F^s)$ is F-automatic, as desired. □ Theorem 2.54

2.7 Bibliographical notes

Section 2.1 consists entirely of general knowledge from automata theory. The definition of d-automatic sets is standard. The generalizations of d-automaticity to $\mathbb{Z}$ and $\mathbb{Z}^m$ is from [1], and was inspired by [20]. In fact in [1] the authors define a notion of d-automaticity in an arbitrary finitely generated abelian group, not just $\mathbb{Z}^m$; but this is outside of our context, since the mapping $a \mapsto da$ isn't typically injective when there's torsion.

Spanning sets and automaticity were introduced in [6] in the context of finitely generated abelian groups, and in [5] it was observed that they can be used in the more general context of abelian groups. My spanning sets generalize theirs slightly, however: whereas my axiom

(iii) only requires that sums of three elements of Σ land in $\Sigma + F\Sigma$, they demand that this hold for sums of five elements; likewise whereas my axiom (iv) is only about binary sums, they demand that it hold for ternary sums as well. This relaxation requires a different proof of the fact that the graph of addition, and more generally any F-invariant subgroup, is F-automatic. Most of the results in Section 2.3 appear in [6]; the exceptions are the fact that $\lambda_\Sigma(F^r a) \geq \lambda_\Sigma(a) + r - 1$ (Proposition 2.19) and the example of a (Γ, F) where F^2 admits a spanning set but F does not (Example 2.16).

The characterization of automaticity in terms of spanning set kernels (Proposition 2.25) appears in [6], as does a restricted version of the fact that F-automaticity is independent of the spanning set chosen: they show that $A \subseteq \Gamma$ is (F, Σ)-automatic if and only if it is (F, Σ')-automatic, provided Σ and Σ' are both F-spanning (or more generally both F^r-spanning for the same r). The relaxation (Corollary 2.27) of the kernel characterization to the case where S merely contains a representative of each coset of $F\Gamma$, rather than being a spanning set, is original. The full independence of (F^r, Σ)-automaticity from both r and Σ (also Corollary 2.27) is original, though it can be deduced from the results in [6] without going through the more general kernel characterization. Closure of F-automatic sets under finite unions appears in [6], and closure under complements follows immediately from their work. The fact (Proposition 2.32) that the diagonal in Γ^2, the graph of addition, and the graph of F are F-automatic appears in [6], but my more general kernel characterization of automaticity allowed me to simplify the proof. The fact that $[L]_{F^r}$ is F-automatic whenever L is regular (Proposition 2.33) appears in [6], as does the fact (Corollary 2.30) that F-automaticity agrees with classical d-automaticity when F is multiplication by d. The observation that F-automatic sets are closed under projection is new, but it is a direct adaptation of the proof of the analogous fact for d-automatic sets, as appearing for example in [7].

Everything in Section 2.5 is original except the definition of height functions and the fact that the existence of a height function implies the existence of a spanning set (part of Theorem 2.43); these appeared in [6]. The notion of height function described here differs cosmetically from the one described in [6]: we include the weak ultrametric inequality $h(a+b) \leq \alpha \max(h(a), h(b)) + \kappa$, whereas they use the triangle inequality $h(a+b) \leq \alpha(h(a) + h(b)) + \kappa$. However, the former directly implies the latter, and if h satisfies the latter with respect to α, κ then it satisfies the former with respect to $2\alpha, \kappa$. So the two definitions capture the same class of functions.

I am indebted to Luke Anthony Franceschini for helping me work through the details of Example 2.36.

My presentation of F-sets differs slightly from how they were introduced in [18]: they first define F-*cycles*, which are sets of the form $C(a;r) = \{\, a + F^r a + \cdots + F^{(n-1)r} a : n > 0 \,\}$; so $K(a;F^r) = C(a;r) \cup \{\, 0 \,\}$. They then define groupless F-sets to be finite unions of sets of the form

$$\alpha + C(a_1;r_1) + \cdots + C(a_n;r_n)$$

and F-sets to be finite unions of sets of the form

$$\alpha + C(a_1;r_1) + \cdots + C(a_n;r_n) + H$$

where $H \le \Gamma$ is F-invariant. (The "elementary F-set" terminology was introduced here for convenience; it doesn't occur in [18].) In fact the two notions of groupless F-sets coincide, and likewise with the two notions of F-sets: this follows from the fact that

$$\begin{aligned} C(a;r) &= a + K(F^r a; F^r) \\ K(a;F^r) &= C(a;r) \cup \{\, 0 \,\}. \end{aligned}$$

Our context for F-sets is more general than the one in which they were introduced in [18]: they require that $\mathbb{Z}[F] \subseteq \operatorname{End}_{\mathbb{Z}}(\Gamma)$ be a finite extension of $\mathbb{Z}$ in which F is not a zero divisor and

$$\bigcap_{i=0}^{\infty} (F^i) = 0$$

(where (F^i) is the ideal generated by F^i in $\mathbb{Z}[F]$), and moreover that Γ be a finitely generated $\mathbb{Z}[F]$-module. These assumptions weren't necessary in this chapter, but they will come up in later chapters, in which we will want to make use of the results of [18].

The fact that F-sets are F-automatic (Theorem 2.54) was shown in [6] when Γ is finitely generated; however, since we don't require that, and due to our relaxation of the spanning set axioms, their proof that F-invariant subgroups of Γ are F-automatic falls through in this context. The proof that appears here instead relies on our characterization of spanning sets in terms of height functions via Proposition 2.55, which is original.

Chapter 3

F-sparsity

In this chapter, we introduce a notion of sparsity for F-automatic sets. We develop the basic properties of F-sparse sets in Section 3.1, and we show in Theorem 3.19 that we can characterize F-sparsity among the F-automatic sets using the asymptotics of the length functions we defined in Chapter 2 (see Definition 2.38).

We continue to assume that Γ is an infinite abelian group, that $F\colon \Gamma \to \Gamma$ is an injective endomorphism, and that there is an F^r-spanning set for some $r > 0$.

3.1 Sparse languages and F-sparse sets

Recall the notion of sparsity in the context of regular languages:

Definition 3.1. If Σ is a finite alphabet, we say $L \subseteq \Sigma^*$ is *sparse* if it is regular and $|\{\, \sigma \in L : |\sigma| \leq x \,\}|$ grows polynomially in x; i.e., there is $f \in \mathbb{R}[x]$ and $C, M \geq 0$ such that if $x \geq M$ then $|\{\, \sigma \in L : |\sigma| \leq x \,\}| \leq Cf(x)$.

Remark 3.2. If $L_1, L_2 \subseteq \Sigma^*$ are sparse and $L' \subseteq \Sigma^*$ is regular, then $L_1 \cup L_2$, L_1L_2, and $L_1 \cap L'$ are sparse. Indeed, that these are regular follows from the definition of regularity and the closure of regular languages under Boolean combinations (see Corollary 2.5); for sparsity, we then note that

$$\begin{aligned}
|\{\, \sigma \in L_1 \cup L_2 : |\sigma| \leq x \,\}| &\leq |\{\, \sigma \in L_1 : |\sigma| \leq x \,\}| + |\{\, \sigma \in L_2 : |\sigma| \leq x \,\}| \\
|\{\, \sigma \in L_1 L_2 : |\sigma| \leq x \,\}| &\leq |\{\, \sigma_1 \in L_1 : |\sigma_1| \leq x \,\}| \cdot |\{\, \sigma_1 \in L_2 : |\sigma_2| \leq x \,\}| \\
|\{\, \sigma \in L_1 \cap L' : |\sigma| \leq x \,\}| &\leq |\{\, \sigma \in L_1 : |\sigma| \leq x \,\}|
\end{aligned}$$

all grow polynomially in x.

The sparse languages are very well understood: the following fact follows from the results of [22, 16].

Fact 3.3. *Suppose Σ is a finite alphabet and $L \subseteq \Sigma^*$. The following are equivalent:*

1. *L is sparse.*
2. *L is a finite union of sets of the form*

$$u_0 v_1^* u_1 v_2^* \cdots u_{n-1} v_n^* u_n = \{\, u_0 v_1^{k_1} u_1 v_2^{k_2} \cdots u_{n-1} v_n^{k_n} u_n : k_1, \ldots, k_n \in \mathbb{N} \,\}$$

   for some $u_0, \ldots, u_n, v_1, \ldots, v_n \in \Sigma^*$. We call sets of the form $u_0 v_1^* u_1 v_2^* \cdots u_{n-1} v_n^* u_n$ simple sparse.
3. *There is a DFA $(\Sigma, Q, q_0, \Omega, \delta)$ recognizing L such that if $q \in Q$ satisfies:*
   - *q is reachable from q_0; that is, there is $\nu \in \Sigma^*$ such that $\delta(q_0, \nu) = q$.*
   - *q is not a* dead state*; that is, there is $\nu \in \Sigma^*$ such that $\delta(q, \nu) \in \Omega$.*

   then for any $n \in \mathbb{N}$ there is at most one $\sigma \in \Sigma^{(n)}$ satisfying $\delta(q, \sigma) = q$.

Informally, the last condition states that there is a DFA such that no state has a "double loop" back to itself.

We study the natural adaptation of sparsity to the setting of F-automatic sets:

Definition 3.4. If Σ is an F^r-spanning set, we say $A \subseteq \Gamma$ is (F^r, Σ)-*sparse* if there is some sparse $L \subseteq \Sigma^*$ such that $A = [L]_{F^r}$. We say A is F-*sparse* if it is (F^r, Σ)-sparse for some $r > 0$ and F^r-spanning set Σ.

It will follow from work in this chapter (see Corollary 3.20) that F-sparsity does not depend on the choice of r and Σ. Note that F-sparse sets are F-automatic as sparse languages are regular by definition (and using Proposition 2.33).

Example 3.5. Some examples of F-sparse sets:

1. Finite subsets of Γ are F-sparse as finite languages are sparse.

2. Fix $a \in \Gamma$; we show that the F-orbit $F^{\mathbb{N}}a$ of a is F-sparse. Fix an F^r-spanning set Σ. By Proposition 2.15 we may assume $a, Fa, \ldots, F^{r-1}a \in \Sigma$. Then

$$\begin{aligned} F^{\mathbb{N}}a &= \{\, [0^n a]_F : n \in \mathbb{N} \,\} \\ &= \{\, [(0^r)^i 0^j a]_F : i \in \mathbb{N}, j < r \,\} \\ &= \{\, [0^i (F^j a)]_{F^r} : i \in \mathbb{N}, j < r \,\} \\ &= \left[\bigcup_{j=0}^{r-1} 0^*(F^j a) \right]_{F^r}, \end{aligned}$$

and $\bigcup_{j=0}^{r-1} 0^*(F^j a) \subseteq \Sigma^*$ is sparse by our characterization of sparse languages (Fact 3.3). So $F^{\mathbb{N}}a$ is (F^r, Σ)-sparse, and hence F-sparse.

3. Consider the case $\Gamma = (\mathbb{Z}, +)$ with $F\colon \mathbb{Z} \to \mathbb{Z}$ multiplication by some $d \geq 2$; we show that the ordering $A = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} : a, b \in d^{\mathbb{N}}, a < b \right\}$ on $d^{\mathbb{N}}$ is F-sparse in Γ^2. Indeed, recall that $\Sigma = \{\, -d+1, -d+2, \ldots, d-1 \,\}$ is F-spanning for Γ (see Example 2.13 (1)), and hence Σ^2 is F-spanning for Γ^2 (see Example 2.13 (2)). Moreover,

$$\begin{aligned} A &= \left\{ \begin{pmatrix} d^i \\ d^j \end{pmatrix} : i, j \in \mathbb{N}, i < j \right\} \\ &= \left\{ \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}^i \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}^{j-i-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]_F : i, j \in \mathbb{N}, i < j \right\} \\ &= \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}^* \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}^* \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]_F, \end{aligned}$$

and hence A is F-sparse.

We can verify some closure properties immediately from the definition:

Proposition 3.6.

1. *If $A \subseteq \Gamma$ and $r > 0$ then A is F-sparse if and only if it is F^r-sparse.*
2. *If $A, B \subseteq \Gamma$ are F-sparse then so is $A \cup B$.*
3. *If $A \subseteq \Gamma$ is F-sparse and $X \subseteq \Gamma$ is F-automatic then $A \cap X$ is F-sparse.*

Proof.

1. That F^r-sparsity implies F-sparsity is by definition. For the converse, we show for an F^s-spanning set Σ that (F^s, Σ)-sparsity implies $(F^{rs}, [\Sigma^{(r)}]_{F^s})$-sparsity, which suffices. (Recall from Proposition 2.17 that $[\Sigma^{(r)}]_{F^s}$ is indeed F^{rs}-spanning.) Replacing F with F^s, we may assume $s = 1$.

   Suppose then that A is (F, Σ)-sparse; so $A = [L]_F$ for some sparse $L \subseteq \Sigma^*$. Let $\Sigma' = [\Sigma^{(r)}]_F$; note that strings over Σ' correspond to strings over Σ with length in $r\mathbb{N}$. Our strategy will then be to show that we can force all elements of L to have length r, and then use this correspondence to produce a sparse $L' \subseteq (\Sigma')^*$ such that $A = [L']_{F^r}$.

   We have $A = [L]_F = [L0^* \cap \Sigma^{(r\mathbb{N})}]_F$, and since L is sparse so too is $L0^* \cap \Sigma^{(r\mathbb{N})}$; this follows from the closure properties of sparse languages noted in Remark 3.2. Given $\sigma \in \Sigma^{(r\mathbb{N})}$, say $\sigma = \sigma_0 \cdots \sigma_{n-1}$ with $\sigma_0, \ldots, \sigma_{n-1} \in \Sigma^{(r)}$, we let $\sigma' = [\sigma_0]_F \cdots [\sigma_{n-1}]_F \in (\Sigma')^*$. Then $|\sigma'| = r^{-1}|\sigma|$, and

   $$[\sigma']_{F^r} = [\sigma_0]_F + F^r[\sigma_1]_F + \cdots + F^{r(n-1)}[\sigma_{n-1}]_F = [\sigma_0\sigma_1\cdots\sigma_{n-1}]_F = [\sigma]_F.$$

   So if $L' = \{\, \sigma' : \sigma \in L \,\} \subseteq (\Sigma')^*$, then $[L']_{F^r} = A$; it then suffices to show that L' is sparse. Fix a DFA $M = (\Sigma, Q, q_0, \Omega, \delta)$ recognizing L, and define an NFA $M' = (\Sigma', Q, q_0, \Omega, \delta')$ where

   $$\delta'(q, a) = \{\, \delta(q, \sigma) : \sigma \in \Sigma^{(r)}, [\sigma]_F = a \,\}$$

   for $a \in \Sigma$. Then given $\tau \in (\Sigma')^*$ of length n we get that M' accepts τ if and only if there are $\sigma_0, \ldots, \sigma_{n-1} \in \Sigma^{(r)}$ such that $\tau = [\sigma_0]_F \cdots [\sigma_{n-1}]_F$ and $\delta(q_0, \sigma_0 \cdots \sigma_{n-1}) \in \Omega$; i.e., if and only if $\tau \in L'$. So L' is regular. For sparsity of L', we note that

   $$|\{\, \sigma' \in L' : |\sigma'| \leq x \,\}| = |\{\, \sigma \in L : |\sigma| \leq rx \,\}|.$$

   So $L' \subseteq (\Sigma')^*$ is sparse, and hence $A = [L']_{F^r}$ is (F^r, Σ')-sparse.

2. Suppose $A, B \subseteq \Gamma$ are F-sparse. We argued above that (F, Σ)-sparsity implies $(F^r, [\Sigma^{(r)}]_F)$-sparsity for any $r > 0$. Applying this to A and B, we find that there is $r > 0$ and F^r-spanning sets Σ and Σ' such that A is (F^r, Σ)-sparse and B is (F^r, Σ')-sparse; say $A = [L]_{F^r}$ and $B = [L']_{F^r}$ for $L \subseteq \Sigma^*$ and $L' \subseteq (\Sigma')^*$ sparse. By Proposition 2.15 there is an F^r-spanning set Θ containing $\Sigma \cup \Sigma'$. Then $L \cup L' \subseteq \Theta^*$ is sparse (see Remark 3.2), and $A \cup B = [L \cup L']_{F^r}$; so $A \cup B$ is (F^r, Θ)-sparse, and is thus F-sparse.

3. Suppose $A \subseteq \Gamma$ is F-sparse and $X \subseteq \Gamma$ is F-automatic; say there is an F^r-spanning set Σ and a sparse $L \subseteq \Sigma^*$ such that $A = [L]_{F^r}$. Since F-automaticity implies (F^r, Σ)-automaticity, we get that $L' := \{ \sigma \in \Sigma^* : [\sigma]_{F^r} \in X \}$ is regular. Hence by our closure properties of sparse languages (Remark 3.2) we get that $L \cap L' \subseteq \Sigma^*$ is sparse. So, since $[L \cap L']_{F^r} = A \cap X$, we get that $A \cap X$ is F-sparse. □ Proposition 3.6

A consequence of this is that the F-sparse sets are closed under relative complement: if $A, B \subseteq \Gamma$ are F-sparse, then $\Gamma \backslash B$ is F-automatic by closure of F-automaticity under Boolean combinations Corollary 2.29, and hence by the previous result we get that $A \backslash B = A \cap (\Gamma \backslash B)$ is F-sparse.

Because of our characterization of sparse languages (Fact 3.3), we can write an F-sparse set as a finite union of sets of the form $[u_0 v_1^* u_1 \cdots v_n^* u_n]_{F^r}$. The following lemma and corollary tell us that we can dispense with the u_i, at the cost of introducing a translation.

Lemma 3.7. *Suppose $r > 0$ and $u_0, \ldots, u_n, v_1, \ldots, v_n \in \Gamma^*$; suppose $v_1, \ldots, v_n$ all have the same length N. Then there is $\alpha, a_1, \ldots, a_n \in \Gamma$ such that for $k_1, \ldots, k_n \in \mathbb{N}$ we have*

$$[u_0 v_1^{k_1} u_1 \cdots v_n^{k_n} u_n]_{F^r} = \alpha + [a_1^{k_1} a_2^{k_2} \cdots a_n^{k_n}]_{F^{rN}}.$$

Proof. Replacing F with F^r, we may assume $r = 1$. We apply induction on n; the base case $n = 0$ is vacuous. For the induction step, use the induction hypothesis to find $\beta, b_2, \ldots, b_n \in \Gamma$ such that for $k_2, \ldots, k_n \in \mathbb{N}$ we have

$$[u_1 v_2^{k_2} u_2 \cdots v_n^{k_n} u_n]_F = \beta + [b_2^{k_2} \cdots b_n^{k_n}]_{F^N}.$$

If we let $a_i = F^{|u_0|} b_i$ for $2 \leq i \leq n$, then

$$\begin{aligned}
[u_0 v_1^{k_1} u_1 \cdots v_n^{k_n} u_n]_F &= [u_0 v_1^{k_1}]_F + F^{|u_0| + k_1 N} [u_1 v_2^{k_2} u_2 \cdots v_n^{k_n} u_n]_F \\
&= [u_0]_F + F^{|u_0|} [v_1^{k_1}]_F + F^{|u_0| + k_1 N} \beta + F^{|u_0| + k_1 N} [b_2^{k_2} \cdots b_n^{k_n}]_{F^N} \\
&= [u_0]_F + F^{|u_0|} [v_1^{k_1}]_F + F^{|u_0| + k_1 N} \beta + F^{k_1 N} [a_2^{k_2} \cdots a_n^{k_n}]_{F^N}.
\end{aligned}$$

We wish to write $F^{|u_0|} [v_1^{k_1}]_F + F^{|u_0| + k_1 N} \beta$ as a translate of something of the form $[a_1^{k_1}]_{F^N}$.

We can do so via a telescoping sum: if we let $a_1 = F^{|u_0|}([v_1]_F - \beta + F^N\beta)$, then

$$
\begin{aligned}
& F^{|u_0|}[v_1^{k_1}]_F + F^{|u_0|+k_1N}\beta \\
= {} & F^{|u_0|}[v_1^{k_1}]_F + F^{|u_0|}\Big(\beta + (-\beta + F^N\beta) + (-F^N\beta + F^{2N}\beta) + \cdots + (-F^{(k_1-1)N}\beta + F^{k_1N}\beta)\Big) \\
= {} & F^{|u_0|}\beta \\
& + F^{|u_0|}([v_1]_F - \beta + F^N\beta) \\
& + F^{|u_0|+N}([v_1]_F - \beta + F^N\beta) \\
& + \cdots \\
& + F^{|u_0|+(k_1-1)N}([v_1]_F - \beta + F^N\beta) \\
= {} & F^{|u_0|}\beta + [a_1^{k_1}]_{F^N}.
\end{aligned}
$$

Hence if we let $\alpha = [u_0]_F + F^{|u_0|}\beta$, then substituting this into the above yields

$$
\begin{aligned}
[u_0v_1^{k_1}u_1\cdots v_n^{k_n}u_n]_F &= [u_0]_F + F^{|u_0|}\beta + [a_1^{k_1}]_{F^N} + F^{k_1N}[a_2^{k_2}\cdots a_n^{k_n}]_{F^N} \\
&= \alpha + [a_1^{k_1}a_2^{k_2}\cdots a_n^{k_n}]_{F^N},
\end{aligned}
$$

as desired. $\square$ Lemma 3.7

Corollary 3.8. *Suppose $A \subseteq \Gamma$ is F-sparse. Then there is $s_0 > 0$ such that if $s_0 \mid s$ then A is a finite union of sets of the form $\alpha + [a_1^*\cdots a_n^*]_{F^s}$ where $\alpha, a_1, \ldots, a_n \in \Gamma$.*

Proof. Replacing F with some power thereof, it suffices to check the case where A is (F,Σ)-sparse for some F-spanning set Σ; say $A = [L]_F$ for some sparse $L \subseteq \Sigma^*$. By our characterization Fact 3.3 of sparse languages, we get that L is a finite union of simple sparse languages $u_0v_1^*u_1\cdots v_n^*u_n$. Let N be the least common multiple of all the $|v_i|$ across the union, and let $s_0 = N$; suppose $s_0 \mid s$. If we let $N_i = \frac{s}{|v_i|}$ then

$$
u_0v_1^*u_1\cdots v_n^*u_n = \bigcup_{j_1=0}^{N_1-1}\cdots\bigcup_{j_n=0}^{N_n-1} u_0(v_1^{N_1})^*v_1^{j_1}u_1\cdots(v_n^{N_n})^*v_n^{j_n}u_n,
$$

and each $|v_i^{N_i}| = s$. So L is a finite union of sets of the form $u_0v_1^*u_1\cdots v_n^*u_n$ with $|v_1| = \cdots = |v_n| = s$. Given such u_i, v_i, Lemma 3.7 yields $\alpha, a_1, \ldots, a_n \in \Gamma$ such that $[u_0v_1^{k_1}u_1\cdots v_n^{k_n}u_n]_F = \alpha + [a_1^{k_1}\cdots a_n^{k_n}]_{F^s}$ for all $k_1, \ldots, k_n \in \mathbb{N}$. So $A = [L]_F$ is a finite union of sets of the form $\alpha + [a_1^*\cdots a_n^*]_{F^s}$, as desired. $\square$ Corollary 3.8

Remark 3.9. The converse holds when there is an F^s-spanning set Σ[1] (which by Proposition 2.17 can be assumed by replacing s with a multiple thereof). Indeed, by closure

[1] In fact the converse holds more generally, but we will not need this.

of F-sparsity under finite unions (Proposition 3.6) it suffices to show that sets of the form $\alpha + [a_1^* \cdots a_n^*]_{F^s}$ are F-sparse. By Proposition 2.15 we may assume $\alpha, a_1, \ldots, a_n, \alpha + a_1, \ldots, \alpha + a_n \in \Sigma$. Then

$$L := \{ \alpha \} \cup \bigcup_{i=1}^{n} (\alpha + a_i) a_i^* a_{i+1}^* \cdots a_n^* \subseteq \Sigma^*$$

is sparse by Fact 3.3, and hence $\alpha + [a_1^* \cdots a_n^*]_{F^s} = [L]_{F^s}$ is F-sparse.

In particular, we get that the F-sparse sets are closed under translation. Indeed, by Corollary 3.8 any F-sparse set A can be written as a finite union of sets of the form $\alpha + [a_1^* \cdots a_n^*]_{F^s}$ for some s for which there is an F^s-spanning set. But then for $\gamma \in \Gamma$ we get that $A + \gamma$ is a finite union of sets of the form $\gamma + \alpha + [a_1^* \cdots a_n^*]_{F^s}$, which is then F-sparse by the above.

We can now show that the F-sparse sets are closed under sums:

Corollary 3.10. *If $A, B \subseteq \Gamma$ are F-sparse then so is $A + B$.*

Proof. By Corollary 3.8 there is $s > 0$ such that A can be written as a finite union of sets of the form $\alpha + [a_1^* \cdots a_n^*]_{F^s}$ for some $\alpha, a_1, \ldots, a_n \in \Gamma$, and likewise B is a finite union of $\beta + [b_1^* \cdots b_{n'}^*]_{F^s}$. Moreover Proposition 2.17 tells us that by replacing s with a multiple thereof we may assume there is an F^s-spanning set Σ. Since F-sparsity is closed under finite unions, it suffices to show that a set of the form $\alpha + \beta + [a_1^* \cdots a_n^*]_{F^s} + [b_1^* \cdots b_{n'}^*]_{F^s}$ is F-sparse. We noted in Remark 3.9 that translates of F-sparse sets are F-sparse; so it suffices to show that

$$[a_1^* \cdots a_n^*]_{F^s} + [b_1^* \cdots b_{n'}^*]_{F^s}$$

is F-sparse.

Note by Proposition 2.15 that we can assume $a_i, b_j, a_i + b_j \in \Sigma$ for all i, j. Now, for $s_0 \cdots s_{k-1}, t_0 \cdots t_{\ell-1} \in \Sigma^*$, say with $k \leq \ell$, we let

$$s_0 \cdots s_{k-1} \oplus t_0 \cdots t_{\ell-1} = (s_0 + t_0)(s_1 + t_1) \cdots (s_{k-1} + t_{k-1}) t_k t_{k+1} \cdots t_{\ell-1}$$

be their characterwise sum; note then that

$$\begin{aligned} [s_0 \cdots s_{k-1} \oplus t_0 \cdots t_{\ell-1}]_{F^s} &= (s_0 + t_0) + \cdots + F^{k-1}(s_{k-1} + t_{k-1}) + F^k t_k + \cdots + F^{\ell-1} t_{\ell-1} \\ &= [s_0 \cdots s_{k-1}]_{F^s} + [t_0 \cdots t_{\ell-1}]_{F^s}. \end{aligned}$$

Moreover if $\sigma \in a_1^* \cdots a_n^*$ and $\tau \in b_1^* \cdots b_{n'}^*$ then $\sigma \oplus \tau \in \Sigma^*$. So $L := a_1^* \cdots a_n^* \oplus b_1^* \cdots b_{n'}^* \subseteq \Sigma^*$ and $[L]_{F^s} = [a_1^* \cdots a_n^*]_{F^s} + [b_1^* \cdots b_{n'}^*]_{F^s}$. It then suffices to show that L is sparse; we do so by induction on (n, n'). The base case $n = n' = 0$ is simply that $\{\varepsilon\}$ is sparse. For the induction step, note given $\sigma = a_1^{k_1} \cdots a_n^{k_n}$ and $\tau = b_1^{\ell_1} \cdots b_{n'}^{\ell_{n'}}$ that either $k_1 \leq \ell_1$, in which case

$$\sigma \oplus \tau = (a_1 + b_1)^{k_1} (a_2^{k_2} \cdots a_n^{k_n} \oplus b_1^{\ell_1 - k_1} b_2^{\ell_2} \cdots b_{n'}^{\ell_{n'}}),$$

or $k_1 \geq \ell_1$, in which case

$$\sigma \oplus \tau = (a_1 + b_1)^{\ell_1} (a_1^{k_1 - \ell_1} a_2^{k_2} \cdots a_n^{k_n} \oplus b_2^{\ell_2} \cdots b_{n'}^{\ell_{n'}}).$$

We thus see that

$$L = (a_1 + b_1)^* (a_2^* \cdots a_n^* \oplus b_1^* b_2^* \cdots b_{n'}^*) \cup (a_1 + b_1)^* (a_1^* a_2^* \cdots a_n^* \oplus b_2^* \cdots b_{n'}^*),$$

which is sparse by the induction hypothesis (and Remark 3.2, which tells us that the concatenation and union of sparse languages is sparse). □ Corollary 3.10

Corollary 3.11. *Groupless F-sets are F-sparse.*

Recall from Definition 2.52 that a *groupless F-set* is a finite union of translates of finite sums of sets of the form $K(a; F^r) = [a^*]_{F^r}$.

Proof. Note that if $s > 0$ is such that there is an F^s-spanning set Σ, then $K(b; F^s)$ is F-sparse: indeed, by Proposition 2.15 we may assume $b \in \Sigma$, and $K(b; F^s) = [b^*]_{F^s}$. Moreover, we showed in the proof of Theorem 2.54 that given any $a \in \Gamma$ and $r > 0$ we can write $K(a; F^r)$ as a finite union of translates of $K(b; F^s)$ for some $s > 0$ for which there is an F^s-spanning set. So $K(a; F^r)$ is a finite union of translates of F-sparse sets, and hence by Proposition 3.6 and Corollary 3.10 we get that $K(a; F^r)$ is F-sparse. But a groupless F-set is a finite union of translates of finite sums of $K(a; F^r)$; so again by Proposition 3.6 and Corollary 3.10 we get that any groupless F-set is F-sparse. □ Corollary 3.11

3.2 F-sparsity via length functions

The definition of F-sparsity raises the question of independence from the chosen spanning set: that is, given an F^r-spanning set Σ and an $F^{r'}$-spanning set Σ', if A is (F^r, Σ)-sparse, must it also be $(F^{r'}, \Sigma')$-sparse? More concretely, one might also wonder whether there even exist F-automatic sets that aren't F-sparse. In this section, we will characterize F-sparsity in terms of length functions, and use this characterization to answer both questions in the affirmative.

Lemma 3.12. *Suppose Σ is F-spanning. Suppose λ is a length function for (Γ, F) with associated constants C, D, E such that $C \geq D$. Then $\lambda \in O(\lambda_\Sigma)$.*

Proof. Let $M = \max\{\, \lambda(a) : a \in \Sigma \,\}$. For $\sigma \in \Sigma^*$ non-empty, we show by induction on $|\sigma|$ that $\lambda([\sigma]_F) \leq M + (|\sigma| - 1)C + D$. The base case $|\sigma| = 1$ is just the definition of M. For the induction step, suppose $|\sigma| > 1$, and write $\sigma = \sigma_0 a$ for some non-empty $\sigma_0 \in \Sigma^*$ and $a \in \Sigma$. By the induction hypothesis we get that $\lambda([\sigma_0]_F) \leq M + (|\sigma_0| - 1)C + D$, and by definition of M we get that $\lambda(a) \leq M$. So by the ultrametric inequality and the logarithmic property we get that

$$\begin{aligned}
\lambda([\sigma]_F) &= \lambda([\sigma_0]_F + F^{|\sigma_0|}a) \\
&\leq \max(\lambda([\sigma_0]_F), \lambda(F^{|\sigma_0|}a)) + D \\
&\leq \max(M + (|\sigma_0| - 1)C + D, M + |\sigma_0|C) + D \\
&= M + |\sigma_0|C + D \\
&= M + (|\sigma| - 1)C + D
\end{aligned}$$

(since $C \geq D$ by assumption).

Hence if $0 \neq a \in \Gamma$, say $a = [\sigma]_F$ with $\varepsilon \neq \sigma \in \Sigma^*$ of length $\lambda_\Sigma(a)$, then

$$\lambda(a) \leq M + (|\sigma| - 1)C + D \leq M + (\lambda_\Sigma(a) - 1)C + D = C\lambda_\Sigma(a) + M - C + D.$$

So $\lambda \in O(\lambda_\Sigma)$, as desired. □ Lemma 3.12

Corollary 3.13. *If λ is a length function for (Γ, F^r) and λ' is a length function for $(\Gamma, F^{r'})$, then $\lambda \in \Theta(\lambda')$.*

(Recall that $f \in \Theta(g)$ if $f \in O(g)$ and $g \in O(f)$.)

Proof. By symmetry, it suffices to show that $\lambda \in O(\lambda')$. Recall from Proposition 2.41 that $h(a) := 2^{\lambda'(a)}$ is a height function for $(\Gamma, F^{r'})$; recall further from Theorem 2.43 and Remark 2.47 that from h we can derive an F^s-spanning set Σ for some s such that $\lambda_\Sigma \in O(\log \circ h) = O(\lambda')$. It then suffices to show that $\lambda \in O(\lambda_\Sigma)$.

Let the constants associated to λ be C, D, E. By possibly replacing r with a multiple thereof, we may assume that $C \geq D$ (see Remark 2.42) and that $s \mid r$. Recall from Proposition 2.17 and Remark 2.18 that $\Sigma' := [\Sigma^{\frac{r}{s}}]_{F^s}$ is an F^r-spanning set, and $\lambda_{\Sigma'}(a) \leq \left\lceil \lambda_\Sigma(a) \cdot \frac{s}{r} \right\rceil$. So $\lambda_{\Sigma'} \in O(\lambda_\Sigma)$, and it suffices to prove that $\lambda \in O(\lambda_{\Sigma'})$. But this is precisely the previous lemma. □ Corollary 3.13

In regular languages, the defining condition of sparsity is that $|\{\,\sigma \in L : |\sigma| \leq x\,\}|$ grows polynomially in x. We would like to adapt this to the context of F-automatic sets using length functions in place of $|\sigma|$; the above corollary tells us that it won't matter which length function we choose.

Definition 3.14. We say $A \subseteq \Gamma$ is F*-meagre* if $|\{\,a \in A : \lambda(a) \leq x\,\}|$ grows polynomially in x for some (equivalently, any) length function λ for (Γ, F^r) for some (equivalently, any) $r > 0$.

To see that this is truly independent of the λ chosen, suppose $|\{\,a \in A : \lambda(a) \leq x\,\}| \in O(f(x))$ for some polynomial f. Suppose λ' is another length function. By Corollary 3.13 we get $\lambda' \in \Theta(\lambda)$; so there is $M > 0$ such that $\lambda(a) \leq M\lambda'(a)$ for a outside some finite set X. Then

$$|\{\,a \in A : \lambda'(a) \leq x\,\}| \leq |\{\,a \in A : \lambda(a) \leq Mx\,\}| + |X| \in O(f(Mx))$$

grows polynomially in x.

Remark 3.15. F-meagreness coincides with F^r-meagreness. Indeed, fix $s > 0$ for which there is a length function λ for (Γ, F^s). Recall from Remark 2.42 that λ is also a length function for (Γ, F^{rs}); so if $A \subseteq \Gamma$ then

$$A \text{ is } F\text{-meagre} \iff |\{\,a \in A : \lambda(a) \leq x\,\}| \text{ grows polynomially in } x \iff A \text{ is } F^r\text{-meagre.}$$

F-meagreness is intended to capture the polynomial growth condition of F-sparsity without any implication of F-automaticity. We first show:

Lemma 3.16. *If $A \subseteq \Gamma$ is F-sparse then it is F-meagre.*

We will need the following reverse ultrametric inequality:

Remark 3.17. Suppose λ is a length function for (Γ, F) with associated constants C, D, E. If $a, b \in \Gamma$ satisfy $\lambda(b) > \lambda(a) + D$ then $\lambda(a + b) \geq \lambda(b) - D$. Indeed, otherwise

$$\lambda(b) = \lambda((a + b) + (-a)) \leq \max(\lambda(a + b), \lambda(a)) + D < \lambda(b),$$

a contradiction.

Proof of Lemma 3.16. Fix $r > 0$, an F^r-spanning set Σ, and sparse $L \subseteq \Sigma^*$ such that $A = [L]_{F^r}$. Replacing F with F^r, we may assume $r = 1$. It is clear that F-meagreness is closed under finite union; so it suffices to check the case where $L = u_0 v_1^* u_1 \cdots v_n^* u_n$ is simple sparse. We do so by induction on n; the case $n = 0$ is just that singletons are F-meagre.

For the induction step, note first that if $[v_n^* u_n]_F$ is finite, then

$$[L]_F = \bigcup_{a \in [v_n^* u_n]_F} [u_0 v_1^* u_1 \cdots v_{n-1}^* u_{n-1} a]_F$$

is F-meagre by the induction hypothesis. Suppose then that $[v_n^* u_n]_F$ is infinite; we will show that $|\{ a \in [L]_F : \lambda_\Sigma(a) \leq x \}|$ grows polynomially in x. As a notational convenience, let $f(k_1, \ldots, k_n) = u_0 v_1^{k_1} \cdots u_{n-1} v_n^{k_n}$; so we are interested in $\lambda_\Sigma([f(k_1, \ldots, k_n) u_n]_F)$ for $k_1, \ldots, k_n \in \mathbb{N}$. Pick N such that $\lambda_\Sigma([v_n^N u_n]_F) > 2$. Then

$$[L]_F = \left(\bigcup_{i<N} [u_0 v_1^* u_1 \cdots v_{n-1}^* u_{n-1} v_n^i u_n]_F \right) \cup \{ [f(k_1, \ldots, k_n + N) u_n]_F : k_1, \ldots, k_n \in \mathbb{N} \}.$$

and by the induction hypothesis each $[u_0 v_1^* u_1 \cdots v_{n-1}^* u_{n-1} v_n^i u_n]_F$ is F-meagre; it thus suffices to show that $B := \{ [f(k_1, \ldots, k_n + N) u_n]_F : k_1, \ldots, k_n \in \mathbb{N} \}$ is F-meagre.

We are thus interested in placing a lower bound on

$$\lambda_\Sigma([f(k_1, \ldots, k_n + N) u_n]_F) = \lambda_\Sigma([f(k_1, \ldots, k_n)]_F + F^{|f(k_1, \ldots, k_n)|}[v_n^N u_n]_F).$$

We apply the reverse ultrametric inequality. Recall (Proposition 2.39) that the constants associated to λ_Σ are $C = D = E = 1$, and the exceptional set is Σ. We then see from the logarithmic property that

$$\begin{aligned} \lambda_\Sigma(F^{|f(k_1, \ldots, k_n)|}[v_n^N u_n]_F) &\geq \lambda([v_n^N u_n]_F) + |f(k_1, \ldots, k_n)| - 1 \\ &> |f(k_1, \ldots, k_n)| + 1 \\ &\geq \lambda_\Sigma([f(k_1, \ldots, k_n)]_F) + 1 \end{aligned}$$

(since $\lambda_\Sigma([v_n^N u_n]_F) > 2$ implies that $[v_n^N u_n]_F \notin \Sigma$ is not exceptional). So the reverse ultrametric inequality applies, and thus

$$\lambda_\Sigma([f(k_1, \ldots, k_n + N) u_n]_F) \geq \lambda_\Sigma(F^{|f(k_1, \ldots, k_n)|}[v_n^N u_N]_F) - 1 \geq |f(k_1, \ldots, k_n)|.$$

So if $\lambda_\Sigma([f(k_1, \ldots, k_n + N) u_n]_F) \leq x$ then $x \geq |f(k_1, \ldots, k_n)| \geq \max\{ k_1, \ldots, k_n \}$ (under the assumption that no v_i is empty, which is harmless), and there are at most $(x+1)^n$ such $(k_1, \ldots, k_n)$. So B is F-meagre, and thus so too is A. $\square$ Lemma 3.16

The following lemma provides a partial converse, and moreover tells us that F-automaticity and F-sparsity can be verified by looking at a "minimal" set of representations.

Lemma 3.18. *Suppose Σ is F^r-spanning and $L \subseteq \Sigma^*$ is regular. Suppose $\preceq$ is a linear ordering on Σ; we identify $\preceq$ with the length-lexicographical order it induces on Σ^*.*[2] *Then*

$$\widetilde{L} := \{\, \sigma \in L : \sigma \preceq \tau \text{ for every } \tau \in L \text{ such that } [\sigma]_{F^r} = [\tau]_{F^r} \,\}$$

is regular. Moreover the following are equivalent:

1. *$\widetilde{L}$ is sparse.*
2. *$[L]_{F^r}$ is F-sparse.*
3. *$[L]_{F^r}$ is F-meagre.*

We will primarily be interested in the case where L is the set of all representations of some F-automatic set, but the full statement has found some use in the literature; see the bibliographical notes.

Proof. We first verify that $\widetilde{L}$ is regular. By closure under Boolean combinations (Corollary 2.5) it suffices to show that $\Sigma^* \setminus \widetilde{L} = (\Sigma^* \setminus L) \cup (L \setminus \widetilde{L})$ is regular; since L is regular, it thus suffices to show that $L \setminus \widetilde{L}$ is regular. Note that given $\sigma \in L$, the question of whether $\sigma \in L \setminus \widetilde{L}$ is an existential one: we are asking whether there is $\tau \in L$ such that $[\sigma]_{F^r} = [\tau]_{F^r}$ and $\tau \prec \sigma$. We thus look to write $L \setminus \widetilde{L}$ as a projection of a regular language. Let L' be the set of $\begin{pmatrix} \sigma \\ \tau \end{pmatrix} \in (\Sigma^2)^*$ such that $\sigma \in L$, $[\sigma]_{F^r} = [\tau]_{F^r}$, and either

- $\tau \in L$ and $\tau \prec \sigma$, or
- $\tau \in \{\, \tau_0 0^i : \tau_0 \in L, i > 0 \,\} = L00^*$.

I claim that $L \setminus \widetilde{L}$ is the projection of L' onto its first coordinate. Indeed, if $\sigma \in L \setminus \widetilde{L}$ then there is $\tau \in L$ such that $[\tau]_{F^r} = [\sigma]_{F^r}$ and $\tau \prec \sigma$. If $|\tau| < |\sigma|$ then $\begin{pmatrix} \sigma \\ \tau 0^{|\sigma|-|\tau|} \end{pmatrix} \in L'$; if on the other hand $|\tau| = |\sigma|$ then $\begin{pmatrix} \sigma \\ \tau \end{pmatrix} \in L'$. Conversely, suppose $\begin{pmatrix} \sigma \\ \tau \end{pmatrix} \in L'$ for some τ.

[2] Recall that σ precedes τ in the length-lexicographical order induced by $\preceq$ if $|\sigma| < |\tau|$ or if $|\sigma| = |\tau|$ and σ precedes τ in the lexicographical order induced by $\preceq$.

If $\tau \in L$ and $\tau \prec \sigma$, then τ itself witnesses that $\sigma \notin \widetilde{L}$. If on the other hand $\tau = \tau_0 0^i$ for some $\tau_0 \in L$ and $i > 0$, then $\tau_0 \prec \sigma$ since $|\tau_0| < |\tau| = |\sigma|$; so, since $[\tau_0]_{F^r} = [\tau]_{F^r} = [\sigma]_{F^r}$, we get that $\sigma \notin \widetilde{L}$.

So by closure of regular languages under projections (Corollary 2.8) it suffices to show that L' is regular. By closure under Boolean combinations, it suffices to show that L, $L00^*$, $\left\{ \begin{pmatrix} \sigma \\ \tau \end{pmatrix} \in (\Sigma^2)^* : [\sigma]_{F^r} = [\tau]_{F^r} \right\}$, and $\left\{ \begin{pmatrix} \sigma \\ \tau \end{pmatrix} \in (\Sigma^2)^* : \sigma \prec \tau \right\}$ are regular. The first is by hypothesis; the second is by definition of regularity; the third is by Proposition 2.32 and Corollary 2.27; and the fourth is $\Delta^* X (\Sigma^2)^*$, where $\Delta \subseteq \Sigma^2$ is the diagonal and $X = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \Sigma^2 : a \prec b \right\}$, and is thus regular by definition.

We now show the "moreover" clause.

<u>**(1) $\Longrightarrow$ (2)**</u> This is just the definition of F-sparsity, since $\widetilde{L}$ is sparse and $[\widetilde{L}]_{F^r} = [L]_{F^r}$.

<u>**(2) $\Longrightarrow$ (3)**</u> This is Lemma 3.16.

<u>**(3) $\Longrightarrow$ (1)**</u> Note that $[\cdot]_{F^r}$ is a bijection $\widetilde{L} \to [L]_{F^r}$ with the property that $\lambda_\Sigma([\sigma]_{F^r}) \leq |\sigma|$. So

$$|\{ \sigma \in \widetilde{L} : |\sigma| \leq x \}| \leq |\{ a \in [L]_{F^r} : \lambda_\Sigma(a) \leq x \}|$$

grows polynomially in x since $[L]_{F^r}$ is F-meagre. □ Lemma 3.18

We have thus proven:

Theorem 3.19. *$A \subseteq \Gamma$ is F-sparse if and only if it is F-automatic and F-meagre.*

Proof. It suffices to show that if A is F-automatic then A is F-sparse if and only if A is F-meagre. Fix $r > 0$ and an F^r-spanning set Σ; so $L := \{ \sigma \in \Sigma^* : [\sigma]_{F^r} \in A \}$ is regular. The result is then the "moreover" clause of the previous lemma. □ Theorem 3.19

Corollary 3.20. *If $A \subseteq \Gamma$ is F-sparse then it is (F^r, Σ)-sparse for all $r > 0$ and all F^r-spanning sets Σ.*

Proof. Suppose A is F-sparse; suppose $r > 0$ and Σ is an F^r-spanning set. Let $L = \{ \sigma \in \Sigma^* : [\sigma]_{F^r} \in A \}$; so L is regular, since A is F-automatic (using Corollary 2.27). Fix any linear ordering $\preceq$ on Σ, and let $\widetilde{L}$ be as in Lemma 3.18. Then since $A = [L]_{F^r}$ is F-sparse, Lemma 3.18 yields that $\widetilde{L}$ is sparse. Then since $A = [\widetilde{L}]_{F^r}$, it follows that A is (F^r, Σ)-sparse. □ Corollary 3.20

As a consequence, we get that F-sparsity interacts nicely with F-invariant subgroups:

Corollary 3.21. *Suppose $H \leq \Gamma$ is F-invariant and $A \subseteq H$. Then A is F-sparse in Γ if and only if it is $(F{\restriction}H)$-sparse in H.*

Proof. Recall from Proposition 2.55 that there is a spanning set for some power of $F{\restriction}H$, and that A is $(F{\restriction}H)$-automatic in H if and only if it is F-automatic in Γ. Furthermore if $r > 0$ and λ is any length function for (Γ, F^r) then $\lambda{\restriction}H$ is a length function for $(H, F{\restriction}H)$; so A is F-meagre in Γ if and only if A is $(F{\restriction}H)$-meagre in H. Theorem 3.19 then yields that A is F-sparse in Γ if and only it is $(F{\restriction}H)$-sparse in H. □ Corollary 3.21

Finally, we see that there are sets that aren't F-sparse.

Proposition 3.22. *Suppose $A \subseteq \Gamma$ is F-sparse. Then:*

1. *A does not contain a coset of an infinite F-invariant subgroup.*
2. *A does not contain $\mathbb{Z}a$ for any non-torsion $a \in \Gamma$.*

Proof.

1. We first show that Γ itself is not F-meagre. Replacing F with a power thereof, which is harmless by Remark 3.15, we may assume there is an F-spanning set Σ. We will show that $|\{ a \in \Gamma : \lambda_\Sigma(a) \leq x \}|$ grows exponentially in x. Recall from Proposition 2.39 that the constants associated to λ_Σ are $C = D = E = 1$; moreover, we noted in Lemma 2.49 that by increasing E we can assume the exceptional set only contains elements of finite F-orbit.

   Fix $a \in \Sigma$ with $F^{\mathbb{N}}a$ infinite; so the logarithmic property applies to a. Note that such a must exist, since otherwise Lemma 2.50 would yield that all $a \in \Sigma$ are torsion, and hence that
   $$\Gamma = \{ [\sigma]_F : \sigma \in \Sigma^* \} = \sum_{a \in \Sigma} \sum_{b \in F^{\mathbb{N}}a} \mathbb{Z}b$$
   is finite, a contradiction.

   Fix $s > E + 1$. For $n \in \mathbb{N}$ consider $A_n := \{ [\tau]_{F^s} : \tau \in \{ 0, a \}^{(n)} \}$. Note for $\tau = t_0 \cdots t_{n-1} \in \{ 0, a \}^{(n)}$ we have that
   $$[\tau]_{F^s} = t_0 + F^s t_1 + \cdots + F^{s(n-1)} t_{n-1} = [t_0 0^{s-1} t_1 0^{s-1} \cdots t_{n-2} 0^{s-1} t_{n-1}]_F,$$

and hence that $\lambda_\Sigma([\tau]_{F^s}) \leq s(|\tau|-1)+1$. So $\lambda_\Sigma(b) \leq s(n-1)+1$ for $b \in A_n$. I claim that $|A_n| = 2^n$. This will then imply that $|\{ b \in \Gamma : \lambda_\Sigma(a) \leq s(x-1)+1 \}| \geq 2^n$ for $x \in \mathbb{N}$; hence Γ isn't F-meagre, and hence by Theorem 3.19 we get that Γ isn't F-sparse.

To prove this, we wish to show that the map $\Phi\colon \{ 0, a \}^{(n)} \to \Gamma$ given by $\tau \mapsto [\tau]_{F^s}$ is injective. Note, however, that given distinct $\sigma, \tau \in \{ 0, a \}^{(n)}$ we can write $[\sigma]_{F^s} - [\tau]_{F^s} = [\nu]_{F^s}$ for some $\nu \in \{ -a, 0, a \}^* \setminus \{ 0 \}^*$. It then suffices to show that given $\nu \in \{ -a, 0, a \}^* \setminus \{ 0 \}^*$ we must have $[\nu]_{F^s} \neq 0$.

Suppose we are given such ν; we will show that $\lambda_\Sigma([\nu]_{F^s}) \neq 0$. The result is clear when $|\nu| = 1$; suppose then that $|\nu| \geq 2$. By negating and removing trailing zeroes, we may assume ν ends in a; say $\nu = \nu_0 a$ for some $\nu_0 \in \{ -a, 0, a \}^*$. Then $[\nu]_{F^s} = [\nu_0]_{F^s} + F^{s|\nu_0|}a$; we look to apply the reverse ultrametric inequality. Applying our earlier bound to ν_0, we get that

$$\lambda_\Sigma([\nu_0]_{F^s}) \leq s(|\nu_0|-1)+1 = \lambda_\Sigma(a)+s|\nu_0|-s < \lambda_\Sigma(a)+s|\nu_0|-E-1 \leq \lambda_\Sigma(F^{s|\nu_0|}a)-1$$

(using the logarithmic property, and recalling that $C = D = 1$ and $s > E+1$). So, by the reverse ultrametric inequality, we get that

$$\lambda_\Sigma([\nu]_{F^s}) \geq \lambda_\Sigma(F^{s|\nu_0|}a) - 1 \geq s(|\nu_0|-1)+1 \geq 1$$

(since by assumption $|\nu| \geq 2$, and thus $|\nu_0| \geq 1$). So $[\nu]_{F^s} \neq 0$.

It then follows that Φ is injective. So $2^n = |\Phi(\{ 0, a \}^{(n)})| \leq |A_n|$, and thus Γ is not F-meagre, as desired.

We now prove the full statement. By closure of F-sparsity under translation (Remark 3.9) it suffices to show that A doesn't contain an infinite F-invariant subgroup; for this, it suffices to show that if $H \leq \Gamma$ is F-invariant then H isn't F-meagre. Recall from Proposition 2.55 that there is a spanning set for some power of $(F{\restriction}H)$; so our work above applies to $(H, F{\restriction}H)$, and hence H is not $(F{\restriction}H)$-meagre. But if $r > 0$ and λ is any length function for (Γ, F^r), then $\lambda{\restriction}H$ is a length function for $(H, (F{\restriction}H)^r)$, and hence $|\{ b \in H : \lambda(b) \leq x \}|$ grows exponentially in x; so H is not F-meagre, as desired.

2. Suppose $a \in \Gamma$ is non-torsion; we will show that $\mathbb{Z}a$ is not F-meagre, and hence by Theorem 3.19 that no superset of $\mathbb{Z}a$ is F-sparse. Replacing F with a power thereof, we may assume there is a length function λ for (Γ, F), say with associated constants C, D, E. We show by induction on $k \in \mathbb{N}$ that if $1 \leq i \leq 2^k$ then $\lambda(ia) \leq \lambda(a) + kD$. The base case $k = 0$ is clear. For the induction step, suppose

the claim holds of k. If $1 \leq i \leq 2^k$ then the induction hypothesis yields that $\lambda(ia) \leq \lambda(a) + kD \leq \lambda(a) + (k+1)D$. If on the other hand $2^k \leq i \leq 2^{k+1}$ then the ultrametric inequality and the induction hypothesis yield that

$$\begin{aligned}
\lambda(ia) &= \lambda(2^k a + (i - 2^k)a) \\
&\leq \max(\lambda(2^k a), \lambda((i - 2^k)a)) + D \\
&\leq \max(\lambda(a) + kD, \lambda(a) + kD) + D \\
&= \lambda(a) + (k+1)D,
\end{aligned}$$

as desired.

Hence $|\{\, b \in \mathbb{Z}a : \lambda(b) \leq \lambda(a) + xD \,\}| \geq 2^x$ for $x \in \mathbb{N}$; so $\mathbb{Z}a$ isn't F-meagre, and thus no superset of $\mathbb{Z}a$ is F-sparse. □ Proposition 3.22

3.3 Bibliographical notes

The definitions of sparsity and F-sparsity were taken from [6]. Corollary 3.11 is proven as part of [6, Theorem 7.4].

All other definitions and results in this chapter are original. This includes the closure properties of F-sparse sets (Proposition 3.6 and Corollary 3.10); the definition of F-meagreness; the characterization of F-sparsity in terms of F-meagreness (Theorem 3.19); the independence of F-sparsity from the chosen spanning set (Corollary 3.20); and the fact that F-sparse sets cannot contain certain infinite subgroups (Proposition 3.22).

When F-sparsity was introduced in [6], the handling was somewhat cumbersome because the authors couldn't use independence from the chosen spanning set (Corollary 3.20). They also asked whether a translate of an F-sparse set is necessarily F-sparse, which we answered here in the affirmative in Corollary 3.10.

The (2) $\implies$ (1) direction of Lemma 3.18 found use in [5], as did Proposition 3.22 (2).

Chapter 4

Stable sparse sets

In this chapter, we consider the question of which F-sparse subsets of Γ are *stable*. Stability is an important model-theoretic tameness condition that rules out certain infinite combinatorial configurations. We briefly recall it here; more information can be found in [23, Chapter 8].

Definition 4.1. Suppose X and Y are sets, and $R \subseteq X \times Y$ is a binary relation. If $N \in \mathbb{N}$ then an N*-ladder* for R is a tuple $(a_0, \dots, a_{N-1}; b_0, \dots, b_{N-1})$ with each $a_i \in X$ and $b_j \in Y$ such that $(a_i, b_j) \in R$ if and only if $i \leq j$. We say R is *stable* if there is a bound on the N for which there is an N-ladder for R.

If $(G, \cdot, e)$ is a group and $A \subseteq G$, we say A is *stable in the group* G if the binary relation $xy \in A$ on G^2 is stable.

If T is a complete first-order theory and $\varphi(x_1, \dots, x_k; y_1, \dots, y_\ell)$ is a partitioned formula in the associated signature, we say φ is *stable (in T)* if there is $M \models T$ such that the relation on $M^k \times M^\ell$ defined by φ is stable. (Since T is complete, it decides the existence of N-ladders for φ; hence we can replace "there is $M \models T$" in the above definition with "for any $M \models T$".) We say T is *stable* if every partitioned formula in the associated signature is stable in T.

Suppose $(\Gamma, +)$ is an infinite abelian group, $F \colon \Gamma \to \Gamma$ is an injective endomorphism, and Γ admits an F^r-spanning set for some $r > 0$. We consider the question of which F-sparse $A \subseteq \Gamma$ are stable in $(\Gamma, +)$. Note first that not all of them are:

Example 4.2. Let $\Gamma = (\mathbb{Z}^2, +)$ and $F \colon \Gamma \to \Gamma$ be multiplication by some $d \geq 2$. Let $A = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{Z}^2 : a, b \in d^{\mathbb{N}}, a < b \right\}$. We saw in Example 3.5 (3) that A is F-sparse.

However, if $i, j \in \mathbb{N}$ then

$$\begin{pmatrix} d^i \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ d^{j+1} \end{pmatrix} \in A \iff i \leq j,$$

and thus if $N \in \mathbb{N}$ then $\left(\begin{pmatrix} d^0 \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} d^{N-1} \\ 0 \end{pmatrix}; \begin{pmatrix} 0 \\ d^1 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ d^N \end{pmatrix} \right)$ forms an N-ladder for $x + y \in A$. So A isn't stable in Γ.

Our characterization of the stable F-sparse sets will make use of the following result of Moosa and Scanlon, which under certain conditions gives a sufficient condition for stability:

Fact 4.3 ([18, Theorem A])**.** *Let $\mathbb{Z}[F] \subseteq \mathrm{End}(\Gamma)$ be the $\mathbb{Z}$-subalgebra generated by F. Suppose that $\mathbb{Z}[F]$ is a finite extension of $\mathbb{Z}$, that Γ is a finitely generated $\mathbb{Z}[F]$-module, and that*

$$\bigcap_{i \in \mathbb{N}} F^i \mathbb{Z}[F] = \{ 0 \}.$$

*Then $\mathrm{Th}(\Gamma, \mathcal{F})$ is stable and admits quantifier elimination. (Recall that $(\Gamma, \mathcal{F})$ is the F-*structure *on Γ, which has domain Γ and a predicate for every F-set in every Γ^m.)*

The fact that this last hypothesis holds in our context will come from our assumption that Γ admits an F^r-spanning set for some $r > 0$.

So under these hypotheses, the groupless F-sets are stable in Γ. Moreover, we showed in Corollary 3.11 that groupless F-sets are F-sparse. So (under some additional assumptions on $\mathbb{Z}[F]$) we have a class of examples of stable F-sparse sets, namely the groupless F-sets.

It turns out that when Γ is finitely generated these are the only stable F-sparse sets, up to Boolean combinations:

Theorem. *Suppose Γ is finitely generated and $A \subseteq \Gamma$ is F-sparse. Then A is stable in Γ if and only if A is a Boolean combination of elementary F-sets.*

This is the goal of the chapter; it appears as part of Theorem 4.11 below.

We continue to assume that Γ is an infinite abelian group, that $F \colon \Gamma \to \Gamma$ is an injective endomorphism, and that Γ admits an F^r-spanning set for some $r > 0$.

4.1 An auxiliary structure on $\mathbb{N}$

To characterize the stable F-sparse sets, we will need to understand certain stable formulas on natural numbers. Consider the first-order signatures $L_{\text{div}} := \{ 0, S \} \cup \{ P_\delta : \delta \in \mathbb{N}, \delta \geq 2 \}$ and $L_{\text{div},<} := L_{\text{div}} \cup \{ < \}$, where

- 0 is a constant symbol,
- S is a unary function symbol,
- each P_δ is a unary relation symbol, and
- $<$ is a binary relation symbol.

Let $\mathfrak{N} = (\mathbb{N}, 0, S, (\delta\mathbb{N})_{\delta \geq 2}, <)$, where S is the successor function and P_δ is interpreted as the set $\delta\mathbb{N}$ of natural numbers divisible by δ. We fix $\mathrm{Th}(\mathfrak{N})$ as our ambient theory.

The order relation is the only obvious source of instability in $\mathfrak{N}$, and indeed:

Proposition 4.4. *If $\varphi \in L_{\text{div},<}$ is quantifier-free and stable under any partition of its variables then φ is equivalent to a quantifier-free L_{div}-formula.*[1]

Proof. Write $\varphi = \varphi(\mathbf{x}) = \varphi(x_1, \ldots, x_n)$. We apply induction on n.

The base case $n = 0$ is vacuous, since any sentence is equivalent to either $\exists x(x = x)$ or $\exists x(x \neq x)$. (A word on what stability means in absence of variables: according to the definition, stability of $\varphi(;)$ says that there is $M \models T$ such that φ defines a stable relation on $M^0 \times M^0$. But there is no unstable relation on $M^0 \times M^0$; so φ is automatically stable.)

We describe the atomic $L_{\text{div},<}$-formulas. An $L_{\text{div},<}$-term takes the form $S^e x$ or $S^e 0$ (which we simply write as e) for $e \in \mathbb{N}$. An atomic $L_{\text{div},<}$-formula thus takes one of the following forms, for some $e, f \in \mathbb{N}$, some $\delta \geq 2$, and distinct variables x, y:

(i) $e = f$, $P_\delta(e)$, $e < f$, $S^e x = S^f x$, or $S^e x < S^f x$

(ii) $S^e x = f$

(iii) $S^e x = S^f y$

[1] In fact both $\mathrm{Th}(\mathfrak{N})$ and $\mathrm{Th}(\mathbb{N}, 0, S, (\delta\mathbb{N})_{\delta \geq 2})$ admit quantifier elimination, but we will not make use of this.

(iv) $P_\delta(S^e x)$

(v) $S^e x < f$

(vi) $e < S^f x$

(vii) $S^e x < S^f y$

So φ is a Boolean combination of formulas of the above form. We now describe how to remove or simplify some of these forms. The arguments below are arranged so that later simplifications will not ruin earlier ones.

- Formulas of form (i) are either inconsistent or necessarily true, and can thus be removed from φ.

- For form (vii), we first note that $S^e x < S^f y$ is equivalent to $\neg((S^e x = S^f y) \vee (S^f y < S^e x))$. Thus if $S^e x < S^f y$ occurs in φ, then by possibly applying the above substitution we can reduce to the case where $e \geq f$. Moreover, in this case we get that $S^e x < S^f y$ is equivalent to $S^{e-f} x < y$. We may thus assume that $f = 0$ whenever $S^e x < S^f y$ appears in φ.

- For form (vi), we note that $e < S^f x$ is equivalent to $\neg((S^f x = e) \vee (S^f x < e))$. Applying this substitution, we may assume that form (vi) does not occur in φ.

- For form (v), if $e \geq f$ then this is inconsistent, and can thus be removed; otherwise we can replace $S^e x < f$ with $\bigvee_{k<f} S^e x = k$. We may thus assume that form (v) does not occur in φ.

- For form (ii), if $e > f$ then $S^e x = f$ is inconsistent, and can thus be removed; otherwise we can replace $S^e x = f$ with $x = f - e$, and thus assume $e = 0$.

- For form (iii), by symmetry we may assume $e \leq f$, in which case $S^e x = S^f y$ is equivalent to $x = S^{f-e} y$; we may thus assume that $e = 0$ whenever $S^e x = S^f y$ appears in φ.

- For form (iv), note if $k > 0$ that $P_\delta(x)$ is equivalent to $\bigvee_{e<k} P_{k\delta}(S^{e\delta} x)$. Thus if δ' is the least common multiple of all the δ for which P_δ appears in φ, we can replace any P_δ occurring in φ with a Boolean combination of $P_{\delta'}$; so we can assume that all P_δ occurring in φ share the same δ.

  For ease of comprehension we will write $P_\delta(S^e x)$ as $x \equiv e' \pmod{\delta}$, where $0 \leq e' < \delta$ and $e \equiv -e' \pmod{\delta}$.

Altogether, there is some δ for which we can write $\varphi = \varphi(\mathbf{x}) = \varphi(x_1, \ldots, x_n)$ as a Boolean combination of formulas of the following forms:

$$x_i = e \qquad x_i = S^e x_j \qquad S^e x_i < x_j \qquad x_i \equiv e \pmod{\delta}$$

for $e \in \mathbb{N}$, $\delta \geq 2$, and distinct $i, j \in \{1, \ldots, n\}$. Let M be the largest e for which a formula of the above form appears in φ, and let Δ be the collection of formulas of the above forms with $e \leq M$; so Δ is finite. So whether φ holds of some $\mathbf{a} \in \mathbb{N}^n$ is determined by $\operatorname{tp}_\Delta(\mathbf{a}) := \{\psi \in \Delta : \mathfrak{N} \models \psi(\mathbf{a})\}$. We can thus write $\varphi(\mathbf{x})$ as

$$\bigvee_{p \in X} \underbrace{\left(\bigwedge_{\psi \in p} \psi(\mathbf{x}) \wedge \bigwedge_{\psi \in \Delta \setminus p} \neg\psi(\mathbf{x}) \right)}_{\theta_p}$$

for some $X \subseteq \mathcal{P}(\Delta)$. We may further assume that θ_p is consistent for each $p \in X$. We will show given $p \in X$ that there is a quantifier-free L_{div}-formula $\theta'_p(\mathbf{x})$ such that $\theta_p(\mathfrak{N}) \subseteq \theta'_p(\mathfrak{N}) \subseteq \varphi(\mathfrak{N})$. It will then follow that $\varphi(\mathbf{x})$ is equivalent to the quantifier-free L_{div}-formula $\bigvee_{p \in X} \theta'_p(\mathbf{x})$, completing the proof.

There are two cases:

Case 1. Suppose p contains a formula of the form $x_i = e$ or $x_i = S^e x_j$; assume for notational convenience that $i = 1$. Define the L_{div}-term t to be e in the former case and $S^e x_j$ in the latter. Let $\chi(x_2, \ldots, x_n)$ be $\varphi(t, x_2, \ldots, x_n)$.

Then χ is stable under any partition of its variables. Indeed, suppose otherwise; for notational convenience, assume the unstable partition is $\chi(x_2, \ldots, x_{n_0}; x_{n_0+1}, \ldots, x_n)$ for some n_0. So for any $N \in \mathbb{N}$ there are $a_{k2}, \ldots, a_{kn_0}, b_{k(n_0+1)}, \ldots, b_{kn} \in \mathbb{N}$ for $k < N$ such that

$$\mathfrak{N} \models \chi(a_{k2}, \ldots, a_{kn_0}; b_{\ell(n_0+1)}, \ldots, b_{\ell n}) \iff k \leq \ell.$$

If $t = e$ for some $e \in \mathbb{N}$ then this means that $\varphi(x_1, x_2, \ldots, x_{n_0}; x_{n_0+1}, \ldots, x_n)$ is unstable, since

$$\mathfrak{N} \models \varphi(e, a_{k2}, \ldots, a_{kn_0}; b_{\ell(n_0+1)}, \ldots, b_{\ell n}) \iff k \leq \ell.$$

If $t = x_j$ for some $2 \leq j \leq n_0$ then we again get that $\varphi(x_1, x_2, \ldots, x_{n_0}; x_{n_0+1}, \ldots, x_n)$ is unstable, since

$$\mathfrak{N} \models \varphi(a_{kj}, a_{k2}, \ldots, a_{kn_0}; b_{\ell(n_0+1)}, \ldots, b_{\ell n}) \iff k \leq \ell.$$

If $t = x_j$ for some $n_0 + 1 \leq j \leq n$, we instead get that $\varphi(x_2, \dots, x_{n_0}; x_1, x_{n_0+1}, \dots, x_n)$ is unstable, since

$$\mathfrak{N} \models \varphi(a_{k2}, \dots, a_{kn_0}; b_{\ell j}, b_{\ell(n_0+1)}, \dots, b_{\ell n}) \iff k \leq \ell.$$

Thus in any case we get that φ is unstable with respect to some partition of its variables, a contradiction.

So χ has one fewer variable than φ, and is also stable under any partition of its variables. So by the induction hypothesis we get that χ is equivalent to some quantifier-free $\chi' \in L_{\text{div}}$. We then take θ'_p to be $(x_1 = t) \wedge \chi'(x_2, \dots, x_n)$. Then $\theta_p(\mathbf{x})$ implies $x_1 = t$ and $\varphi(\mathbf{x})$ by assumption, and hence implies $\varphi(t, x_2, \dots, x_n)$; so $\theta_p(\mathfrak{N}) \subseteq \theta'_p(\mathfrak{N})$. Moreover by definition of χ we get that θ'_p implies $(x_1 = t) \wedge \varphi(t, x_2, \dots, x_n)$, and hence $\varphi(\mathbf{x})$; so $\theta'_p(\mathfrak{N}) \subseteq \varphi(\mathfrak{N})$. So this choice of θ'_p works as desired.

Case 2. Suppose p contains no formula of the form $x_i = e$ or $x_i = S^e x_j$. Let $\chi(\mathbf{x})$ be the conjunction of the negations of such formulas; so χ asserts that no $x_i \leq M$ and no x_i is within M of some x_j for $i \neq j$. Examining Δ, and using the fact that θ_p is consistent and decides every formula in Δ, we see that θ_p can be written in the form

$$\underbrace{\chi(\mathbf{x}) \wedge \left(\bigwedge_{i=1}^{n} x_i \equiv e_i \pmod{\delta} \right)}_{\theta'_p} \wedge (x_{\sigma(1)} < \cdots < x_{\sigma(n)})$$

for some $e_1, \dots, e_n < \delta$ and some $\sigma \in S_n$ (the symmetric group on $\{1, \dots, n\}$). (Note that formulas of the form $S^e x_{\sigma(i)} < x_{\sigma(j)}$ for $e \leq M$ and $i < j$ are implied by $\chi(\mathbf{x})$ and $x_{\sigma(i)} < x_{\sigma(j)}$, and can thus be omitted.)

We will show that this θ'_p satisfies the desired properties. It is clear that $\theta_p(\mathfrak{N}) \subseteq \theta'_p(\mathfrak{N})$; it remains to show that $\theta'_p(\mathfrak{N}) \subseteq \varphi(\mathfrak{N})$. Let

$$T = \{ \tau \in S_n : \mathfrak{N} \models \forall \mathbf{x}((\theta'_p(\mathbf{x}) \wedge (x_{\tau(1)} < \cdots < x_{\tau(n)})) \to \varphi(\mathbf{x})) \}.$$

Since $\theta_p(\mathfrak{N}) \subseteq \varphi(\mathfrak{N})$ we get that $\sigma \in T$. To show that $\theta'_p(\mathfrak{N}) \subseteq \varphi(\mathfrak{N})$ it suffices to show that $T = S_n$. Indeed, if $\mathbf{a} \in \theta'_p(\mathfrak{N})$ then $\mathbf{a} \in \chi(\mathfrak{N})$, and thus the a_i are all distinct; so there is $\tau \in S_n$ such that $\mathfrak{N} \models \theta'_p(\mathbf{a}) \wedge (a_{\tau(1)} < \cdots < a_{\tau(n)})$, and hence if $T = S_n$ we can conclude that $\mathbf{a} \in \varphi(\mathfrak{N})$. Since S_n is generated by permutations of the form $(j\ j+1)$ for $1 \leq j < n$, and since we know T is non-empty, it suffices to show that if $\tau \in T$ then so too is $(j\ j+1)\tau$.

Suppose for contradiction that we had $\tau \in T$ with $(j\ j+1)\tau \notin T$; assume for notational convenience that $\tau = \mathrm{id}$. Then

$$\theta'_p(\mathbf{x}) \wedge (x_{(j\ j+1)(1)} < \cdots < x_{(j\ j+1)(n)})$$

decides every formula in Δ, and thus takes the form θ_q for some $q \subseteq \Delta$. In particular, since by assumption we have that $\theta_q(\mathfrak{N}) \not\subseteq \varphi(\mathfrak{N})$, and since whether $\mathfrak{N} \models \varphi(\mathbf{a})$ is determined by $\mathrm{tp}_\Delta(\mathbf{a})$, we get that $\theta_q(\mathfrak{N}) \cap \varphi(\mathfrak{N}) = \emptyset$. We will use this to contradict the stability of $\varphi(x_1, \ldots, x_j; x_{j+1}, \ldots, x_n)$.

Fix $N < \omega$; we construct an N-ladder for $\varphi(x_1, \ldots, x_j; x_{j+1}, \ldots, x_n)$. Inductively choose

$$a_1, \ldots, a_{j-1}, \quad a_{j,0}, \ldots, a_{j,N-1}, \quad a_{j+1,0}, \ldots, a_{j+1,N-1}, \quad a_{j+2}, \ldots, a_n \in \mathbb{N}$$

such that

- $a_i \equiv e_i \pmod{\delta}$ for $i \in \{1, \ldots, n\} \setminus \{j, j+1\}$;
- $a_{j,k} \equiv e_j \pmod{\delta}$ for $k < N$;
- $a_{j+1,\ell} \equiv e_{j+1} \pmod{\delta}$ for $k < N$; and
- The $a_i, a_{j,k}, a_{j+1,\ell}$ are ordered as in the following diagram:

$$\begin{tikzcd}
0 \arrow[r] & a_1 \arrow[r] & \cdots \arrow[r] & a_{j-1} \arrow[r] & a_{j,0} \arrow[r] & a_{j+1,0} \arrow[dl] \\
& & & & a_{j,1} \arrow[r] & a_{j+1,1} \arrow[dl] \\
& & & & \cdots & {} \arrow[dl] \\
& & & & a_{j,N-1} \arrow[r] & a_{j+1,N-1} \arrow[dllll] \\
& a_{j+2} \arrow[r] & \cdots \arrow[r] & a_n & &
\end{tikzcd}$$

where an arrow indicates that the target exceeds the source plus M.

Now, for $k < N$ let $\mathbf{a}_k = (a_1, \ldots, a_{j-1}, a_{j,k})$ and $\mathbf{b}_k = (a_{j+1,k}, a_{j+2}, \ldots, a_n)$. Then by construction for any k, ℓ we get that $\mathfrak{N} \models \theta'_p(\mathbf{a}_k, \mathbf{b}_\ell)$. Moreover if $k \leq \ell$ then $(\mathbf{a}_k, \mathbf{b}_\ell)$

satisfies $x_1 < \cdots < x_n$; so since we assumed $\mathrm{id} = \tau \in T$, it follows that $\mathfrak{N} \models \varphi(\mathbf{a}_k, \mathbf{b}_\ell)$. If on the other hand $k > \ell$ then $(\mathbf{a}_k, \mathbf{b}_\ell)$ satisfies $x_{(j\ j+1)(1)} < \cdots < x_{(j\ j+1)(n)}$; so $\mathfrak{N} \models \theta_q(\mathbf{a}_k, \mathbf{b}_\ell)$, and thus $\mathfrak{N} \not\models \varphi(\mathbf{a}_k, \mathbf{b}_\ell)$.

So $\mathfrak{N} \models \varphi(\mathbf{a}_k, \mathbf{b}_\ell)$ if and only if $k \leq \ell$, and we have exhibited an N-ladder for $\varphi(x_1, \ldots, x_j; x_{j+1}, \ldots, x_n)$. But this contradicts our assumption that φ is stable under any partition of its variables. So no such j exists. So T is all of S_n, and thus $\theta'_p(\mathfrak{N}) \subseteq \varphi(\mathfrak{N})$, as desired. $\square$ Proposition 4.4

We will need the following general fact relating regular languages to $\mathfrak{N}$:

Lemma 4.5. *Suppose Σ is a finite alphabet and $a_1, \ldots, a_n \in \Sigma$. Suppose $L \subseteq \Sigma^*$ is regular. Then the relation $\{\, (k_1, \ldots, k_n) \in \mathbb{N}^n : a_1^{k_1} \cdots a_n^{k_n} \in L \,\}$ is definable by a Boolean combination of formulas of the form $x_i = K$ and $x_i \equiv K \pmod{\delta}$ for $K \in \mathbb{N}$ and $\delta \geq 2$.*

Proof. Fix a DFA $(\Sigma, Q, q_0, \Omega, \delta)$ recognizing L. We will show by induction on n that for all $q, q' \in Q$ the relation $\{\, (k_1, \ldots, k_n) \in \mathbb{N}^n : \delta(q, a_1^{k_1} \cdots a_n^{k_n}) = q' \,\}$ is definable by a formula of the desired form. Note that the result then follows, since

$$a_1^{k_1} \cdots a_n^{k_n} \in L \iff \bigvee_{q \in \Omega} \delta(q_0, a_1^{k_1} \cdots a_n^{k_n}) = q.$$

The base case $n = 0$ is vacuous. For the induction step, let $q_i = \delta(q, a_1^i)$ for $i \in \mathbb{N}$; so $\delta(q, a_1^{k_1} \cdots a_n^{k_n}) = \delta(q_{k_1}, a_2^{k_2} \cdots a_n^{k_n})$. Note that since Q is finite and $q_{i+1} = \delta(q_i, a_i)$ we get that q_i is ultimately periodic in i; say with M, μ such that $q_{i+\mu} = q_i$ for $i \geq M$. Then $\delta(q, a_1^{k_1} \cdots a_n^{k_n}) = q'$ if and only if

$$\bigvee_{i<M} ((k_1 = i) \wedge (\delta(q_i, a_2^{k_2} \cdots a_n^{k_n}) = q')) \vee \bigvee_{i=M}^{M+\mu-1} ((k_1 \equiv i \pmod{\mu}) \wedge (\delta(q_i, a_2^{k_2} \cdots a_n^{k_n}) = q')),$$

which can be written in the desired form by the induction hypothesis. $\square$ Lemma 4.5

4.2 Characterizing the stable sparse sets

We now work towards our desired classification of the F-sparse sets that are stable.

We will need the following refinement of Corollary 3.8:

Lemma 4.6. *Suppose $A \subseteq \Gamma$ is F-sparse. Then there is $s_0 > 0$ such that if $s_0 \mid s$ then A is a finite union of sets of the form $\alpha + \{ [b_1^{e_1}]_{F^s} + \cdots + [b_n^{e_n}]_{F^s} : e_1 \leq \cdots \leq e_n \}$ where $\alpha, b_1, \ldots, b_n \in \Gamma$.*

Proof. By Corollary 3.8 there is $s_0 > 0$ such that if $s_0 \mid s$ then A can be written as a finite union of sets of the form $\alpha + [a_1^* \cdots a_n^*]_{F^s}$ for $\alpha, a_1, \ldots, a_n \in \Gamma$. Given such $\alpha, a_1, \ldots, a_n$, fix $b_1, \ldots, b_n \in \Gamma$ such that $a_i = b_i + b_{i+1} + \cdots + b_n$. Note then for $e_1 \leq \cdots \leq e_n$ in $\mathbb{N}$ that

$$\begin{aligned}
[b_1^{e_1}]_F + \cdots + [b_n^{e_n}]_F = {} & [b_1^{e_1}]_F + \\
& [b_2^{e_1}]_F + F^{e_1}[b_2^{e_2-e_1}]_F + \\
& [b_3^{e_1}]_F + F^{e_1}[b_3^{e_2-e_1}]_F + F^{e_2}[b_3^{e_3-e_2}]_F + \\
& \vdots \\
& [b_n^{e_1}]_F + F^{e_1}[b_n^{e_2-e_1}]_F + F^{e_2}[b_n^{e_3-e_2}]_F + \cdots + F^{e_{n-1}}[b_n^{e_n-e_{n-1}}]_F \\
= {} & [a_1^{e_1}]_F + F^{e_1}[a_2^{e_2-e_1}]_F + F^{e_2}[a_3^{e_3-e_2}]_F + \cdots + F^{e_{n-1}}[a_n^{e_n-e_{n-1}}]_F \\
= {} & [a_1^{e_1} a_2^{e_2-e_1} \cdots a_n^{e_n-e_{n-1}}]_F.
\end{aligned}$$

Hence $\alpha + [a_1^* \cdots a_n^*]_F = \alpha + \{ [b_1^{e_1}]_F + \cdots + [b_n^{e_n}]_F : e_1 \leq \cdots \leq e_n \}$, and A can be written in the desired form. □ Lemma 4.6

Using this, we can relate the F-sparse stable sets to F-sets. In fact, this connection doesn't require that Γ be finitely generated.

Proposition 4.7. *If $A \subseteq \Gamma$ is F-sparse and stable in Γ then A is definable in $(\Gamma, \mathcal{F})$.*

Proof of Proposition 4.7. By Lemma 4.6 and Proposition 2.17 there is s for which

- Γ admits an F^s-spanning set, and
- A can be written as a finite union $\bigcup_{i=1}^{M} A_i$ where each A_i takes the form $\alpha + \{ [b_1^{e_1}]_{F^s} + \cdots + [b_n^{e_n}]_{F^s} : e_1 \leq \cdots \leq e_n \}$ for some $\alpha, b_1, \ldots, b_n \in \Gamma$.

Since F-sparsity coincides with F^s-sparsity (Proposition 3.6) and the F^s-structure on Γ is a reduct of the F-structure on Γ (which follows from the definition), we can replace F with F^s, and thus assume $s = 1$. Fix i, and write $A_i = \alpha + \{ [b_1^{e_1}]_F + \cdots + [b_n^{e_n}]_F : e_1 \leq \cdots \leq e_n \}$. We will produce $A_i' \subseteq \Gamma$ definable in $(\Gamma, \mathcal{F})$ such that $A_i \subseteq A_i' \subseteq A$. It will then follow that $A = \bigcup_{i=1}^{M} A_i'$ is definable in $(\Gamma, \mathcal{F})$.

Let $A_i' = A \cap (\alpha + K(b_1;F) + \cdots + K(b_n;F)) = A \cap (\alpha + [b_1^*]_F + \cdots + [b_n^*]_F)$. It is clear that $A_i \subseteq A_i' \subseteq A$; it remains to show that A_i' is definable in $(\Gamma, \mathcal{F})$. For this we will consider

$$X := \{ (e_1, \ldots, e_n) \in \mathbb{N}^n : \alpha + [b_1^{e_1}]_F + \cdots + [b_n^{e_n}]_F \in A \}$$

and look to apply our result (Proposition 4.4) on stable subsets of $\mathfrak{N} = (\mathbb{N}, 0, S, (\delta\mathbb{N})_{\delta \geq 2}, <)$. We first check that X is quantifier-free definable in $\mathfrak{N}$. Indeed, fix $f \in S_n$, and let

$$X_f := X \cap \{ (e_1, \ldots, e_n) \in \mathbb{N}^n : e_{f(1)} \leq \cdots \leq e_{f(n)} \}.$$

We will show that X_f is quantifier-free definable in $\mathfrak{N}$, and hence that $X = \bigcup_{f \in S_n} X_f$ is as well.

Assume for notational simplicity that $f = \mathrm{id}$. Then as we argued above, if $a_i = b_i + b_{i+1} + \cdots + b_n$ then for $e_1 \leq \cdots \leq e_n$ we have

$$[b_1^{e_1}]_F + \cdots + [b_n^{e_n}]_F = [a_1^{e_1} a_2^{e_2 - e_1} \cdots a_n^{e_n - e_{n-1}}]_F.$$

Recall by assumption that Γ admits an F-spanning set, say Σ; by Proposition 2.15 we may assume $a_1, \ldots, a_n \in \Sigma$. Since A is F-sparse, in particular it is F-automatic; closure of F-automaticity under translation (Corollary 2.35) then yields that $A - \alpha$ is F-automatic, and thus (F, Σ)-automatic (by Corollary 2.27). So $\{ \sigma \in \Sigma^* : [\sigma]_F \in A - \alpha \}$ is regular. Then Lemma 4.5 yields that $\{ (k_1, \ldots, k_n) \in \mathbb{N}^n : [a_1^{k_1} a_2^{k_2} \cdots a_n^{k_n}]_F \in A - \alpha \}$ is definable by some Boolean combination of formulas of the form $x_i = K$ and $x_i \equiv K \pmod{\delta}$. So for $e_1 \leq \cdots \leq e_n$ we get that

$$\alpha + [a_1^{e_1} a_2^{e_2 - e_1} \cdots a_n^{e_n - e_{n-1}}]_F \in A$$

is equivalent to some Boolean combination of statements of the following forms:

- $e_1 = K$.
- $e_1 \equiv K \pmod{\delta}$.
- $e_{j+1} - e_j = K$; note that this is equivalent to $e_{j+1} = S^K e_j$.
- $e_{j+1} - e_j \equiv K \pmod{\delta}$; note that this is equivalent to

$$\bigvee_{\ell=0}^{\delta-1} ((e_j \equiv \ell \pmod{\delta}) \wedge (e_{j+1} \equiv \ell + K \pmod{\delta})).$$

In particular, $\alpha + [a_1^{e_1} a_2^{e_2 - e_1} \cdots a_n^{e_n - e_{n-1}}]_F \in A$ is equivalent to some quantifier-free L_{div}-formula under the assumption that $e_1 \leq \cdots \leq e_n$. So X_f is quantifier-free definable in $\mathfrak{N}$ by

$$(e_1 \leq \cdots \leq e_n) \wedge (\alpha + [a_1^{e_1} a_2^{e_2 - e_1} \cdots a_n^{e_n - e_{n-1}}]_F \in A),$$

as desired. So X is quantifier-free definable in $\mathfrak{N}$, say by $\varphi_X(\mathbf{x})$.

We now check that the formula φ_X is stable under any partition of the variables. Indeed, suppose otherwise; assume for notational convenience that we have a partition $\varphi_X(x_1, \dots, x_{n_0}; x_{n_0+1}, \dots, x_n)$ that is unstable for some n_0. So for $N \in \mathbb{N}$ there are $e_{k1}, \dots, e_{kn} \in \mathbb{N}$ for $k < N$ such that

$$\mathfrak{N} \models \varphi_X(e_{k1}, \dots, e_{kn_0}; e_{\ell(n_0+1)}, \dots, e_{\ell n}) \iff k \leq \ell.$$

But by definition of X, and since $X = \varphi_X(\mathfrak{N})$, this means that

$$(\alpha + [b_1^{e_{k1}}]_F + \cdots + [b_{n_0}^{e_{kn_0}}]_F) + ([b_{n_0+1}^{e_{\ell(n_0+1)}}]_F + \cdots + [b_n^{e_{\ell n}}]_F) \in A \iff k \leq \ell,$$

contradicting our assumption that A is stable in Γ. So φ_X is stable under any partition of the variables.

So Proposition 4.4 applies, and thus φ_X can be taken to be a quantifier-free L_{div}-formula. Fix $a \in \Gamma$ such that $F^{\mathbb{N}}a$ is infinite. Note that such a must exist: if λ is any length function for any power of F, then whenever $a \in \Gamma$ is such that $F^{\mathbb{N}}a$ is finite, we must have that a fails the logarithmic property, and thus that a lies in the finite exceptional set of λ. So since Γ is infinite there must exist $a \in \Gamma$ with $F^{\mathbb{N}}a$ infinite.

Consider the map $\Phi\colon \mathbb{N} \to \Gamma$ given by $e \mapsto [a^e]_F$. Note that Φ embeds $\mathbb{N}$ into Γ as the 0-definable set $K(a; F)$. Indeed, if $[a^{e_1}]_F = [a^{e_2}]_F$ for some $e_1, e_2 \in \mathbb{N}$, then applying $F - 1$ yields that $F^{e_1}a - a = F^{e_2}a - a$, and hence since $F^{\mathbb{N}}a$ is infinite that $e_1 = e_2$.

We will show that Φ induces a 0-definable interpretation of $\mathfrak{N}_0 := (\mathbb{N}, 0, S, (\delta\mathbb{N})_{\delta \geq 2})$ in $(\Gamma, \mathcal{F})$. From this, and the fact that X is definable in $\mathfrak{N}_0$, it will follow that A_i' is indeed definable in $(\Gamma, \mathcal{F})$.

We must show that $\{\, \Phi(0) \,\} = \{\, 0 \,\}$ is 0-definable in $(\Gamma, \mathcal{F})$; this is simply because $\{\, 0 \,\} = K(0; F)$. We must also show that $\left\{ \begin{pmatrix} \Phi(e) \\ \Phi(Se) \end{pmatrix} : e \in \mathbb{N} \right\}$ is 0-definable; for this we note that it is precisely $\begin{pmatrix} 0 \\ a \end{pmatrix} + K\left(\begin{pmatrix} a \\ Fa \end{pmatrix}; F \right)$. Finally, we must show that $\{\, \Phi(e) : e \in \delta\mathbb{N} \,\}$ is 0-definable; but this is $K([a^\delta]_F; F^\delta)$. Since all of these are F-sets, we get that all are 0-definable in $(\Gamma, \mathcal{F})$.

So Φ does indeed induce a 0-definable interpretation of $\mathfrak{N}_0$ in $(\Gamma, \mathcal{F})$. Thus

$$\Phi(X) = \left\{ \begin{pmatrix} [a^{e_1}]_F \\ \vdots \\ [a^{e_n}]_F \end{pmatrix} : e_1, \dots, e_n \in \mathbb{N}, \alpha + [b_1^{e_1}]_F + \cdots + [b_n^{e_n}]_F \in A \right\}$$

is 0-definable in $(\Gamma, \mathcal{F})$. Moreover, for each $1 \leq j \leq n$ the map $g_j \colon K(a; F) \to K(b_j; F)$ given by $[a^e]_F \mapsto [b_j^e]_F$ is definable in $(\Gamma, \mathcal{F})$: its graph is $K\left(\begin{pmatrix} a \\ b_j \end{pmatrix}; F\right)$. Finally, recall that addition is definable in $(\Gamma, \mathcal{F})$, since its graph is an F-invariant subgroup of Γ^3. Putting these together, we find that

$$\begin{aligned} A_i' &= A \cap (\alpha + K(b_1; F) + \cdots + K(b_n; F)) \\ &= \{ \alpha + [b_1^{e_1}]_F + \cdots + [b_n^{e_n}]_F : (e_1, \dots, e_n) \in X \} \\ &= \left\{ \alpha + [b_1^{e_1}]_F + \cdots + [b_n^{e_n}]_F : \begin{pmatrix} [a^{e_1}]_F \\ \vdots \\ [a^{e_n}]_F \end{pmatrix} \in \Phi(X) \right\} \end{aligned}$$

is definable in $(\Gamma, \mathcal{F})$, namely by the formula

$$\exists y_1 \cdots \exists y_n \left(\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \Phi(X) \wedge (x = \alpha + g_1(y_1) + \cdots + g_n(y_n)) \right).$$

So $A = \bigcup_{i=1}^{M} A_i'$ is definable in $(\Gamma, \mathcal{F})$, as desired. $\square$ Proposition 4.7

In the finitely generated torsion-free case, we can do better:

Proposition 4.8. *Suppose $\Gamma = (\mathbb{Z}^m, +)$ for some m. Then:*

1. $\mathrm{Th}(\Gamma, \mathcal{F})$ *is stable.*
2. *If $A \subseteq \Gamma$ is F-sparse and stable in Γ then A is a Boolean combination of elementary F-sets.*

We will need the following result of Moosa and Scanlon:

Fact 4.9 ([18, Proposition 3.9])**.** *Suppose that $\mathbb{Z}[F]$ is a finite extension of $\mathbb{Z}$, that Γ is a finitely generated $\mathbb{Z}[F]$-module, and that*

$$\bigcap_{i\in\mathbb{N}} F^i\mathbb{Z}[F] = \{\,0\,\}.$$

If $A, B \subseteq \Gamma$ are F-sets then so is $A \cap B$.

Proof of Proposition 4.8. We will appeal to Fact 4.3; we verify that the hypotheses hold. It is clear that $\mathbb{Z}[F]$ extends $\mathbb{Z}$, and since Γ is a finitely generated group it is certainly a finitely generated $\mathbb{Z}[F]$-module. To see that $\mathbb{Z}[F]$ is a finite extension of $\mathbb{Z}$, we note that F is given by some matrix T with integer entries; so if p_T is the characteristic polynomial of T then p_T is monic, $p_T \in \mathbb{Z}[t]$, and by the Cayley-Hamilton theorem $p_T(T) = 0$. So $p_T(F) = 0$, and F is integral over $\mathbb{Z}$; so $\mathbb{Z}[F]$ is a finite extension of $\mathbb{Z}$. Finally, we must show that

$$\bigcap_{i\in\mathbb{N}} F^i\mathbb{Z}[F] = \{\,0\,\}.$$

This will use our standing assumption that Γ admits an F^r-spanning set for some $r > 0$. Indeed, by Theorem 2.43 there is a length function λ for (Γ, F^r) for some $r > 0$, say with associated constants C, D, E. By Lemma 2.50 we may assume the exceptional set contains only torsion elements, and is thus $\{\,0\,\}$; so the logarithmic property applies to all non-zero elements of Γ. Suppose $G \in F^i\mathbb{Z}[F]$ for all $i \in \mathbb{N}$; so for all i we get that $G \in F^{ir}\mathbb{Z}[F]$, and there is $G_i \in \mathbb{Z}[F]$ such that $G = F^{ir}G_i$. Suppose for contradiction we had $a \in \Gamma$ such that $Ga \neq 0$. Then $G_i a \neq 0$, so by the logarithmic property we get that $\lambda(Ga) \geq \lambda(G_i a) + iC - E$. But this grows without bound as i grows, a contradiction. So $Ga = 0$ for all $a \in \Gamma$, and thus $G = 0$.

So Fact 4.3 applies, and thus $\mathrm{Th}(\Gamma, \mathcal{F})$ is stable. Suppose now that $A \subseteq \Gamma$ is F-sparse and stable in Γ. Proposition 4.7 yields that A is definable in $(\Gamma, \mathcal{F})$, which admits quantifier elimination by Fact 4.3. So A is a Boolean combination of F-sets; using disjunctive normal form, we can write A as a finite union of sets of the form

$$B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots \cup C_\ell)$$

for F-sets $B_1, \ldots, B_k, C_1, \ldots, C_\ell$. We wish to show that the B_i, C_j can be taken to be groupless.

Since A is F-sparse, Lemma 4.6 yields that there is some s for which A can be written as a finite union of sets of the form

$$\alpha + \{\,[b_1^{e_1}]_{F^s} + \cdots + [b_n^{e_n}]_{F^s} : e_1 \leq \cdots \leq e_n\,\}.$$

In particular, A is contained in a finite union of sets of the form

$$\alpha + K(b_1; F^s) + \cdots + K(b_n; F^s),$$

and is thus contained in a groupless F-set, say $\widetilde{A}$.

Now, if $k > 0$ then

$$\begin{aligned} B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots \cup C_\ell) &= \widetilde{A} \cap (B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots \cup C_\ell)) \\ &= \widetilde{A} \cap B_1 \cap \cdots \cap B_k \setminus (\widetilde{A} \cap (C_1 \cup \cdots \cup C_\ell)) \\ &= (\widetilde{A} \cap B_1) \cap \cdots \cap (\widetilde{A} \cap B_k) \setminus ((\widetilde{A} \cap C_1) \cup \cdots \cup (\widetilde{A} \cap C_\ell)). \end{aligned}$$

By Fact 4.9 we get that each $\widetilde{A} \cap B_i$ is an F-set. Fix i, and write $\widetilde{A} \cap B_i$ as a finite union of sets of the form $D + H$ for some elementary F-set D and some F-invariant $H \leq \Gamma$. If H is finite then $D + H = \bigcup_{h \in H}(h + D)$, and we can replace $D + H$ in the union with a union of elementary F-sets. If H is infinite, then $\widetilde{A} \supseteq \widetilde{A} \cap B_i \supseteq D + H$ contains a coset of the infinite F-invariant subgroup H, and thus isn't F-sparse by Proposition 3.22. But we showed in Corollary 3.11 that groupless F-sets are F-sparse, and $\widetilde{A}$ is a groupless F-set, a contradiction.

So $\widetilde{A} \cap B_i$ can be written as a union of elementary F-sets; similarly with $\widetilde{A} \cap C_j$. So if $k > 0$ then we can write $B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots \cup C_\ell)$ as a Boolean combination of elementary F-sets. If on the other hand $k = 0$, then

$$B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots \cup C_\ell) = \Gamma \setminus (C_1 \cup \cdots \cup C_\ell) = \widetilde{A} \setminus ((\widetilde{A} \cap C_1) \cup \cdots \cup (\widetilde{A} \cap C_\ell)),$$

and by a similar argument we can deduce that each $\widetilde{A} \cap C_j$ is a finite union of elementary F-sets. So in this case too we can write $B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots \cup C_\ell)$ as a Boolean combination of elementary F-sets. Thus, since A is a finite union of sets of the form $B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots \cup C_\ell)$ for F-sets B_i, C_j, we get that A can be written as a Boolean combination of elementary F-sets. □ Proposition 4.8

A general finitely generated abelian group is obtained by allowing finitely many torsion elements; we now study how this affects automaticity.

Lemma 4.10. *Suppose Γ is finitely generated; write $\Gamma = \Gamma_0 \times H$ for some torsion-free $\Gamma_0 \leq \Gamma$ and some finite H. Then Γ_0 is F-automatic. Moreover, if $a \in \Gamma$ then $\pi_H([a^i]_F)$ is ultimately periodic in i, where $\pi_H \colon \Gamma \to H$ is the projection.*

The primary difficulty here is that most of our previous results on automaticity and spanning sets in subgroups required that the subgroup be F-invariant, which need not hold of Γ_0.

Proof. Suppose first that there is an F-spanning set Σ. We will produce a DFA $(\Sigma, Q, q_0, \Omega, \delta)$ for which there exists a map $f\colon Q \to H$ such that $f(\delta(q_0, \sigma)) = \pi_H([\sigma]_F)$ for $\sigma \in \Sigma^*$; that is to say, $\pi_H([\sigma]_F)$ can be determined from $\delta(q_0, \sigma)$.

Fix generators $\gamma_1, \dots, \gamma_m$ for Γ, and write $F\gamma_j = a_{1j}\gamma_1 + \cdots + a_{mj}\gamma_m$ for some $a_{ij} \in \mathbb{Z}$. Observe that if $b = b_1\gamma_1 + \cdots + b_m\gamma_m \in \Gamma$ then

$$\pi_H(b) = b_1\pi_H(\gamma_1) + \cdots + b_m\pi_H(\gamma_m) = (b_1 + |H|\mathbb{Z})\pi_H(\gamma_1) + \cdots + (b_m + |H|\mathbb{Z})\pi_H(\gamma_m)$$

since $|H| \cdot H = 0$. In particular, to know $\pi_H(b)$ it suffices to know $b_1 + |H|\mathbb{Z}, \dots, b_m + |H|\mathbb{Z}$. So if we want to determine $\pi_H(F^k b)$ then it suffices to determine

$$\begin{aligned}
&\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix}^k \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} + |H|\mathbb{Z}^m \\
&= \underbrace{\begin{pmatrix} a_{11} + |H|\mathbb{Z} & \cdots & a_{1m} + |H|\mathbb{Z} \\ \vdots & \ddots & \vdots \\ a_{m1} + |H|\mathbb{Z} & \cdots & a_{mm} + |H|\mathbb{Z} \end{pmatrix}}_{X}{}^k \begin{pmatrix} b_1 + |H|\mathbb{Z} \\ \vdots \\ b_m + |H|\mathbb{Z} \end{pmatrix} \in (\mathbb{Z}/|H|\mathbb{Z})^m.
\end{aligned}$$

Thus given $\sigma \in \Sigma^*$ and $b = b_1\gamma_1 + \cdots + b_m\gamma_m \in \Sigma$, we can determine $\pi_H([\sigma b]_F) = \pi_H([\sigma]_F) + \pi_H(F^{|\sigma|}b)$ from $\pi_H([\sigma]_F)$ and $X^{|\sigma|}$. But $\pi_H([\sigma]_F) \in H$ and $X^{|\sigma|} \in M_m(\mathbb{Z}/|H|\mathbb{Z})$ both take on one of finitely many possible values, and can thus be tracked by a DFA.

Indeed, let

$$\begin{aligned}
Q &= H \times M_m(\mathbb{Z}/|H|\mathbb{Z}) \\
q_0 &= (0, I) \\
\Omega &= \{\, (0, Y) : Y \in M_m(\mathbb{Z}/|H|\mathbb{Z}) \,\}.
\end{aligned}$$

For the transition map δ, suppose we are given $(h, Y) \in Q$ and $b = b_1\gamma_1 + \cdots + b_m\gamma_m \in \Sigma$. Let c_i be the i^{th} component of $Y \begin{pmatrix} b_1 + |H|\mathbb{Z} \\ \vdots \\ b_m + |H|\mathbb{Z} \end{pmatrix}$; we then let

$$\delta((h, Y), b) = (h + c_1\pi_H(\gamma_1) + \cdots + c_m\pi_H(\gamma_m), YX).$$

We show by induction on $|\sigma|$ that if $\sigma \in \Sigma^*$ then $\delta(q_0, \sigma) = (\pi_H([\sigma]_F), X^{|\sigma|})$. The base case is immediate. For the induction step, suppose we are given $\sigma \in \Sigma^*$ such that

$\delta(q_0, \sigma) = (\pi_H([\sigma]_F), X^{|\sigma|})$, and suppose $b = b_1\gamma_1 + \cdots + b_m\gamma_m \in \Sigma$. Let c_i be the i^{th} component of $X^{|\sigma|}\begin{pmatrix} b_1 + |H|\mathbb{Z} \\ \vdots \\ b_m + |H|\mathbb{Z}\end{pmatrix}$; so $F^{|\sigma|}b \in (c_1 + |H|\mathbb{Z})\gamma_1 + \cdots + (c_m + |H|\mathbb{Z})\gamma_m$, and thus $\pi_H(F^{|\sigma|}b) = c_1\pi_H(\gamma_1) + \cdots + c_m\pi_H(\gamma_m)$. So

$$\begin{aligned}
\delta(q_0, \sigma b) &= \delta((\pi_H([\sigma]_F), X^{|\sigma|}), b) \\
&= (\pi_H([\sigma]_F) + c_1\pi_H(\gamma_1) + \cdots + c_m\pi_H(\gamma_m), X^{|\sigma|+1}) \\
&= (\pi_H([\sigma]_F) + \pi_H(F^{|\sigma|}b), X^{|\sigma|+1}) \\
&= (\pi_H([\sigma b]_F), X^{|\sigma b|}),
\end{aligned}$$

as desired.

Thus if we take $f\colon Q \to H$ to be $(h, Y) \mapsto h$, we get that $f(\delta(q_0, \sigma)) = \pi_H([\sigma]_F)$ for $\sigma \in \Sigma^*$; so we have constructed our desired DFA. Moreover examining Ω we see that this DFA accepts Γ_0; so if Γ admits an F-spanning set Σ then Γ_0 is (F, Σ)-automatic.

We now do the general case. By assumption there is $r > 0$ for which Γ admits an F^r-spanning set Σ; then by the above applied to F^r we get that Γ_0 is (F^r, Σ)-automatic, and thus F-automatic in Γ. Moreover if $a \in \Gamma$ then by Proposition 2.15 we may assume that $a_i := [a^i]_F \in \Sigma$ for $i \leq r$. Let $(\Sigma, Q, q_0, \Omega, \delta)$ be the DFA constructed above. Then since Q is finite we get that $\delta(q_0, a_r^j)$ is eventually periodic in j. So for each $i < r$ we get that $\delta(q_0, a_r^j a_i)$ is eventually periodic in j; hence since we can determine $\pi_H([a_r^j a_i]_{F^r})$ from $\delta(q_0, a_r^j a_i)$ (using the function f above), it follows that

$$\pi_H([a_r^j a_i]_{F^r}) = \pi_H([a^r]_F + F^r[a^r]_F + \cdots + F^{(j-1)r}[a^r]_F + F^{jr}[a^i]_F) = \pi_H([a^{jr+i}]_F)$$

is also eventually periodic in j. So, taking the LCMs of the periods of the $\pi_H([a^{jr+i}]_F)$ for $i < r$, we see that $\pi_H([a^i]_F)$ is ultimately periodic in i. $\square$ Lemma 4.10

We can now prove our main theorem.

Theorem 4.11. *Suppose Γ is finitely generated and $A \subseteq \Gamma$ is F-sparse. The following are equivalent:*

1. *$(\Gamma, +, A)$ is stable.*
2. *A is stable in $(\Gamma, +)$.*
3. *A is a Boolean combination of elementary F-sets.*

Proof. By the fundamental theorem of finitely generated abelian groups, we can write $\Gamma = \Gamma_0 \times H$ with H a finite group and Γ torsion-free and finitely generated. Let π_0, π_H be the projections $\Gamma \to \Gamma_0, \Gamma \to H$, respectively. Let $F_0\colon \Gamma_0 \to \Gamma_0$ be $(\pi_0 \circ F)\restriction\Gamma_0$; so by Lemma 2.51 F_0 is injective and Γ_0 admits a spanning set for some power of F_0. Our general strategy will be to apply Proposition 4.8 to (Γ_0, F_0), and then use that to deduce the result on (Γ, F).

(1) $\Longrightarrow$ (2) This follows from the definition of stable theories.

(2) $\Longrightarrow$ (3) Suppose A is stable in Γ. For $h \in H$ let $A_h := \{ a \in \Gamma_0 : a + h \in A \}$. We will show that the A_h are F_0-sparse and stable in $(\Gamma_0, +)$, and hence that Proposition 4.8 applies.

To see that A_h is stable in $(\Gamma_0, +)$, suppose we have an N-ladder $(a_0, \ldots, a_{N-1}; b_0, \ldots, b_{N-1})$ for $x + y \in A_h$ in Γ_0; so $a_i + b_j \in A_h$ if and only if $i \leq j$. Then $a_i + (b_j + h) \in A$ if and only if $i \leq j$, and thus $(a_0, \ldots, a_{N-1}; b_0 + h, \ldots, b_{N-1} + h)$ is an N-ladder for $x + y \in A$ in Γ. So since A is stable in Γ we get that A_h is stable in $(\Gamma_0, +)$.

To see that A_h is F_0-sparse, note by Lemma 4.10 that Γ_0 is F-automatic in Γ. Closure of F-sparsity under translation (Remark 3.9) yields that $A - h$ is F-sparse. Hence by closure properties of F-sparsity (Proposition 3.6 (3)) we get that $A_h = \Gamma_0 \cap (A - h)$ is F-sparse; say there is an F^r-spanning set Σ such that A_h is (F^r, Σ)-sparse. Then by Fact 3.3 we get that A_h is a finite union of sets of the form $[u_0 v_1^* u_1 \cdots v_n^* u_n]_{F^r}$ for $u_0, \ldots, u_n, v_1, \ldots, v_n \in \Sigma^*$. So $A_h = \pi_0(A_h)$ is a finite union of sets of the form $\pi_0([u_0 v_1^* u_1 \cdots v_n^* u_n]_{F^r}) = [\pi_0(u_0)(\pi_0(v_1))^* \pi_0(u_1) \cdots (\pi_0(v_n))^* \pi_0(u_n)]_{F_0^r}$ by Lemma 2.51 (2) (where we let $\pi_0(s_0 \cdots s_{n-1}) = (\pi_0(s_0)) \cdots (\pi_0(s_{n-1}))$). But Lemma 2.51 (3) yields that $\pi_0(\Sigma)$ is an F_0^r-spanning set for Γ_0; so A_h is $(F_0^r, \pi_0(\Sigma))$-sparse, and is F_0-sparse.

Thus Proposition 4.8 applies, and we get that A_h is a Boolean combination of elementary F_0-sets. So to show that A_h is a Boolean combination of elementary F-sets, it will suffice to show that given $a \in \Gamma_0$ and $s > 0$ we can write $K(a; F_0^s)$ as a finite union of elementary F-sets. Replacing F with a power thereof, it suffices to check the case $s = 1$. Recall from Lemma 2.51 (2) that

$$\pi_0([a^i]_F) = [(\pi_0(a))^i]_{F_0} = [a^i]_{F_0};$$

so if we let $h_i = \pi_H([a^i]_F)$ then $[a^i]_F - h_i = [a^i]_{F_0}$. Moreover we showed in Lemma 4.10 that h_i is ultimately periodic in i; say with M, μ such that $h_i = h_{i+\mu}$ for $i \geq M$.

Then

$$\begin{aligned}
K(a;F_0) &= \{\, [a^i]_{F_0} : i \in \mathbb{N} \,\} \\
&= \bigcup_{j<M} \{\, [a^j]_{F_0} \,\} \cup \bigcup_{j=M}^{M+\mu-1} \{\, [a^{j+i\mu}]_{F_0} : i \in \mathbb{N} \,\} \\
&= \bigcup_{j<M} \{\, [a^j]_{F_0} \,\} \cup \bigcup_{j=M}^{M+\mu-1} (\{\, [a^{j+i\mu}]_{F} : i \in \mathbb{N} \,\} - h_j),
\end{aligned}$$

and for $M \leq j < M+\mu$ we have

$$\begin{aligned}
\{\, [a^{j+i\mu}]_F : i \in \mathbb{N} \,\} &= [a^j]_F + \{\, F^j[a^{i\mu}]_F : i \in \mathbb{N} \,\} \\
&= [a^j]_F + \{\, F^j([a^\mu]_F + F^{|\mu|}[a^\mu]_F + \cdots + F^{(i-1)|\mu|}[a^\mu]_F) : i \in \mathbb{N} \,\} \\
&= [a^j]_F + K(F^j[a^\mu]_F; F^\mu).
\end{aligned}$$

So

$$K(a;F_0) = \bigcup_{j<M} \{\, [a^j]_{F_0} \,\} \cup \bigcup_{j=M}^{M+\mu-1} (K(F^j[a^\mu]_F; F^\mu) + [a^j]_F - h_j)$$

is a finite union of elementary F-sets (noting that $\{\, [a^j]_{F_0} \,\} = [a^j]_{F_0} + K(0;F)$ is an elementary F-set). From this, and the fact that sum of two elementary F-sets is by definition again an elementary F-set, it follows that an elementary F_0-set is a finite union of elementary F-sets, and hence that each A_h is a Boolean combination of elementary F-sets. Thus

$$A = \bigcup_{h \in H} (A_h + h)$$

is a Boolean combination of elementary F-sets.

(3) $\Longrightarrow$ (1) From Proposition 4.8 we know that the theory of the F_0-structure $(\Gamma_0, \mathcal{F}_0)$ on Γ_0 is stable. We will show that $(\Gamma, +, A)$ is definably interpretable in $(\Gamma_0, \mathcal{F}_0)$; since stability of a theory is defined as the absence of an unstable definable set, this implies that $\mathrm{Th}(\Gamma, +, A)$ is stable.

For each $h \in H$ fix some $b_h \in \Gamma_0$ such that $b_h \neq b_{h'}$ for $h \neq h'$. Let $\Phi \colon \Gamma \to \Gamma_0$ be $a + h \mapsto \begin{pmatrix} a \\ b_h \end{pmatrix}$; we show that Φ induces a definable interpretation of $(\Gamma, +, A)$ in $(\Gamma_0, \mathcal{F}_0)$. It is clear that Φ is well-defined and injective, and since H is finite we get that

$$\Phi(\Gamma) = \left\{ \begin{pmatrix} a \\ b_h \end{pmatrix} : a \in \Gamma_0, h \in H \right\} \subseteq \Gamma_0^2$$

is definable in $(\Gamma_0, \mathcal{F}_0)$. We must show that addition in Γ corresponds to a definable set in $(\Gamma_0, \mathcal{F}_0)$; for this we note that

$$\begin{aligned}(a_1+h_1)+(a_2+h_2)=a_3+h_3 &\iff (a_1+a_2=a_3)\wedge(h_1+h_2=h_3)\\ &\iff (a_1+a_2=a_3)\wedge \begin{pmatrix} b_{h_1}\\ b_{h_2}\\ b_{h_3}\end{pmatrix} \in \underbrace{\left\{ \begin{pmatrix} b_h\\ b_{h'}\\ b_{h+h'}\end{pmatrix} : h,h'\in H \right\}}_{H_+}\end{aligned}$$

(and H_+ is definable in $(\Gamma_0, \mathcal{F}_0)$ since H, and hence H_+, is finite).

It remains to show that $\Phi(A)$ is definable in $(\Gamma_0, \mathcal{F}_0)$. We have already seen that addition in Γ is definable in $(\Gamma_0, \mathcal{F}_0)$ under this interpretation; so, since A is a Boolean combination of translates of sums of sets of the form $K(a; F^r)$, it suffices to show that $\Phi(K(a; F^r))$ is definable in $(\Gamma_0, \mathcal{F}_0)$ for $a \in \Gamma$ and $r > 0$. Replacing F with F^r, it suffices to check the case $r = 1$. Fix $a \in \Gamma$; we will show for $h \in H$ that

$$B_h := \{\, c \in \Gamma_0 : c + h \in K(a; F) \,\}$$

is definable; this will suffice, since

$$\Phi(K(a;F)) = \bigcup_{h\in H} \left\{ \begin{pmatrix} c\\ b_h\end{pmatrix} : c \in B_h \right\}.$$

Let $a_0 = \pi_0(a)$, and as before let $h_i = \pi_H([a^i]_F)$ for $i \in \mathbb{N}$. Again Lemma 4.10 yields that h_i is ultimately periodic in i; say with M, μ such that $h_i = h_{i+\mu}$ for $i \geq M$. Then using Lemma 2.51 (2) we find that

$$\begin{aligned} B_h &= ([a^*]_F - h) \cap \Gamma_0\\ &= \{\, [a^i]_F - h : i \in \mathbb{N}, h_i = h \,\}\\ &= \{\, \pi_0([a^i]_F) : i \in \mathbb{N}, h_i = h \,\}\\ &= \{\, [a_0^i]_{F_0} : i \in \mathbb{N}, h_i = h \,\}\\ &= \{\, [a_0^i]_{F_0} : i < M, h_i = h \,\} \cup \bigcup_{\substack{M \leq j < M+\mu\\ h_j = h}} \{\, [a_0^{j+i\mu}]_{F_0} : i \in \mathbb{N} \,\}\\ &= \{\, [a_0^i]_{F_0} : i < M, h_i = h \,\} \cup \bigcup_{\substack{M \leq j < M+\mu\\ h_j = h}} ([a_0^j]_{F_0} + K(F_0^j[a_0^\mu]_{F_0}; F_0^\mu)) \end{aligned}$$

(where the argument that $\{\, [a_0^{j+i\mu}]_{F_0} : i \in \mathbb{N} \,\} = [a_0^j]_{F_0} + K(F_0^j[a_0^\mu]_{F_0}; F_0^\mu)$ is as in the proof of (2) $\implies$ (3)). So B_h is definable in $(\Gamma_0, \mathcal{F}_0)$; so Φ indeed induces a definable interpretation of $(\Gamma, +, A)$ in $(\Gamma_0, \mathcal{F}_0)$, and $\mathrm{Th}(\Gamma, +, A)$ is stable. $\square$ Theorem 4.11

We can adapt Theorem 4.11 to give a characterization of the stable F-sparse sets among all subsets of Γ, rather than among the F-sparse sets.

Corollary 4.12. *Suppose Γ is finitely generated and $A \subseteq \Gamma$. The following are equivalent:*

1. *A is F-sparse and stable in Γ.*
2. *A is a finite union of sets of the form*

$$B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots \cup C_\ell)$$

for B_i, C_j elementary F-sets, $k > 0$, and $\ell \geq 0$.

Proof.

(1) $\Longrightarrow$ (2) Suppose A is F-sparse and stable in Γ. By Theorem 4.11 we get that A is a Boolean combination of elementary F-sets, and hence by disjunctive normal form is a finite union of sets of the form

$$B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots C_\ell)$$

for B_i, C_j elementary F-sets and $k, \ell \geq 0$. Suppose for contradiction that some such $k = 0$; so A contains $\Gamma \setminus (C_1 \cup \cdots \cup C_\ell)$. By Corollary 3.11 we get that each C_j is F-sparse, and hence by closure under union (Proposition 3.6) that $\Gamma = A \cup C_1 \cup \cdots \cup C_\ell$ is F-sparse, contradicting Proposition 3.22. So all such $k > 0$, and A takes the desired form.

(2) $\Longrightarrow$ (1) Suppose we are given such a set $B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots \cup C_\ell)$ (in particular with $k > 0$); then $B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots \cup C_\ell) \subseteq B_1$. Moreover B_1 is a groupless F-set, and is thus F-sparse by Corollary 3.11. Furthermore $B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots \cup C_\ell)$ is a Boolean combination of F-sets, and is thus F-automatic by Theorem 2.54 and Corollary 2.29. So by Proposition 3.6 (3) we get that $B_1 \cap \cdots \cap B_k \setminus (C_1 \cup \cdots \cup C_\ell)$ is F-sparse. But A is a finite union of such sets; hence Proposition 3.6 (2) yields that A is F-sparse. That A is stable in Γ then follows from Theorem 4.11. □ Corollary 4.12

4.3 Bibliographical notes

Stability is a standard concept; see for instance [23, Chapter 8]. The definition of stability in a group is adapted from a similar definition in [9]. The term "ladder" is from [15]. As noted in-text, Facts 4.3 and 4.9 are Theorem A and Proposition 3.9 of [18], respectively. Their results require that F not be a zero divisor in $\mathbb{Z}[F]$, which we did not explicitly demand; however, in our context this follows from injectivity of F. Proposition 4.4 is presumably not new, given how elementary it is, but I was unable to find a reference for it in the literature. All other results in this section are original.

Chapter 5

Stable automatic sets in $(\mathbb{Z}, +)$

We now specialize back to the case $\Gamma = (\mathbb{Z}, +)$ and F is given by multiplication by d for some $d \geq 2$. In this context, we abbreviate F-automaticity, F-sparsity, F-sets, etc. as d-automaticity, d-sparsity, d-sets, etc. In principle, this introduces a conflict: we already had a definition of d-automaticity for the integers in Section 2.2. However, we saw in Corollary 2.30 that it coincides with F-automaticity, so there is no ambiguity.

We ask which d-automatic $A \subseteq \mathbb{Z}$ are stable in the group $(\mathbb{Z}, +)$ (which we abbreviate to "stable in $\mathbb{Z}$"). A result of stable group theory lets us reduce to the case where A is not *generic*: that is, to the case where no finite union of translates of A covers $\mathbb{Z}$. If A is d-sparse, then our question is answered by Theorem 4.11; we thus consider the case where A is neither generic nor d-sparse. We will show in Theorem 5.3 that in fact such A are never stable in $\mathbb{Z}$. This leads to our characterization of the d-automatic $A \subseteq \mathbb{Z}$ that are stable in $\mathbb{Z}$:

Theorem. *If $A \subseteq \mathbb{Z}$ then A is d-automatic and stable in $(\mathbb{Z}, +)$ if and only if A is a Boolean combination of elementary d-sets and cosets of subgroups of $\mathbb{Z}$.*

This appears as part of our full characterization, Theorem 5.17 below.

Recall from Example 2.13 that there is a natural d-spanning set for $(\mathbb{Z}, +)$, namely $\Sigma_d^{\pm} := \{ -d+1, \ldots, d-1 \}$. We will often use the fact that $a \in \mathbb{Z}$ is represented by a word over $\Sigma_d^{\pm}$ of length N if and only if $|a| < d^N$.

5.1 The non-generic case

Definition 5.1. We say $A \subseteq \mathbb{Z}$ is *generic* if some finite union of translates of A covers $\mathbb{Z}$.

Note that $A \subseteq \mathbb{Z}$ is not generic if and only if it has arbitrarily large gaps; that is, arbitrarily large runs of integers that are omitted from A. Indeed, if A is generic, say with $a_1, \ldots, a_n \in \mathbb{Z}$ such that $\mathbb{Z} = (a_1 + A) \cup \cdots \cup (a_n + A)$, then given $b \in \mathbb{Z}$ there is $1 \leq i \leq n$ and $a \in A$ such that $b = a_i + a$; so $|b - a| = |a_i| \leq \max(|a_1|, \ldots, |a_n|)$. Thus any gap in A has size at most $2\max(|a_1|, \ldots, |a_n|)$. Conversely if there is N such that any gap in A has size at most N, then $\mathbb{Z} \subseteq A \cup (1 + A) \cup \cdots \cup (N + A)$, and A is generic.

Genericity should be thought of as a "largeness" condition, and indeed generic sets are never sparse. Suppose otherwise; suppose we had d-sparse $C \subseteq \mathbb{Z}$ and $a_1, \ldots, a_n \in \mathbb{Z}$ such that $\mathbb{Z} = (a_1 + C) \cup \cdots \cup (a_n + C)$. Then each $a_i + C$ is d-sparse by Remark 3.9, and hence $\mathbb{Z}$ itself is d-sparse by Proposition 3.6 (2); but this contradicts Proposition 3.22.

Some examples and non-examples of generic sets:

Example 5.2.

1. $2\mathbb{Z}$ is generic, since $2\mathbb{Z} \cup (1 + 2\mathbb{Z}) = \mathbb{Z}$. More generally any coset of any non-trivial subgroup of $\mathbb{Z}$ is generic.

2. $\mathbb{N}$ is not generic, since for any $a_1, \ldots, a_n \in \mathbb{Z}$ we have that $(a_1 + \mathbb{N}) \cup \cdots \cup (a_n + \mathbb{N})$ doesn't contain $\min(a_1, \ldots, a_n) - 1$.

3. Recall our alphabet $\Sigma_d := \{ 0, \ldots, d-1 \} \subseteq \Sigma_d^{\pm}$. Let $A = [\Sigma_d^* 1]_d \cup [(-\Sigma_d)^*(-1)]_d$ be the set of naturals whose most significant digit base d is a 1, together with the negatives of such. Then A is not generic if $d > 2$, since if $k \in \mathbb{N}$ then $\{ [\sigma 2]_d : \sigma \in \Sigma_d^{(k)} \}$ is a gap in A of size d^k.

4. Let $B = \{ [\sigma]_d : \sigma \in \Sigma_d^*, |\sigma| \in 2\mathbb{N} + 1, \sigma \text{ has no trailing zeroes} \}$. Then $B \cup -B$ is not generic, since if $k \in \mathbb{N}$ then $\{ [\sigma 1]_d : \sigma \in \Sigma_d^{(2k)} \}$ is a gap in $B \cup -B$ of size d^{2k}.

One can check that all of the above examples are d-automatic, and that none is d-sparse; the latter can be proven by showing that they aren't d-meagre using the length function associated with the spanning set $\Sigma_d^{\pm}$. So the non-generic d-automatic sets strictly include the d-sparse sets. It turns out, however, that passing from d-sparse sets to non-generic d-automatic sets adds no new stable sets:

Theorem 5.3. *Suppose $A \subseteq \mathbb{Z}$ is d-automatic and is not generic. If A is stable in $\mathbb{Z}$ then A is d-sparse.*

It will be convenient to work in $\mathbb{N}$, rather than $\mathbb{Z}$. The main advantage is that we can work with representations over $\Sigma_d = \{ 0, \dots, d-1 \}$, rather than $\Sigma_d^{\pm}$, and representations are essentially unique over the former: if $\sigma, \tau \in \Sigma_d^*$ have the same length and $[\sigma]_d = [\tau]_d$, we must have that $\sigma = \tau$. (Note that this fails in $(\Sigma_d^{\pm})^*$: for instance, we have $[(d-1)0]_d = [(-1)1]_d$.) Note that as with $\Sigma_d^{\pm}$ we get that $a \in \mathbb{N}$ has a representation in Σ_d^* of length N if and only if $a < d^N$.

A set $A \subseteq \mathbb{N}$ is d-automatic in $(\mathbb{Z}, +)$ if and only if $\{ \sigma \in \Sigma_d^* : [\sigma]_d \in A \}$ is regular. (This is in fact the classical definition of d-automaticity for subsets of $\mathbb{N}$, which we saw in Section 2.2.) Indeed, if A is d-automatic then since $\Sigma_d^{\pm}$ is d-spanning, we get that $\{ \sigma \in (\Sigma_d^{\pm})^* : [\sigma]_d \in A \}$ is regular, and hence since regular languages are closed under intersection that

$$\{ \sigma \in \Sigma_d^* : [\sigma]_d \in A \} = \{ \sigma \in (\Sigma_d^{\pm})^* : [\sigma]_d \in A \} \cap \Sigma_d^*$$

is regular. The converse is by Proposition 2.33. Moreover, if $A \subseteq \mathbb{N}$ is d-automatic but not d-sparse in $(\mathbb{Z}, +)$, then the definition of d-sparsity yields that $\{ \sigma \in \Sigma_d^* : [\sigma]_d \in A \}$ is not sparse. So we can convert d-automaticity and d-sparsity to properties of representations over Σ_d^*.

What should stability and genericity mean for subsets of $\mathbb{N}$? For stability, we simply consider the monoid structure rather than the group structure: we say that $A \subseteq \mathbb{N}$ is *stable in* $(\mathbb{N}, +)$ if the binary relation $x + y \in A$ on $\mathbb{N}^2$ is stable. For genericity, we say that $A \subseteq \mathbb{N}$ is *generic in* $\mathbb{N}$ if there are $a_1, \dots, a_n \in \mathbb{Z}$ such that $\mathbb{N} \subseteq (a_1 + A) \cup \cdots \cup (a_n + A)$. Note that we allow negative translates of A; this is to ensure that e.g. $\mathbb{N} \setminus \{ 0 \}$ is generic in $\mathbb{N}$. As in $\mathbb{Z}$, we get that $A \subseteq \mathbb{N}$ is not generic in $\mathbb{N}$ if and only if it has arbitrarily large gaps.

We characterize genericity in $\mathbb{N}$ in terms of when $L := \{ \sigma \in \Sigma_d^* : [\sigma]_d \in A \}$ has a *forbidden suffix*: that is, some $\sigma \in \Sigma_d^*$ such that if $\tau \in \Sigma_d^*$ has σ as a suffix then $\tau \notin L$.

Proposition 5.4. *Suppose $A \subseteq \mathbb{N}$, and let $L = \{ \sigma \in \Sigma_d^* : [\sigma]_d \in A \}$. The following are equivalent:*

1. *A is not generic in $\mathbb{N}$.*
2. *There are $s > 0$ and $r \in \mathbb{N}$ such that $L \cap \Sigma_d^{(r+s\mathbb{N})}$ has a forbidden suffix.*
3. *There is $s > 0$ such that if $0 \leq r < s$ then $L \cap \Sigma_d^{(r+s\mathbb{N})}$ has a forbidden suffix.*

We will need the following fact from automata theory:

Fact 5.5 (Pumping lemma, [24, Lemma 4.1])**.** *Suppose Σ is an alphabet and $L \subseteq \Sigma^*$ is regular. There is a* pumping length $p \in \mathbb{N}$ *for L such that if $\sigma \in L$ is of length $\geq p$ then we can write $\sigma = uvw$ for $u, v, w \in \Sigma^*$ such that*

- $v \neq \varepsilon$,
- $|uv| \leq p$, *and*
- $uv^*w \subseteq L$.

Proof of Proposition 5.4.

(1) $\Longrightarrow$ (2) Suppose A is not generic in $\mathbb{N}$. Let \$ be a symbol not in Σ_d, which we will use as a separator, and consider

$$S := \{ 0^i\$\sigma : i \in \mathbb{N}, \sigma \in \Sigma_d^*, L \cap \Sigma_d^{(i)}\sigma = \emptyset \} \subseteq (\Sigma_d \cup \{ \$ \})^*.$$

That is, S is the set of $0^i\$\sigma$ such that σ is a forbidden suffix for $L \cap \Sigma_d^{(i+|\sigma|)}$. We wish to apply the pumping lemma to S; so we must show that S is regular.

We will use the fact (Corollary 2.8) that regular languages are closed under projection. Let $S_2 \subseteq ((\Sigma_d \cup \{ \$ \})^2)^*$ be the set of $\begin{pmatrix} 0^{|\tau|} \\ \tau \end{pmatrix} \begin{pmatrix} \$ \\ \$ \end{pmatrix} \begin{pmatrix} \sigma \\ 0^{|\sigma|} \end{pmatrix}$ for $\sigma, \tau \in \Sigma_d^*$ such that $\tau\sigma \in L$. So $\begin{pmatrix} 0^{|\tau|}\$\sigma \\ \tau\$0^{|\sigma|} \end{pmatrix} \in S_2$ if and only if $\tau\sigma$ witnesses that $0^{|\tau|}\$\sigma \notin S$. So if $\pi\colon ((\Sigma_d \cup \{ \$ \})^2)^* \to (\Sigma_d \cup \{ \$ \})^*$ is the projection to the first coordinate, then $S = 0^*\$\Sigma_d^* \setminus \pi(S_2)$. Hence by Corollaries 2.5 and 2.8 if S_2 is regular then so is S.

We show that S_2 is regular; fix a DFA $M = (\Sigma_d, Q, q_0, \Omega, \delta)$ recognizing L. We construct a new DFA $M' = ((\Sigma_d \cup \{ \$ \})^2, Q', q_0', \Omega', \delta')$ that will recognize S_2. Let q_{dead} be some new state, which we will use to indicate that whatever partial input M' has received precludes the input being accepted. Let $Q' = (Q \times \{ 0, 1 \}) \cup \{ q_{\text{dead}} \}$, $q_0' = (q_0, 0)$, and $\Omega' = \Omega \times \{ 1 \}$. For $q \in Q$ and $a \in \Sigma_d$, we let

$$\delta'\left((q,0), \begin{pmatrix} 0 \\ a \end{pmatrix}\right) = (\delta(q,a), 0)$$
$$\delta'\left((q,0), \begin{pmatrix} \$ \\ \$ \end{pmatrix}\right) = (q, 1)$$
$$\delta'\left((q,1), \begin{pmatrix} a \\ 0 \end{pmatrix}\right) = (\delta(q,a), 1);$$

we let $\delta'(q', \mathbf{a}) = q_{\text{dead}}$ for all other $q' \in Q'$ and $\mathbf{a} \in (\Sigma_d \cup \{ \$ \})^2$. Informally, M' verifies that its input takes the form $\begin{pmatrix} 0^{|\tau|}\$\sigma \\ \tau\$0^{|\sigma|} \end{pmatrix}$, and if so it feeds $\tau\sigma$ through M. So M' recognizes S_2, and thus S_2 is regular; so S is regular.

So the pumping lemma applies to S. We wish for some $0^p\$\sigma \in S$ where p is the pumping length of S. Since A has arbitrarily large gaps, it must have a gap I of size at least $2d^p$; so I contains two multiples of d^p, say ad^p and $(a+1)d^p$. Pick $\sigma \in \Sigma_d^*$ such that $[\sigma]_d = a$. Then σ is a forbidden suffix for $L \cap \Sigma_d^{(p+|\sigma|)}$: if $\tau \in \Sigma_d^{(p)}$ then $[\tau\sigma]_d = ad^p + [\tau]_d$ lies between ad^p and $(a+1)d^p$, and hence lies in $I \subseteq A^c$. So $0^p\$\sigma \in S$.

So by the pumping lemma we can write $0^p\$\sigma = uvw$ for strings u, v, w such that $v \neq \varepsilon$, $|uv| \leq p$, and $uv^*w \subseteq S$; so u and v consist entirely of zeroes. Let $s = |v| > 0$ and $r = p - s$. Then for $i \in \mathbb{N}$ we have $0^{r+si}\$\sigma = uv^iw \in S$, and thus σ is a forbidden suffix for $L \cap \Sigma_d^{(r+si+|\sigma|)}$. So σ is forbidden suffix for $L \cap \Sigma_d^{(r+|\sigma|+s\mathbb{N})}$.

(2) $\Longrightarrow$ (1) Suppose there are $r \in \mathbb{N}$ and $s > 0$ for which $L \cap \Sigma_d^{(r+s\mathbb{N})}$ has a forbidden suffix $\sigma \in \Sigma_d^*$. If $k \in \mathbb{N}$ satisfies $k \equiv r - |\sigma| \pmod{s}$ then $\{ \tau\sigma : \tau \in \Sigma_d^{(k)} \} \subseteq \Sigma^{(r+s\mathbb{N})}$ only contains strings that have σ as a suffix, and hence is disjoint from L. So $\{ [\tau\sigma]_d : \tau \in \Sigma_d^{(k)} \}$ is a gap in A of size d^k (since $[\cdot]_d$ is injective when restricted to $\Sigma_d^{(k)}$). Thus A has arbitrarily large gaps, and A is non-generic in $\mathbb{N}$.

(2) $\Longrightarrow$ (3) Suppose there is $r \in \mathbb{N}$ and $s > 0$ such that $L \cap \Sigma_d^{(r+s\mathbb{N})}$ has a forbidden suffix σ. Note that $0^i\sigma$ is also a forbidden suffix for $L \cap \Sigma_d^{(r+s\mathbb{N})}$, since any string that has $0^i\sigma$ as a suffix must also have σ as a suffix. Hence by prepending zeroes to σ, we may assume $|\sigma| \geq r$.

Fix $0 \leq r_0 < r$ such that $r \equiv r_0 \pmod{s}$. Then σ is a forbidden suffix for $L \cap \Sigma_d^{(r_0+s\mathbb{N})}$. Indeed, if $\tau \in \Sigma_d^{(r_0+s\mathbb{N})}$ has σ as a suffix, then since $|\sigma| \geq r$ we get that $|\tau| \geq r$, and thus that $\tau \in \Sigma_d^{(r+s\mathbb{N})}$; so, since τ has σ as a suffix, it follows that $\tau \notin L$. Thus, replacing r with r_0, we may assume $0 \leq r < s$.

Suppose $0 \leq r' < s$; fix $k \in \mathbb{N}$ such that $r' \equiv r + k \pmod{s}$. Then $\sigma 0^k$ is a forbidden suffix for $L \cap \Sigma_d^{(r'+s\mathbb{N})}$. Indeed, if $\tau \in \Sigma_d^*$ and $\tau\sigma 0^k \in L \cap \Sigma_d^{(r'+s\mathbb{N})}$ then $[\tau\sigma]_d = [\tau\sigma 0^k]_d \in A$, and thus $\tau\sigma \in L$ as well. But $|\tau\sigma| \in (r' - k + s\mathbb{N})$, and hence $|\tau\sigma| \equiv r' - k \equiv r \pmod{s}$; so $\tau\sigma \in L \cap \Sigma_d^{(r+s\mathbb{N})}$, contradicting our assumption that σ is a forbidden suffix for $L \cap \Sigma_d^{(r+s\mathbb{N})}$.

(3) $\Longrightarrow$ (2) This is clear. $\square$ Proposition 5.4

Suppose we are given d-automatic $A \subseteq \mathbb{N}$ that is neither d-sparse nor generic in $\mathbb{N}$; so our goal is to show that A is unstable in $\mathbb{N}$. Rather than exhibit ladders for the relation $x+y \in A$ on $\mathbb{N}^2$, we will instead consider an associated relation on strings. Given $K \in \mathbb{N}$, we define a partial binary operation $+_K$ on Σ_d^* as follows: given $\sigma, \tau \in \Sigma_d^*$, we set $\sigma +_K \tau$ to be the unique $\nu \in \Sigma_d^{(K)}$ such that $[\nu]_d = [\sigma]_d + [\tau]_d$, if such ν exists. Let $L = \{\, \sigma \in \Sigma_d^* : [\sigma]_d \in A \,\}$, and consider the binary relation $x +_K y \in L$ on $(\Sigma_d^*)^2$ consisting of those $\sigma, \tau \in \Sigma_d^*$ such that $\sigma +_K \tau$ is defined and lies in L. We will show that for every $N \in \mathbb{N}$ there is $K > 0$ such that there is an N-ladder for $x +_K y \in L$; we will then conclude that $x + y \in A$ is unstable in $\mathbb{N}$.

It is inconvenient for us that L need not be closed under concatenation; that is, it need not hold that $L = L^*$. To remedy this, we pass from L to an associated language. Fix a DFA $(\Sigma_d, Q, q_0, \Omega, \delta)$ recognizing L. Given $q \in Q$, we let $L_q = \{\, \sigma \in \Sigma_d^* : \delta(q, \sigma) = q \,\}$; that is, L_q is the set of $\sigma \in \Sigma_d^*$ such that if the machine starts in state q and receives σ as input then it returns to state q. Then L_q is regular: it is recognized by the DFA $(\Sigma_d, Q, q, \{\, q \,\}, \delta)$. Moreover $L_q = L_q^*$.

We will work with L_q for some appropriate choice of q. Of course, if no accepting state is reachable from q, then L_q doesn't tell us much about A. We will thus be interested in q that are not *dead*:

Definition 5.6. If $(\Sigma, Q, q_0, \Omega, \delta)$ is a DFA and $q \in Q$, we say q is a *dead state* if $\delta(q, \sigma) \notin \Omega$ for all $\sigma \in \Sigma^*$.

The following lemma tells us that if we choose q correctly, the properties that L inherits from A being neither generic in $\mathbb{N}$ nor d-sparse carry over to L_q.

Lemma 5.7. *Suppose $(\Sigma_d, Q, q_0, \Omega, \delta)$ is a DFA recognizing some language $L \subseteq \Sigma_d^*$; suppose that L is not sparse and there is $s > 0$ such that if $0 \le r < s$ then $L \cap \Sigma_d^{(r+s\mathbb{N})}$ has a forbidden suffix. Then there is a non-dead $q \in Q$ such that the same holds of L_q: that is, L_q is not sparse and there is $s > 0$ such that if $0 \le r < s$ then $L_q \cap \Sigma_d^{(r+s\mathbb{N})}$ has a forbidden suffix.*

Proof. By Fact 3.3, we get that there is $q \in Q$ that is not a dead state, that is reachable from q_0, and for which there is $n \in \mathbb{N}$ and distinct $\sigma, \tau \in \Sigma_d^{(n)}$ such that $\delta(q, \sigma) = \delta(q, \tau) = q$; in particular, we get that L_q contains $\{\, \sigma, \tau \,\}^*$, and is thus not sparse. Since q is not a dead state, there is $\nu \in \Sigma_d^*$ such that $q' := \delta(q, \nu) \in \Omega$.

Case 1. Suppose $L_{q'}$ is not sparse. By assumption q is reachable from q_0; so there is $\rho \in \Sigma_d^*$ such that $\delta(q_0, \rho) = q$.

Suppose $0 \leq r < s$; let $0 \leq r' < r$ be such that $r + |\rho\nu| \equiv r' \pmod{s}$. By assumption $L \cap \Sigma_d^{(r'+s\mathbb{N})}$ has a forbidden suffix $\sigma_{r'}$. Then $\sigma_{r'}$ is also a forbidden suffix for $L_{q'} \cap \Sigma_d^{(r+s\mathbb{N})}$. Indeed, suppose for contradiction that $\tau \in L_{q'} \cap \Sigma_d^{(r+s\mathbb{N})}$ has $\sigma_{r'}$ as a suffix. Then $\delta(q', \tau) = q'$, and thus

$$\delta(q_0, \rho\nu\tau) = \delta(q, \nu\tau) = \delta(q', \tau) = q' \in \Omega;$$

so $\rho\nu\tau \in L$. But $|\rho\nu\tau| \in |\rho\nu| + r + s\mathbb{N} \subseteq r' + s\mathbb{N}$; so $\rho\nu\tau \in L \cap \Sigma_d^{(r'+s\mathbb{N})}$ and $\rho\nu\tau$ has $\sigma_{r'}$ as a suffix, a contradiction.

So $L_{q'} \cap \Sigma_d^{(r+s\mathbb{N})}$ has a forbidden suffix for every $0 \leq r < s$; so q' is our desired state.

Case 2. Suppose $L_{q'}$ is sparse. Note that there cannot exist $\sigma \in \Sigma_d^*$ such that $\delta(q', \sigma) = q$. Indeed, if we had such σ then we would get that $L_{q'} \supseteq \sigma L_q \nu$, and thus that

$$|\{\, \tau \in L_q : |\tau| \leq x \,\}| \leq |\{\, \tau \in L_{q'} : |\tau| \leq x + |\sigma| + |\nu| \,\}|$$

grows polynomially in x; so L_q would be sparse, a contradiction. We will use this to construct $\sigma \in \Sigma_d^*$ that does not occur as a substring of any element of L_q, which in particular implies that σ is a forbidden suffix for L_q.

Enumerate the states $\theta \in Q$ for which there exists $\tau \in \Sigma_d^*$ such that $\delta(\theta, \tau) = q$ as $q_0, \ldots, q_{n-1}$; that is, the q_i are the states from which q is reachable. Note since $\delta(q, \nu) = q'$ that q' is also reachable from each q_i. Inductively construct $\sigma_0, \ldots, \sigma_{n-1} \in \Sigma_d^*$ as follows: having constructed $\sigma_0, \ldots, \sigma_{i-1}$, if $\delta(q_i, \sigma_0 \cdots \sigma_{i-1}) = q_j$ for some j then pick σ_i such that $\delta(q_j, \sigma_i) = q'$, and otherwise let $\sigma_i = \varepsilon$. Let $\sigma = \sigma_0 \cdots \sigma_{n-1}$. Note by construction that q is not reachable from $\delta(q_i, \sigma_0 \cdots \sigma_i)$, and hence is not reachable from $\delta(q_i, \sigma)$. Moreover if $\theta \in Q$ is not one of the q_i then q is not reachable from θ, and hence is not reachable from $\delta(\theta, \sigma)$. So for any $\theta \in Q$ we get that q is not reachable from $\delta(\theta, \sigma)$; hence σ cannot occur as a substring of an element of L_q. So σ is a forbidden suffix for L_q, and thus taking $s = 1$ we see that q is our desired state. $\square$ Lemma 5.7

We now construct our ladders for $x +_K y \in L_q$:

Lemma 5.8. *Suppose $L \subseteq \Sigma_d^*$ is any language satisfying:*

- *L is regular but not sparse,*
- *$L = L^*$, and*

- *there is $s > 0$ such that if $0 \le r < s$ then $L \cap \Sigma_d^{(r+s\mathbb{N})}$ has a forbidden suffix.*

Then for all $N \in \mathbb{N}$ there is $K \in \mathbb{N}$ such that the binary relation $x +_K y \in L$ on Σ_d^* admits an N-ladder $(\mu_0, \dots, \mu_{N-1}; \nu_0, \dots, \nu_{N-1})$ such that $\mu_i +_K \nu_j$ is defined for all $i, j < N$.

The proof of this lemma is fairly involved; for the sake of intuition, we first do a concrete example.

Example 5.9. Consider the case $d = 10$ and $L = \{3, 6, 8\}^*$. Note that L satisfies the hypotheses; in particular, 4 is a forbidden suffix for L, and hence is a forbidden suffix for $L \cap \Sigma_{10}^{(r+s\mathbb{N})}$ for any r, s.

Suppose $N > 0$, and let $K = N + 1$; for notational convenience, we let $L_K = L \cap \Sigma_d^{(K)}$. Consider the preorder $\le_{10}$ on Σ_{10}^* given by $\sigma \le_{10} \tau$ if $[\sigma]_{10} \le [\tau]_{10}$; note that $\le_{10}$ forms a total order when restricted to $\Sigma_{10}^{(K)}$. Observe that there is a "boundary" for L_K at $8^N 3$: namely $8^N 3 \in L_K$, and there is a gap after $8^N 3$ in the sense that $8^N 3$ is the $\le_{10}$-largest string in L_K that ends in 3, and no strings in L_K end in 4. (Here the exponential notation $8^N 3$ refers to iterated concatenation.) More formally,

$$L_K \cap \{ \sigma \in \Sigma_{10}^{(K)} : 8^N 3 <_{10} \sigma \le_{10} 9^N 4 \} = \emptyset. \tag{†}$$

(Note that $9^N 4$ is the $\le_{10}$-largest string of length K that ends in 4.)

Suppose we can find $\mu_i, \nu_j \in \Sigma_{10}^*$ for $i, j \le N$ such that

1. the μ_i are strictly increasing in $\le_{10}$,
2. $[\mu_i]_{10} + [\nu_i]_{10} = [8^N 3]_{10}$ for all i,
3. $[\mu_i]_{10} + [\nu_j]_{10} \le [9^N 4]_{10}$ for all i, j, and
4. $\mu_i +_K \nu_j \in L_K$ for $i \le j$.

Then (3) tells us that $[\mu_i]_{10} + [\nu_j]_{10} < 10^{|9^N 4|} = 10^K$, and hence $\mu_i +_K \nu_j$ is defined for all $i, j \le N$. Moreover for $i > j$ properties (1)–(3) yield that

$$9^N 4 \ge_{10} \mu_i +_K \nu_j >_{10} \mu_j +_K \nu_j = 8^N 3$$

(since $[\mu_i +_K \nu_j]_{10} = [\mu_i]_{10} + [\nu_j]_{10}$). So $\mu_i +_K \nu_j$ lies in the gap (†) in L_K following $8^N 3$, and hence $\mu_i +_K \nu_j \notin L_K$. But property (4) tells us that $\mu_i +_K \nu_j \in L_K$ if $i \le j$; so the μ_i, ν_j form our desired ladder.

To construct a ladder satisfying properties (1)–(4), we will use the fact that $\{6, 8\}^*3 \subseteq L$. For $i \leq N$ we let $\mu_i = 6^{N-i}8^i3$ and let $\nu_i = 2^{N-i}$. Property (1) is clear, and for property (3) we note that

$$[\mu_i]_{10} + [\nu_j]_{10} < [6^{N-i}8^i3]_{10} + 10^{|\nu_j|} \leq [6^{N-i}8^i3]_{10} + 10^N = [6^{N-i}8^i4]_{10} \leq [9^N4]_{10}.$$

For properties (2) and (4), we use place value addition to compute $[\mu_i]_{10} + [\nu_j]_{10}$ for $i \leq j$:

$$\begin{array}{cccccccccccc} & \multicolumn{3}{c}{\overbrace{}^{N-j \text{ copies}}} & \multicolumn{3}{c}{\overbrace{}^{j-i \text{ copies}}} & \multicolumn{3}{c}{\overbrace{}^{i \text{ copies}}} & \\ & 6 & \cdots & 6 & 6 & \cdots & 6 & 8 & \cdots & 8 & 3 \\ + & 2 & \cdots & 2 & & & & & & & \\ \hline & 8 & \cdots & 8 & 6 & \cdots & 6 & 8 & \cdots & 8 & 3 \end{array}$$

So $\mu_i +_K \nu_j = 8^{N-j}6^{j-i}8^i3 \in L_K$ if $i \leq j$, and property (4) holds. Moreover $\mu_i +_K \nu_i = 8^N3$ for all i, and property (2) holds.

So properties (1)–(4) hold, and hence by the above we get that $(\mu_0, \ldots, \mu_N; \nu_0, \ldots, \nu_N)$ forms an $(N+1)$-ladder for $x +_K y \in L$.

Our proof will generalize this example.

Proof of Lemma 5.8. Suppose $N \in \mathbb{N}$; to avoid notational clutter, we will in fact produce an $(N+1)$-ladder with the desired property, rather than just an N-ladder.

We first look for $K \in \mathbb{N}$ for which there is a gap in $L \cap \Sigma_d^{(K)}$ as in the example. To do this, we will use the forbidden suffix property of L. Indeed, fix $s > 0$ such that if $0 \leq r < s$ then $L \cap \Sigma_d^{(r+s\mathbb{N})}$ has a forbidden suffix. Since L is infinite (as it isn't sparse), the pigeonhole principle yields some $0 \leq r < s$ such that $L \cap \Sigma_d^{(r+s\mathbb{N})}$ is infinite; let $\tau \in \Sigma_d^*$ be a forbidden suffix for $L \cap \Sigma_d^{(r+s\mathbb{N})}$. For notational convenience, we let $L_{r+s\mathbb{N}} = L \cap \Sigma_d^{(r+s\mathbb{N})}$; as before, if $K \in \mathbb{N}$ we let $L_K = L \cap \Sigma_d^{(K)}$.

As in the example, let $\leq_d$ be the preorder on Σ_d^* given by $\rho \leq_d \rho'$ if $[\rho]_d \leq [\rho']_d$; again note that $\leq_d$ restricts to a total order on $\Sigma_d^{(K)}$ for any $K \in \mathbb{N}$. Since $L_{r+s\mathbb{N}}$ is infinite, there is some $\sigma \in \Sigma_d^*$ with $|\sigma| = |\tau|$ that occurs as a suffix of some element of $L_{r+s\mathbb{N}}$. Assume there is such a σ with $\sigma <_d \tau$; we will describe afterwards how to deal with the case where there is no such σ.

If $K \in r + s\mathbb{N}$ and $K \geq |\tau|$, then the fact that τ is a forbidden suffix for $L_{r+s\mathbb{N}}$ is enough to produce a gap in L_K: if $\rho \in \Sigma_d^{(K)}$ ends in τ (that is, $0^{K-|\tau|}\tau \leq_d \rho \leq_d (d-1)^{K-|\tau|}\tau$), then since τ is a forbidden suffix for $L_{r+s\mathbb{N}} \supseteq L_K$ we get that $\rho \notin L_K$. For our "boundary" point, we want to choose the $\leq_d$-largest element of L_K that precedes this gap: we let

$$\partial_K = \max\{\, \rho \in L_K : \rho <_d 0^{K-|\tau|}\tau \,\}$$

if such ρ exist (where the maximum is taken under $\leq_d$). So $\partial_K \in L_K$, and there is a gap in L_K following ∂_K: using maximality of ∂_K and the fact that τ is a forbidden suffix, we get that

$$L_K \cap \{\, \rho \in \Sigma_d^{(K)} : \partial_K <_d \rho' \leq_d (d-1)^{K-|\tau|}\tau \,\} = \emptyset. \tag{‡}$$

Suppose $K \in r + s\mathbb{N}$ is such that ∂_K is defined; that is, $K \geq |\tau|$ and there is $\rho \in L_K$ satisfying $\rho <_d 0^{K-|\tau|}\tau$. (We will see in Claim 5.10 that there are infinitely many such K.) Suppose we can find $\mu_i, \nu_i \in \Sigma_d^*$ for $i \leq N$ such that

1. the μ_i are strictly increasing in $\leq_d$,
2. $[\mu_i]_d + [\nu_i]_d = [\partial_K]_d$ for all i,
3. $[\mu_i]_d + [\nu_j]_d \leq [(d-1)^{K-|\tau|}\tau]_d$ for all i, j, and
4. $\mu_i +_K \nu_j \in L$ for $i \leq j$.

Then (3) tells us that $[\mu_i]_d + [\nu_j]_d < d^K$, and hence that $\mu_i +_K \nu_j$ is defined for all $i, j \leq N$. Moreover for $i > j$ properties (1)–(3) yield that

$$(d-1)^{K-|\tau|}\tau \geq_d \mu_i +_K \nu_j >_d \mu_j +_K \nu_j = \partial_K.$$

So $\mu_i +_K \nu_j$ lies in the gap (‡) in L_K that follows $uv^N w \in S$, and thus $\mu_i +_K \nu_j \notin L$. But property (4) guarantees that $\mu_i +_K \nu_j \in L$ for $i \leq j$; so the μ_i, ν_i form our desired ladder.

So we need only find μ_i, ν_i satisfying properties (1)–(4). In our example, our boundary point took on a very simple form, namely $8^N 3$, and we moreover had that $\{\,6, 8\,\}^*3 \subseteq L$; this allowed us to construct such μ_i, ν_i. To mimic this in our more general context, we will appeal to the pumping lemma. Let S be the set of ∂_K for $K \in r + s\mathbb{N}$ such that ∂_K is defined; i.e., such that $K \geq |\tau|$ and there is $\rho \in L_K$ satisfying $\rho <_d 0^{K-|\tau|}\tau$.

Claim 5.10. *S is regular and infinite.*

Proof. Let $S_0 = \{\, \rho \in L_{r+s\mathbb{N}} : |\rho| \geq |\tau|, \rho <_d 0^{|\rho|-|\tau|}\tau \,\}$. Then S is the set of $\partial \in S_0$ such that if $\rho \in S_0$ and $|\partial| = |\rho|$ then $\rho \leq_d \partial$; that is, $S_0 \setminus S$ is the projection to the first coordinate of

$$S_2 := \left\{ \begin{pmatrix} \rho \\ \rho' \end{pmatrix} \in (\Sigma_d^2)^* : \rho, \rho' \in S_0, \rho' >_d \rho \right\}.$$

(One might think that in order for S_2 to project to $S_0 \setminus S$ we should further demand that if $\begin{pmatrix}\rho\\ \rho'\end{pmatrix} \in S_2$ then $|\rho| = |\rho'|$. This actually comes for free when we assume that $\begin{pmatrix}\rho\\ \rho'\end{pmatrix} \in (\Sigma_d^2)^*$: recall that when we view an element of $(\Sigma_d^2)^*$ as a pair of strings, the two strings are necessarily of the same length.)

Thus if we can show that S_0 and S_2 are regular, it will follow from closure properties (Corollaries 2.5 and 2.8) that S is regular. For S_0, we note for $\rho \in \Sigma_d^*$ with $|\rho| \geq |\tau|$, say $\rho = \rho_0\sigma$ for $\sigma \in \Sigma_d^{(|\tau|)}$, that $\rho <_d 0^{|\rho|-|\tau|}\tau$ if and only if $\sigma <_d \tau$. Indeed, if $\sigma <_d \tau$ then

$$[\rho]_d = [\rho_0]_d + d^{|\rho_0|}[\sigma]_d < d^{|\rho_0|} + d^{|\rho_0|}[\sigma]_d \leq d^{|\rho_0|}[\tau]_d = [0^{|\rho_0|}\tau]_d;$$

conversely if $\sigma \geq_d \tau$ then

$$[\rho]_d = [\rho_0]_d + d^{|\rho_0|}[\sigma]_d \geq d^{|\rho_0|}[\sigma]_d \geq d^{|\rho_0|}[\tau]_d = [0^{|\rho_0|}\tau]_d.$$

It follows that S_0 is simply the set of strings in $L_{r+s\mathbb{N}}$ that have σ as a suffix for some $\sigma \in \Sigma_d^{(|\tau|)}$ with $\sigma <_d \tau$; that is,

$$S_0 = L_{r+s\mathbb{N}} \cap \Sigma_d^*\{\, \sigma \in \Sigma_d^{(|\tau|)} : \sigma <_d \tau \,\}.$$

But $L_{r+s\mathbb{N}} = L \cap \Sigma_d^{(r+s\mathbb{N})} = L \cap \Sigma_d^{(r)}(\Sigma_d^{(s)})^*$ is regular since L is, and $\Sigma_d^*\{\, \sigma \in \Sigma_d^{(|\tau|)} : \sigma <_d \tau \,\}$ is regular by definition. So closure of regular languages under intersection yields that S_0 is regular.

For S_2, we note that using a similar argument to the one given in the previous paragraph, we can show that

$$\left\{ \begin{pmatrix}\rho\\ \rho'\end{pmatrix} \in (\Sigma_d^2)^* : \rho <_d \rho' \right\} = \Sigma_d^* X \Delta^*$$

where $\Delta \subseteq \Sigma_d^2$ is the diagonal and $X = \left\{ \begin{pmatrix}a\\ b\end{pmatrix} : a, b \in \Sigma_d, a < b \right\}$; that is, $\rho <_d \rho'$ if and only if on the most significant digit on which ρ and ρ' differ, the digit in ρ' is larger than the digit in ρ. So

$$S_2 = \left\{ \begin{pmatrix}\rho\\ \rho'\end{pmatrix} \in (\Sigma_d^2)^* : \rho, \rho' \in S_0 \right\} \cap \Delta^* X \Sigma_d^*.$$

But $\Delta^* X \Sigma_d^*$ is regular by definition, and we can use a DFA recognizing S_0 (which must exist, as we showed above that S_0 is regular) to construct a DFA recognizing $\left\{ \begin{pmatrix}\rho\\ \rho'\end{pmatrix} \in (\Sigma_d^2)^* : \rho, \rho' \in S_0 \right\}$. So again using our closure properties we get that S_2 is regular, and hence that S is regular.

We now show that S is infinite. Recall by our earlier assumption that there is $\sigma \in \Sigma_d^{(|\tau|)}$ that occurs as a suffix of some $\rho \in L_{r+s\mathbb{N}} = L \cap \Sigma_d^{(r+s\mathbb{N})}$ and satisfies $\sigma <_d \tau$. So, since $L = L^*$, it follows that if $i \in \mathbb{N}$ then $\rho^{1+si} \in L_{r+s\mathbb{N}}$ and ρ^{1+si} has σ as a suffix. Moreover, as we argued in the proof that S_0 is regular, this implies that $\rho^{1+si} <_d 0^{|\rho^{1+si}|-|\tau|}\tau$. So $\partial_{|\rho^{1+si}|}$ is defined for all i; so S is infinite. $\square$ Claim 5.10

Since S is infinite, there is an element of S whose length exceeds the pumping length of S; so by the pumping lemma there are $u, v, w \in \Sigma_d^*$ with $v \neq \varepsilon$ such that $uv^*w \subseteq S$. By prepending some power of v to w we may assume that $|w| \geq |\tau|$ (and that $w \neq \varepsilon$); say $w = w_0\sigma$ for some $\sigma \in \Sigma_d^{(|\tau|)}$ and $w_0 \in \Sigma_d^*$. Then $\sigma <_d \tau$, since otherwise we would have that

$$[uw]_d = [uw_0\sigma]_d \geq d^{|uw_0|}[\sigma]_d \geq d^{|uw_0|}[\tau]_d = [0^{|uw|-|\tau|}\tau]_d,$$

contradicting our assumption that $uw \in S$.

Note that $uv^N w$ is a boundary point ∂_K for $K = |uv^N w|$; this gives us some generalization of the fact that our boundary point in the example was $8^N 3$. To construct a ladder in the example, we further used the fact that $\{ 6, 8 \}^* 3 \subseteq L$; to get an analogous property in this context, we will use the fact that L is not sparse.

Since L is regular but not sparse, there are $\alpha_1, \alpha_2 \in L$ with $\alpha_1 \neq \alpha_2$ and $|\alpha_1| = |\alpha_2| \neq 0$; indeed, otherwise there would be at most one string in L of any given length, and L would be sparse. It follows that $u\{ w\alpha_1 u, w\alpha_2 u, v \}^* w \subseteq L$. Indeed, an element of $u\{ w\alpha_1 u, w\alpha_2 u, v \}^* w$ can be written in the form

$$uv^{k_1}(w\alpha_{i_1}u)v^{k_2} \cdots (w\alpha_{i_{n-1}}u)v^{k_n}w = (uv^{k_1}w)\alpha_{i_1}(uv^{k_2}w) \cdots \alpha_{i_{n-1}}(uv^{k_n}w)$$

for some $k_1, \ldots, k_n \in \mathbb{N}$ and some $i_1, \ldots, i_{n-1} \in \{ 1, 2 \}$. Hence since $uv^*w \subseteq S \subseteq L$ and $\alpha_1, \alpha_2 \in L$, we get that $u\{ w\alpha_1 u, w\alpha_2 u, v \}^* w \subseteq L^* = L$.

Since our only requirement of v is that $uv^*w \subseteq S$ and $v \neq \varepsilon$, we can replace v with a power thereof, and thus assume that $|w\alpha_1 u|$ divides $|v|$; say $|v| = M|w\alpha_1 u|$. Then $(w\alpha_1 u)^M \neq (w\alpha_2 u)^M$, so there is $\beta \in \{ (w\alpha_1 u)^M, (w\alpha_2 u)^M \}$ such that $\beta \neq v$. So $|\beta| = |v|$, and $u\{ \beta, v \}^* w \subseteq L$. We will use $u\{ \beta, v \}^* w$ as our analogue of $\{ 6, 8 \}^* 3$ in the example.

Since $|\beta| = |v|$ and $\beta \neq v$, we get that $[v]_d \neq [\beta]_d$. Using the definition of S, we look to show that $\beta <_d v$. Recall that $w = w_0\sigma$ with $|\sigma| = |\tau|$ and $\sigma <_d \tau$; so

$$[u\beta w]_d = [u\beta w_0\sigma]_d = [u\beta w_0]_d + d^{|u\beta w_0|}[\sigma]_d < d^{|u\beta w_0|} + d^{|u\beta w_0|}[\sigma]_d \leq d^{|u\beta w_0|}[\tau]_d = [0^{|u\beta w_0|}\tau]_d,$$

and hence $u\beta w <_d 0^{|u\beta w_0|}\tau = 0^{|uvw|-|\tau|}\tau$. Moreover $uvw \in S$, so

$$uvw = \partial_{|uvw|} = \max\{ \rho \in L_{|uvw|} : \rho <_d 0^{|uvw|-|\tau|}\tau \}$$

(using $\leq_d$ as the ordering); furthermore $|u\beta w| = |uvw|$ and $u\beta w \in L$. So maximality of uvw yields that $u\beta w \leq_d uvw$, and thus that $\beta \leq_d v$; so, since $\beta \neq v$, we get that $\beta <_d v$, as claimed. Since $0 < [v]_d - [\beta]_d \leq [v]_d < d^{|v|}$, we get that $[v]_d - [\beta]_d$ has a base-d representation of length $|v|$, say $\eta \in \Sigma_d^{(|v|)}$. This η will play the role that 2 did in the example.

Let $K = |uv^N w|$; so since $uv^N w \in S$ we get that $K \in r + s\mathbb{N}$ and $uv^N w = \partial_K$. For $i \leq N$ we let

$$\begin{aligned} \mu_i &= u\beta^{N-i} v^i w \\ \nu_i &= 0^{|u|}\eta^{N-i}. \end{aligned}$$

We show that the μ_i, ν_i satisfy properties (1)–(4). Property (1) follows from the fact that $\beta <_d v$, and for property (3) we note that

$$\begin{aligned} [\mu_i]_d + [\nu_j]_d &< [\mu_N]_d + d^{|\nu_j|} \\ &= [uv^N w_0 \sigma]_d + d^{|uv^{N-i}|} \\ &\leq [uv^N w_0 \sigma]_d + d^{|uv^N w_0|} \\ &= [uv^N w_0]_d + d^{|uv^N w_0|}([\sigma]_d + 1) \\ &\leq [uv^N w_0]_d + d^{|uv^N w_0|}[\tau]_d \\ &= [uv^N w_0 \tau]_d \\ &\leq [(d-1)^{|uv^N w_0|}\tau]_d \\ &\leq [(d-1)^{K-|\tau|}\tau]_d \end{aligned}$$

(recalling that $\sigma <_d \tau$). For properties (2) and (4), we use (some generalization of) place value addition to compute $[\mu_i]_d + [\nu_j]_d$ for $i \leq j$:

$$\begin{array}{cccccc} & u & \overbrace{\beta \ \cdots \ \beta}^{N-j \text{ copies}} & \overbrace{\beta \ \cdots \ \beta}^{j-i \text{ copies}} & \overbrace{v \ \cdots \ v}^{i \text{ copies}} & w \\ + & 0^{|u|} & \eta \ \cdots \ \eta & & & \\ \hline & u & v \ \cdots \ v & \beta \ \cdots \ \beta & v \ \cdots \ v & w \end{array}$$

(since $[\beta]_d + [\eta]_d = [v]_d$). Moreover $|uv^{N-j}\beta^{j-i}v^i w| = |uv^N w| = K \in r + s\mathbb{N}$; so

$$\mu_i +_K \nu_j = uv^{N-j}\beta^{j-i}v^i w \in u\{\,\beta, v\,\}^* w \subseteq L$$

for $i \leq j$, and property (4) holds. Furthermore $\mu_i +_K \nu_i = uv^N w = \partial_K$ for all $i \leq N$, and property (2) holds. So properties (1)–(4) hold, and $(\mu_0, \ldots, \mu_N; \nu_0, \ldots, \nu_N)$ is our desired $(N+1)$-ladder for $x +_K y \in L$.

At the beginning of the proof, we assumed that there was $\sigma \in \Sigma_d^{(|\tau|)}$ with $\sigma <_d \tau$ that occurs as a suffix of some element of $L_{r+s\mathbb{N}}$; we now describe how to deal with the case where no such σ exists. Since $L_{r+s\mathbb{N}}$ is infinite, there must exist $\sigma \in \Sigma_d^{(|\tau|)}$ that occurs as a suffix of some element of $L_{r+s\mathbb{N}}$; it then follows that all such σ satisfy $\sigma >_d \tau$. (Certainly we can't have $\sigma = \tau$ since τ is a forbidden suffix for $L_{r+s\mathbb{N}}$.) So we try to recreate the above proof with the order inverted.

We now let $\partial_K = \min\{\, \rho \in L_K : \rho >_d (d-1)^{K-|\tau|}\tau \,\}$; so there is now a gap *preceding* ∂_K. We deduce that if $K \in r + s\mathbb{N}$ is such that ∂_K is defined, and if we can find μ_i, ν_i for $i \leq N$ satisfying:

1. the μ_i are strictly decreasing in $\leq_d$,
2. $[\mu_i]_d + [\nu_i]_d = [\partial_K]_d$ for all i,
3. $[\mu_i]_d + [\nu_j]_d \geq [0^{K-|\tau|}\tau]_d$ for all i, j, and
4. $\mu_i +_K \nu_j \in L$ for $i \leq j$

then $\mu_i +_K \nu_j$ is defined for all $i, j \leq N$ and $\mu_i +_K \nu_j \in L_K \iff i \leq j$.

Again let S be the set of ∂_K for $K \in r + s\mathbb{N}$ such that ∂_K is defined. As before we find that S is infinite and regular; so the pumping lemma yields u, v, w with $v \neq \varepsilon$ such that $uv^*w \subseteq S$. Again we may assume $|w| \geq |\tau|$; say $w = w_0\sigma$ for some $\sigma \in \Sigma_d^{(|\tau|)}$ and $w_0 \in \Sigma_d^*$. Dually to before we get that $\sigma >_d \tau$.

We again use the fact that L is regular but not sparse to produce $\beta \in \Sigma_d^*$ such that $\beta \neq v$, $|\beta| = |v|$, and $u\{\,\beta, v\,\}^*w \subseteq L$. We deduce using the definition of S that $\beta >_d v$, and hence that there is $\eta \in \Sigma_d^{(|v|)}$ such that $[\eta]_d = [\beta]_d - [v]_d$.

We now run into a small wrinkle in producing the ladder. We would like to take $K = |uv^N w|$ and

$$\begin{aligned} \mu_i &= u\beta^{N-i}v^i w \\ \nu_i &= 0^{|u|}(-\eta^{N-i}) \end{aligned}$$

(where $-\eta \in \Sigma_d^{\pm}$ is the characterwise negation of η) so that dually to before we get for $i \leq j$ that $[\mu_i]_d + [\nu_j]_d = [uv^{N-j}\beta^{j-i}v^i w]_d$, and hence that the μ_i, ν_i satisfy the desired properties. The problem is that since we are working over Σ_d, not $\Sigma_d^{\pm}$, we can't use strings with negative digits. Happily, this is easily circumvented. Recall that $w = w_0\sigma$ and $\sigma >_d \tau$; so in particular $[w]_d \geq 1$. So for $i \leq N$ we get that

$$[\mu_i]_d \geq d^{|uv^N|}[w]_d \geq d^{|uv^N|} \geq d^{|\nu_i|} \geq -[\nu_i]_d.$$

So we can pick $\mu_i', \nu_i' \in \Sigma_d^*$ such that $[\mu_i']_d = [\mu_i]_d - d^{|uv^N|}$ and $[\nu_i']_d = [\nu_i]_d + d^{|uv^N|}$. Then for any i, j we get that $[\mu_i']_d + [\nu_j']_d = [\mu_i]_d + [\nu_j]_d$. That $(\mu_0', \ldots, \mu_N'; \nu_0', \ldots, \nu_N')$ satisfies properties (1)–(4) then follows from the fact that $(\mu_0, \ldots, \mu_N; \nu_0, \ldots, \nu_N)$ does. $\square$ Lemma 5.8

We are now ready to prove Theorem 5.3 in the context of $\mathbb{N}$. Our proof will make use of *minimal automata*:

Fact 5.11 ([24, Theorem 4.7])**.** *Suppose Σ is a finite alphabet and $L \subseteq \Sigma^*$ is regular. There is a DFA $(\Sigma, Q, q_0, \Omega, \delta)$ recognizing L satisfying:*

- *Every state is reachable from the start state: that is, if $q \in Q$ then there is $\sigma \in \Sigma^*$ such that $\delta(q_0, \sigma) = q$.*
- *Given distinct $q, q' \in Q$ there is $\sigma \in \Sigma^*$ such that exactly one of $\delta(q, \sigma)$ and $\delta(q', \sigma)$ is an accepting state.*

We call such an automaton a minimal automaton *for L.*

We will also need the following closure properties of stability:

Fact 5.12.

1. *Boolean combinations of stable relations are stable.*
2. *If $A \subseteq \Gamma$ is stable in a group $(\Gamma, +)$ and $\gamma \in \Gamma$ then $A + \gamma$ is stable in $(\Gamma, +)$.*

The former appears as [23, Exercise 8.2.9], and is a consequence of the counting-types characterization of stability (see for example [23, Theorem 8.2.3]). For the latter, one simply notes that if $(a_0, \ldots, a_{N-1}; b_0, \ldots, b_{N-1})$ is an N-ladder for $x + y \in A + \gamma$, then $(a_0 - \gamma, \ldots, a_{N-1} - \gamma; b_0, \ldots, b_{N-1})$ is an N-ladder for $x + y \in A$.

We are now ready to prove our theorem in the context of $\mathbb{N}$:

Proposition 5.13. *If $A \subseteq \mathbb{N}$ is d-automatic, stable in $\mathbb{N}$, and not generic in $\mathbb{N}$ then A is d-sparse.*

Proof. Suppose A is d-automatic but neither generic in $\mathbb{N}$ nor d-sparse; we show that $x + y \in A$ isn't stable in $\mathbb{N}$. Fix a minimal automaton $(\Sigma_d, Q, q_0, \Omega, \delta)$ for $L := \{\, \sigma \in \Sigma_d^* : [\sigma]_d \in A \,\}$. Since A isn't d-sparse, we get that L isn't sparse, and since A isn't generic in $\mathbb{N}$

Proposition 5.4 yields $s > 0$ such that if $0 \le r < s$ then $L \cap \Sigma_d^{(r+s\mathbb{N})}$ has a forbidden suffix. By Lemma 5.7 there is $q \in Q$ such that L_q isn't sparse and satisfies the same forbidden suffix condition as L. Suppose $N \in \mathbb{N}$; then by Lemma 5.8 there is $K \in \mathbb{N}$ such that there is an N-ladder $(\mu_0, \ldots, \mu_{N-1}; \nu_0, \ldots, \nu_{N-1})$ for $x +_K y \in L_q$ with the property that $\mu_i +_K \nu_j$ is defined for all $i, j < N$. Note that since

$$[\mu_i]_K \le [\mu_i]_K + [\nu_i]_K = [\mu_i +_K \nu_i] < d^K$$

we may assume each $|\mu_i| = K$.

Minimality of $(\Sigma_d, Q, q_0, \Omega, \delta)$ yields that for $q' \neq q$ there is $\sigma_{q'} \in \Sigma_d^*$ and $\varepsilon_{q'} \in \{0, 1\}$ such that

$$(\delta(q, \sigma_{q'}) \in \Omega)^{\varepsilon_{q'}} \wedge (\delta(q', \sigma_{q'}) \in \Omega)^{1-\varepsilon_{q'}}$$

holds (where φ^1 denotes φ and φ^0 denotes $\neg\varphi$). So for $\theta \in Q$ we get that $\theta = q$ if and only if

$$\bigwedge_{q' \neq q} (\delta(\theta, \sigma_{q'}) \in \Omega)^{\varepsilon_{q'}}.$$

By minimality of $(\Sigma_d, Q, q_0, \Omega, \delta)$ there is $\tau \in \Sigma_d^*$ such that $\delta(q_0, \tau) = q$. Then for $i, j < N$ we get

$$\begin{aligned}
i \le j &\iff \mu_i +_K \nu_j \in L_q \\
&\iff \delta(q, \mu_i +_K \nu_j) = q \\
&\iff \bigwedge_{q' \neq q} (\delta(\delta(q, \mu_i +_K \nu_j), \sigma_{q'}) \in \Omega)^{\varepsilon_{q'}} \\
&\iff \bigwedge_{q' \neq q} (\delta(q, (\mu_i +_K \nu_j)\sigma_{q'}) \in \Omega)^{\varepsilon_{q'}} \\
&\iff \bigwedge_{q' \neq q} (\delta(q_0, \tau(\mu_i +_K \nu_j)\sigma_{q'}) \in \Omega)^{\varepsilon_{q'}} \\
&\iff \bigwedge_{q' \neq q} (\tau(\mu_i +_K \nu_j)\sigma_{q'} \in L)^{\varepsilon_{q'}} \\
&\iff \bigwedge_{q' \neq q} ([\tau(\mu_i +_K \nu_j)\sigma_{q'}]_d \in A)^{\varepsilon_{q'}} \\
&\iff \bigwedge_{q' \neq q} ([\tau\mu_i\sigma_{q'}]_d + d^{|\tau|}[\nu_j]_d \in A)^{\varepsilon_{q'}},
\end{aligned}$$

where for the last line we are using our assumption that $|\mu_i| = K$ to deduce that

$$[\tau(\mu_i +_K \nu_j)\sigma_{q'}]_d = [\tau]_d + d^{|\tau|}([\mu_i]_d + [\nu_j]_d) + d^{|\tau|+K}[\sigma_{q'}]_d = [\tau\mu_i\sigma_{q'}]_d + d^{|\tau|}[\nu_j]_d.$$

Let $a_{i,q'} = [\tau\mu_i\sigma_{q'}]_d \in \mathbb{N}$ for $i < N$ and $q' \neq q$; let $\mathbf{a}_i = (a_{i,q'} : q' \neq q)$ and $b_i = d^{|\tau|}[\nu_i]_d$ for $i < N$. Then the above shows that $(\mathbf{a}_0, \ldots, \mathbf{a}_{N-1}; b_0, \ldots, b_{N-1})$ is an N-ladder for the relation defined in $(\mathbb{N}, +, A)$ by the formula

$$\varphi((x_{q'} : q' \neq q), y) := \bigwedge_{q' \neq q} (x_{q'} + y \in A)^{\varepsilon_{q'}}.$$

So φ isn't stable in $\mathrm{Th}(\mathbb{N}, +, A)$. But φ is a Boolean combination of instances of $x + y \in A$; so if $x + y \in A$ were stable then Fact 5.12 would yield that φ is as well. So A isn't stable in $\mathbb{N}$. □ Proposition 5.13

Before proving Theorem 5.3 in general, we will need another closure property of sparse languages:

Fact 5.14 ([24, Theorem 3.8]). *If Σ is a finite alphabet and $L \subseteq \Sigma^*$ is sparse, then*

$$\{\, \sigma \in \Sigma^* : \text{ there is } \tau \in L \text{ such that } \sigma \text{ is a prefix of } \tau \,\}$$

is sparse.

Proof of Theorem 5.3. Suppose $A \subseteq \mathbb{Z}$ is d-automatic but neither d-sparse nor generic (in $\mathbb{Z}$); we show that $x + y \in A$ isn't a stable relation on $\mathbb{Z}^2$.

Note first that one of $A \cap \mathbb{N}, -A \cap \mathbb{N}$ is not generic in $\mathbb{N}$. Indeed, suppose otherwise; say we have $a_1, \ldots, a_k, b_1, \ldots, b_\ell \in \mathbb{Z}$ such that

$$\begin{aligned}
\mathbb{N} &\subseteq (a_1 + A \cap \mathbb{N}) \cup \cdots \cup (a_k + A \cap \mathbb{N}) \\
\mathbb{N} &\subseteq (b_1 + (-A \cap \mathbb{N})) \cup \cdots \cup (b_\ell + (-A \cap \mathbb{N})).
\end{aligned}$$

Then

$$\begin{aligned}
\mathbb{Z} &= \mathbb{N} \cup (-\mathbb{N}) \\
&\subseteq (a_1 + A \cap \mathbb{N}) \cup \cdots \cup (a_k + A \cap \mathbb{N}) \cup (-b_1 + A \cap (-\mathbb{N})) \cup \cdots \cup (-b_\ell + A \cap (-\mathbb{N})) \\
&\subseteq (a_1 + A) \cup \cdots \cup (a_k + A) \cup (-b_1 + A) \cup \cdots \cup (-b_\ell + A)
\end{aligned}$$

and A is generic in $\mathbb{Z}$, a contradiction. Note further that one of $A \cap \mathbb{N}, -A \cap \mathbb{N}$ is not d-sparse. Indeed, suppose otherwise. Note since $a \mapsto -a$ is an automorphism of $(\mathbb{Z}, +)$ that preserves the map $a \mapsto da$ it follows that $A \cap -\mathbb{N} = -(-A \cap \mathbb{N})$ is d-sparse; then closure of d-sparsity under union (Proposition 3.6 (2)) yields that $A = (A \cap \mathbb{N}) \cup -(-A \cap \mathbb{N})$ is d-sparse.

Case 1. Suppose one of $A \cap \mathbb{N}, -A \cap \mathbb{N}$ is neither generic in $\mathbb{N}$ nor d-sparse; by possibly negating, we may assume it is $A \cap \mathbb{N}$. By Proposition 5.13 we get that $A \cap \mathbb{N}$ isn't stable in $\mathbb{N}$; so for $N \in \mathbb{N}$ there is an N-ladder $(a_0, \dots, a_{N-1}; b_0, \dots, b_{N-1})$ for $x + y \in A \cap \mathbb{N}$ (as a binary relation on $\mathbb{N}^2$). Then since each $a_i, b_j \geq 0$ we get for $i, j < N$ that $a_i + b_j \in A \iff a_i + b_j \in A \cap \mathbb{N} \iff i \leq j$; so $(a_0, \dots, a_{N-1}; b_0, \dots, b_{N-1})$ forms an N-ladder in $\mathbb{Z}$ for $x + y \in A$ (as a binary relation on $\mathbb{Z}^2$). So A isn't stable in $(\mathbb{Z}, +)$.

Case 2. Suppose otherwise; so each of $A \cap \mathbb{N}$ and $-A \cap \mathbb{N}$ is either generic in $\mathbb{N}$ or d-sparse. But we showed above that at most one of $A \cap \mathbb{N}$ and $-A \cap \mathbb{N}$ is generic in $\mathbb{N}$, and likewise at most one is d-sparse. So one of $A \cap \mathbb{N}$ is d-sparse, and the other is generic in $\mathbb{N}$. Assume by possibly negating that $-A \cap \mathbb{N}$ is generic in $\mathbb{N}$ and $A \cap \mathbb{N}$ is d-sparse; say we have $a_1, \dots, a_k \in \mathbb{Z}$ such that $\mathbb{N} \subseteq (a_1 + (-A \cap \mathbb{N})) \cup \dots \cup (a_k + (-A \cap \mathbb{N}))$. Let $A' = (-a_1 + A) \cup \dots \cup (-a_k + A)$; so

$$A' = -((a_1 + (-A)) \cup \dots \cup (a_k + (-A))) \supseteq -((a_1 + (-A \cap \mathbb{N})) \cup \dots \cup (a_k + (-A \cap \mathbb{N}))) \supseteq -\mathbb{N}.$$

We show that $A' \cap \mathbb{N}$ is d-sparse. Indeed, if $1 \leq i \leq k$ and $a_i \geq 0$ then

$$(-a_i + A) \cap \mathbb{N} = -a_i + (A \cap \{ a_i, a_i + 1, \dots \}) = -a_i + ((A \cap \mathbb{N}) \cap \{ a_i, a_i + 1, \dots \})$$

is d-sparse (as can be seen, for example, using Corollary 2.35, Proposition 3.6 (3), and Remark 3.9); if on the other hand $a_i \leq 0$ then

$$\begin{aligned} (-a_i + A) \cap \mathbb{N} &= -a_i + (A \cap \{ a_i, a_i + 1, \dots \}) \\ &= (-a_i + (A \cap \mathbb{N})) \cup (-a_i + (A \cap \{ a_i, a_{i+1}, \dots, -1 \})) \end{aligned}$$

is d-sparse (as can be seen, for example, using Example 3.5 (1), Remark 3.9, and Proposition 3.6 (2)). So $(-a_i + A) \cap \mathbb{N}$ is d-sparse for each $1 \leq i \leq k$; so since d-sparse sets are closed under finite unions (Proposition 3.6 (2)) we get that

$$A' \cap \mathbb{N} = ((-a_1 + A) \cap \mathbb{N}) \cup \dots \cup ((-a_k + A) \cap \mathbb{N})$$

is d-sparse.

This implies that $A' \cap \mathbb{N}$ misses some coset:

Claim 5.15. *If $B \subseteq \mathbb{N}$ is d-sparse then there are $s > 0$ and $0 \leq r < s$ such that $B \cap (r + s\mathbb{N}) = \emptyset$.*

Proof. Let $L = \{ \sigma \in \Sigma_d^* : [\sigma]_d \in B \}$; so $L = \Sigma_d^* \cap \{ \sigma \in (\Sigma_d^{\pm})^* : [\sigma]_d \in B \}$ is regular. Let $\preceq$ be the usual ordering on Σ_d. Then by Lemma 3.18 we get that

$$\widetilde{L} := \{ \sigma \in \Sigma_d^* : [\sigma]_d \in B, \sigma \preceq \tau \text{ for all } \tau \in \Sigma_d^* \text{ such that } [\tau]_d \in B \}$$

is sparse. But over Σ_d we have that representations are essentially unique: that is, if $\sigma, \tau \in \Sigma_d^*$ and $[\sigma]_d = [\tau]_d$ then one is obtained from the other by appending zeroes. So $L = \widetilde{L}0^*$, and thus L is sparse by Remark 3.2.

So by Fact 5.14 we get that the set of prefixes of elements of L is sparse, and in particular is not all of Σ_d^*; hence there is $\sigma \in \Sigma_d^*$ that is a forbidden prefix for L. Let $r = [\sigma]_d$ and $s = d^{|\sigma|}$. Suppose we are given $r + sn \in r + s\mathbb{N}$; fix $\tau \in \Sigma_d^*$ such that $[\tau]_d = n$. If we had $r + sn \in B$, then since $[\sigma\tau]_d = [\sigma]_d + d^{|\sigma|}[\tau]_d = r + sn \in B$, we would have $\sigma\tau \in L$, a contradiction. So $B \cap (r + s\mathbb{N}) = \emptyset$, as desired. □ Claim 5.15

So there are $s > 0$ and $0 \leq r < s$ such that $A' \cap (r + s\mathbb{Z}) = (A' \cap -\mathbb{N}) \cap (r + s\mathbb{Z}) = r + s\mathbb{Z}_{<0}$ (since $A' \supseteq -\mathbb{N}$). Thus if we let $b_i = r + is$ and $c_i = -(i+1)s$ for $i \in \mathbb{N}$, then for $i, j \in \mathbb{N}$ we have

$$\begin{aligned} b_i + c_j \in A' &\iff r + (i - j - 1)s \in A' \\ &\iff r + (i - j - 1)s \in A' \cap (r + s\mathbb{Z}) \\ &\iff i - j - 1 < 0 \\ &\iff i \leq j. \end{aligned}$$

So $x + y \in A'$ is an unstable formula in $\mathrm{Th}(\mathbb{Z}, +, A)$. But A' is a union of translates of A; thus Fact 5.12 yields that if A were stable in $\mathbb{N}$ then A' would be as well. So A is unstable in $\mathbb{N}$. □ Theorem 5.3

5.2 Characterizing stable automatic sets

A general characterization of the F-automatic $A \subseteq \mathbb{Z}$ that are stable in $(\mathbb{Z}, +)$ follows quickly from our earlier work, together with the following fact from stable group theory:

Fact 5.16 ([9, Theorem 2.3 (iv)]). *If Γ is a group and $A \subseteq \Gamma$ is stable in Γ then A has non-generic symmetric difference from a union of cosets of a subgroup of Γ.*

(We say $A \subseteq \Gamma$ is *generic in* Γ if there are $a_1, \ldots, a_n \in \Gamma$ such that $\Gamma \subseteq (a_1 + A) \cup \cdots \cup (a_n + A)$.)

Theorem 5.17. *If $A \subseteq \mathbb{Z}$ then the following are equivalent:*

1. *A is d-automatic and $\mathrm{Th}(\mathbb{Z}, +, A)$ is stable.*

2. *A is d-automatic and stable in $(\mathbb{Z}, +)$.*

3. *A is a Boolean combination of elementary d-sets and cosets of subgroups of $\mathbb{Z}$.*

4. *A is definable in $(\mathbb{Z}, +, d^{\mathbb{N}})$.*

Proof.

(1) $\implies$ (2) This is by the definitions of stability.

(2) $\implies$ (3) By Fact 5.16 there is $H \leq \Gamma$ and a union Y of cosets of H such that $A \triangle Y$ is not generic. Moreover cosets are d-sets, and hence by Theorem 2.54 and Fact 4.3 are d-automatic and stable in Γ. (It is easily verified that the hypotheses of Fact 4.3 are satisfied, since $\mathbb{Z}[d] = \mathbb{Z}$.) Furthermore both d-automaticity and stability are closed under Boolean combinations (Corollary 2.29 and Fact 5.12). Thus $A \triangle Y$ is d-automatic and stable in Γ. So Theorem 5.3 yields that A is d-sparse, at which point Theorem 4.11 yields that A is a Boolean combination of elementary d-sets. So $A = (A \triangle Y) \triangle Y$ is a Boolean combination of cosets and elementary d-sets.

(3) $\implies$ (4) Since the only subgroups of $(\mathbb{Z}, +)$ are of the form $s\mathbb{Z}$ for some $s \in \mathbb{N}$, it follows that cosets are definable in $(\mathbb{Z}, +, d^{\mathbb{N}})$. It then suffices to show that elementary d-sets are definable in $(\mathbb{Z}, +, d^{\mathbb{N}})$; since addition is definable in $(\mathbb{Z}, +, d^{\mathbb{N}})$, and since $K(a; d^r) = a\frac{d^{r\mathbb{N}}-1}{d^r-1}$ for $a \in \mathbb{Z}$ and $r > 0$, it suffices to show that $d^{r\mathbb{N}}$ is definable in $(\mathbb{Z}, +, d^{\mathbb{N}})$ for $r \geq 2$.

We show that $a \in d^{r\mathbb{N}}$ if and only if $a \in d^{\mathbb{N}}$ and $d^r - 1 \mid a - 1$. The left-to-right direction is just the geometric series formula. For the right-to-left, suppose we are given $d^n \in d^{\mathbb{N}}$ such that $d^r - 1 \mid d^n - 1$; so $d^n \equiv 1 \pmod{d^r - 1}$. Write $n = ir + j$ for some $i \in \mathbb{N}$ and $0 \leq j < r$; so, since $d^r \equiv 1 \pmod{d^r - 1}$, we get that $1 \equiv d^n \equiv d^j \pmod{d^r - 1}$. Since $r, d \geq 2$, we get that

$$d^{r-1} = d^r - (1 - d^{-1}) \cdot d^r \leq d^r - (1 - 2^{-1}) \cdot 2^2 = d^r - 2 < d^r - 1.$$

So $d^j \leq d^{r-1} < d^r - 1$, and thus since $1 \equiv d^j \pmod{d^r - 1}$ we get that $d^j = 1$. So $j = 0$, and $d^n = d^{ir} \in d^{r\mathbb{N}}$. So $d^{r\mathbb{N}}$ is definable in $(\mathbb{Z}, +, d^{\mathbb{N}})$ for $r \geq 2$, as desired.

(4) $\implies$ (1) Again, the hypotheses of Fact 4.3 are satisfied since $\mathbb{Z}[d] = \mathbb{Z}$; so $\mathrm{Th}(\mathbb{Z}, \mathcal{F}_d)$ (the d-structure on $\mathbb{Z}$) is stable. But $(\mathbb{Z}, +, d^{\mathbb{N}})$ is a reduct of $(\mathbb{Z}, \mathcal{F}_d)$: addition is definable in $(\mathbb{Z}, \mathcal{F}_d)$ as its graph is a d-invariant subgroup of $\mathbb{Z}^3$, and $d^{\mathbb{N}} = K(d-1; d) + 1$ is an elementary d-set. Moreover by hypothesis we get that $(\mathbb{Z}, +, A)$ is a reduct of

$(\mathbb{Z}, +, d^{\mathbb{N}})$. So $(\mathbb{Z}, +, A)$ is a reduct of $(\mathbb{Z}, \mathcal{F}_d)$; thus since $\mathrm{Th}(\mathbb{Z}, \mathcal{F}_d)$ is stable, and since stability is defined as the absence of an unstable definable relation, we get that $\mathrm{Th}(\mathbb{Z}, +, A)$ is stable. $\square$ Theorem 5.17

5.3 Bibliographical notes

Genericity is a standard definition in stable group theory; see for example [9].

I am indebted to Gabriel Conant for pointing out that Theorem 5.3 together with Fact 5.16 imply the (2) $\implies$ (3) direction of Theorem 5.17. All other results in this chapter are original.

Chapter 6

NIP expansions

In Chapter 4, we considered the question of when an F-sparse $A \subseteq \Gamma$ is such that $\mathrm{Th}(\Gamma, +, A)$ is stable. In this chapter, we consider the question of when $\mathrm{Th}(\Gamma, +, A)$ is *NIP*. This is an interesting relaxation of stability; see [21] for a thorough description. We briefly recall the definition here:

Definition 6.1. Suppose X and Y are sets, and $R \subseteq X \times Y$ is a binary relation. We say $X_0 \subseteq X$ is *shattered* by R if for all $S \subseteq X_0$ there is $b_S \in Y$ such that $S = \{ a \in X_0 : (a, b_S) \in R \}$. We say R has the *independence property* (IP) if it shatters arbitrarily large finite subsets of X; otherwise we say R is *NIP*.

If T is a complete first-order theory and $\varphi(x_1, \ldots, x_k; y_1, \ldots, y_\ell)$ is a partitioned formula in the associated signature, we say φ is *NIP* if there is $M \models T$ such that the relation on $M^k \times M^\ell$ defined by φ is NIP. We say T is *NIP* if every partitioned formula in the associated signature is NIP.

As in the definition of stability, completeness of T lets us replace "there is $M \models T$" with "for any $M \models T$" in the definition of a formula being NIP.

Remark 6.2. If $R \subseteq X \times Y$ has IP then it is an unstable relation. Indeed, for $N < \omega$ the independence property implies that there is $X_0 \subseteq X$ that is shattered by R and satisfies $|X_0| = N$; enumerate $X_0 = \{ a_0, \ldots, a_{N-1} \}$. Since R shatters X_0, we get for $j < N$ that there is $b_j \in Y$ such that $(a_i, b_j) \in R$ if and only if $i \leq j$. Then $(a_0, \ldots, a_{N-1}; b_0, \ldots, b_{N-1})$ forms an N-ladder for R.

So stable theories are NIP. A standard example of an NIP theory that isn't stable is that of any ordered abelian group; see [12].

In this chapter, given an infinite abelian group Γ and an injective $F\colon \Gamma \to \Gamma$, we consider the question of which F-automatic $A \subseteq \Gamma$ are such that $\mathrm{Th}(\Gamma, +, A)$ is NIP. Note first that not all such A yield an NIP theory:

Example 6.3. Let $\Gamma = \mathbb{Z}^2$, F be multiplication by 2, and

$$A = \left[\left\{ \begin{pmatrix} s_0 \cdots s_{n-1} \\ t_0 \cdots t_{n-1} \end{pmatrix} \in (\{\,0,1\,\}^2)^* : s_i = t_i = 1 \text{ for some } i \right\}\right]_2 = \left[(\{\,0,1\,\}^2)^* \begin{pmatrix} 1 \\ 1 \end{pmatrix} (\{\,0,1\,\}^2)^*\right]_2.$$

That is, A is the set of pairs $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{N}^2$ for which there is some position i such that the binary representations of a and b both have a 1 in the i^{th} position. Then A is 2-automatic by Proposition 2.33 (using e.g. $\Sigma_2^{\pm} = \{\,-1,0,1\,\}^2$ as our spanning set). Moreover $\mathrm{Th}(\mathbb{Z}^2, +, A)$ isn't NIP: if $N \in \mathbb{N}$ then $\left\{ \begin{pmatrix} 2^0 \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 2^{N-1} \\ 0 \end{pmatrix} \right\}$ is shattered by the relation $x + y \in A$. Indeed, if $S \subseteq \{\,0,\ldots,N-1\,\}$ we let

$$b_S = \sum_{i \in S} 2^i.$$

So the binary representation of b_S has a 1 in position i if and only if $i \in S$. So, since the only 1 in the binary representation of 2^i is at position i, it follows that

$$\begin{pmatrix} 2^i \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ b_S \end{pmatrix} \in A \iff i \in S.$$

So $x + y \in A$ shatters $\left\{ \begin{pmatrix} 2^0 \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 2^{N-1} \\ 0 \end{pmatrix} \right\}$ for all N, and hence shatters arbitrarily large sets. So we have indeed produced an automatic set that yields a theory that isn't NIP.

We will define a class of subsets of Γ called the *F-EDP* sets that will include the F-sparse subsets of Γ. We will then show in Theorem 6.13 that if $(\Gamma, +)$ satisfies a stability hypothesis called *weak minimality* and if $A \subseteq \Gamma$ is F-EDP then $\mathrm{Th}(\Gamma, +, A)$ is NIP. We do so by interpreting the induced structure of $(\Gamma, +)$ on A in Presburger arithmetic $\mathcal{N} := (\mathbb{N}, +)$, which is NIP as it is definable in the ordered abelian group $(\mathbb{Z}, +, <)$.

The general definition of weak minimality is somewhat technical, and not useful to us. We instead use the following characterization for abelian groups:

Fact 6.4 ([10, Proposition 3.1])**.** *If Γ is an abelian group, then $\mathrm{Th}(\Gamma, +)$ is weakly minimal if and only if for all $n > 1$ the subgroups $n\Gamma$ and $\{\, a \in \Gamma : na = 0 \,\}$ either are finite or have finite index in Γ.*

In particular, our results will apply to the case $\Gamma = (\mathbb{Z}^m, +)$.

We then give three applications of Theorem 6.13. We show in Theorem 6.15 that $\mathrm{Th}(\mathbb{Z}, +, d^{\mathbb{N}}, \times\restriction d^{\mathbb{N}})$ is NIP for $d \geq 2$, where $\times\restriction d^{\mathbb{N}}$ is the graph of multiplication restricted to $d^{\mathbb{N}}$, and we show in Theorem 6.16 that $\mathrm{Th}(\mathbb{F}_p[t], +, t^{\mathbb{N}}, \times\restriction t^{\mathbb{N}})$ is NIP for prime $p \geq 9$. Using methods similar to the proof of Theorem 6.13, we show in Theorem 6.17 that $\mathrm{Th}(\mathbb{Z}, +, <, d^{\mathbb{N}})$ is NIP.

We continue to assume that Γ is an infinite abelian group, that $F\colon \Gamma \to \Gamma$ is an injective endomorphism, and that Γ admits an F^r-spanning set for some $r > 0$.

6.1 EDP sets yield NIP expansions

To motivate our definition of F-EDP sets, we describe our approach to proving that $\mathrm{Th}(\Gamma, +, A)$ is NIP. We will use a result of Conant and Laskowski that gives a sufficient condition for an expansion of a weakly minimal structure to be NIP. In our context, their condition is in terms of the *induced structure* A_Γ of $(\Gamma, +)$ on A: this has domain A and a basic relation $X \cap A^m$ for each $X \subseteq \Gamma^m$ definable in $(\Gamma, +)$ with parameters from Γ.

Fact 6.5 ([10, Theorem 2.9])**.** *Suppose* Γ *is a weakly minimal abelian group and* $A \subseteq \Gamma$*. If* $\mathrm{Th}(A_\Gamma)$ *is NIP, then* $\mathrm{Th}(\Gamma, +, A)$ *is NIP.*

Our approach will be to show that A_Γ is interpretable in $\mathcal{N} = (\mathbb{N}, +)$, and hence that A_Γ is NIP. Our definition of F-EDP sets will be set up to guarantee that we can produce such an interpretation.

Before defining F-EDP sets, we introduce a convenient multi-index notation. If $\mathbf{s} = (\sigma_1, \ldots, \sigma_n)$ is a tuple of strings $\sigma_i \in \Gamma^*$ and $\mathbf{k} = (k_1, \ldots, k_n) \in \mathbb{N}^n$, we let $\mathbf{s}^{\mathbf{k}} = \sigma_1^{k_1}\sigma_2^{k_2}\cdots\sigma_n^{k_n} \in \Gamma^*$.

Definition 6.6. An *F-EDP* subset of Γ is a set of the form

$$[\mathbf{s}^{\varphi(\mathcal{N})}]_{F^r} := \{\, [\mathbf{s}^{\mathbf{k}}]_{F^r} : \mathcal{N} \models \varphi(\mathbf{k}) \,\} = \{\, [\sigma_1^{k_1}\cdots\sigma_n^{k_n}]_{F^r} : \mathcal{N} \models \varphi(k_1, \ldots, k_n) \,\}$$

for some $r > 0$, some tuple $\mathbf{s} = (\sigma_1, \ldots, \sigma_n)$ of strings over Γ, and some formula $\varphi(x_1, \ldots, x_n)$ in Presburger arithmetic.

The acronym "EDP" stands for "exponentially definable in Presburger arithmetic"; the idea is that the exponents of the strings come from some Presburger-definable set. We will prove in Proposition 6.11 that if $A \subseteq \Gamma$ is F-EDP then A_Γ is interpretable in $\mathcal{N}$. Roughly speaking, we will show that if $A = [\mathbf{s}^{\varphi(\mathcal{N})}]_{F^r}$ then the map $\varphi(\mathcal{N}) \to A$ given by $\mathbf{k} \mapsto [\mathbf{s}^{\mathbf{k}}]_{F^r}$ induces an intepretation of A_Γ in $\mathcal{N}$.

Proposition 6.7. *F-sparse sets are F-EDP.*

Proof. Suppose $A \subseteq \Gamma$ is F-sparse; say $A = [L]_{F^r}$ for some $r > 0$, some F^r-spanning set Σ, and some sparse $L \subseteq \Sigma^*$.

If L is simple sparse, say $L = u_0 v_1^* u_1 \cdots v_n^* u_n$, then we can let $\mathbf{s} = (u_0, v_1, u_1, \ldots, v_n, u_n)$ and $\varphi(x_0, y_1, x_1, \ldots, y_n, x_n)$ to be $x_0 = x_1 = \cdots = x_n = 1$. Then

$$A = [u_0 v_1^* u_1 \cdots v_n^* u_n]_{F^r} = \{ [u_0^{k_0} v_1^{\ell_1} u_1^{k_1} \cdots v_n^{\ell_n} u_n^{k_n}]_{F^r} : \mathcal{N} \models \varphi(k_0, \ell_1, k_1, \ldots, \ell_n, k_n) \} = [\mathbf{s}^{\varphi(\mathcal{N})}]_{F^r}.$$

In general, we know from Fact 3.3 that L is a finite union of simple sparse languages. The result then follows from the observation that

$$[\mathbf{s}^{\varphi(\mathcal{N})}]_{F^r} \cup [\mathbf{t}^{\psi(\mathcal{N})}]_{F^r} = [(\mathbf{st})^{\chi(\mathcal{N})}]_{F^r}$$

where $\chi(\mathbf{x}, \mathbf{y})$ is $(\mathbf{y} = \mathbf{0} \wedge \varphi(\mathbf{x})) \vee (\mathbf{x} = \mathbf{0} \wedge \psi(\mathbf{y}))$. $\square$ Proposition 6.7

The F-EDP sets include sets that aren't F-sparse, and indeed sets that aren't even F-automatic. For example, if $\Gamma = \mathbb{Z}^2$ and $F \colon \Gamma \to \Gamma$ is multiplication by 2, then

$$A := \left\{ \begin{pmatrix} 2^i \\ 2^{2i} \end{pmatrix} : 0 \neq i \in \mathbb{N} \right\} = \left\{ \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}^i \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}^{i-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]_F : 0 \neq i \in \mathbb{N} \right\}$$

is F-EDP. But the $(F, \{0,1\}^2)$-kernel of A is infinite: one can check that if $0 \neq i \in \mathbb{N}$ and we let $\sigma_i = \begin{pmatrix} 0 \\ 0 \end{pmatrix}^i \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, then $\begin{pmatrix} 0 \\ 2^j \end{pmatrix} \in A_{F, \sigma_i}$ if and only if $j = i - 1$. So A isn't F-automatic by Corollary 2.27.

The following is an analogue of Corollary 3.8 for F-EDP sets:

Proposition 6.8. *Suppose $A \subseteq \Gamma$ is F-EDP. Then there is $s_0 > 0$ such that if $s_0 \mid s$ then A can be written in the form $[\mathbf{a}^{\varphi(\mathcal{N})}]_{F^s}$ where $\mathbf{a}$ is a tuple of elements of Γ (i.e., strings of length 1).*

Proof. We first argue that sets of the desired form are closed under finite union, provided they share the same s. Indeed, given $[\mathbf{a}^{\varphi(\mathcal{N})}]_{F^s}$ and $[\mathbf{b}^{\psi(\mathcal{N})}]_{F^s}$, as we did at the end of the proof of Proposition 6.7 we can let $\chi(\mathbf{x}, \mathbf{y})$ be $(\mathbf{y} = \mathbf{0} \wedge \varphi(\mathbf{x})) \vee (\mathbf{x} = \mathbf{0} \wedge \psi(\mathbf{y}))$. Then

$$[\mathbf{a}^{\varphi(\mathcal{N})}]_{F^s} \cup [\mathbf{b}^{\varphi(\mathcal{N})}]_{F^s} = [(\mathbf{ab})^{\chi(\mathcal{N})}]_{F^s}$$

also takes the desired form, since all the elements of $\mathbf{ab}$ are again strings of length 1.

Write $A = [\mathbf{s}^{\varphi(\mathcal{N})}]_{F^r}$; by replacing F with F^r, we may assume $r = 1$. Write $\mathbf{s} = (\sigma_1, \dots, \sigma_n)$, and let $N = \operatorname{lcm}\{\,|\sigma_1|, \dots, |\sigma_n|\,\}$. Let $s_0 = N$, and suppose $s_0 \mid s$. Let $N_i = \frac{s}{|\sigma|_i}$. Then

$$\begin{aligned}
A &= \{\, [\sigma_1^{k_1} \cdots \sigma_n^{k_n}]_F : \mathcal{N} \models \varphi(k_1, \dots, k_n) \,\} \\
&= \bigcup_{j_1=0}^{N_1-1} \cdots \bigcup_{j_n=0}^{N_n-1} \underbrace{\{\, [\sigma_1^{j_1}(\sigma_1^{N_i})^{\ell_1} \cdots \sigma_n^{j_n}(\sigma_n^{N_n})^{\ell_n}]_F : \mathcal{N} \models \varphi(\ell_1 N_1 + j_1, \dots, \ell_n N_n + j_n) \,\}}_{A_{j_1,\dots,j_n}}
\end{aligned}$$

and each $|\sigma_i^{N_i}| = s$.

Fix $j_1, \dots, j_n$, and let $\chi(x_1, \dots, x_n)$ be $\varphi(x_1 N_1 + j_1, \dots, x_n N_n + j_n)$. Lemma 3.7 then yields $\alpha, a_1, \dots, a_n \in \Gamma$ such that

$$[\sigma_1^{j_1}(\sigma_1^{N_i})^{\ell_1} \cdots \sigma_n^{j_n}(\sigma_n^{N_n})^{\ell_n}]_F = \alpha + [a_1^{\ell_1} \cdots a_n^{\ell_n}]_{F^s}$$

for all $\ell_1, \dots, \ell_n \in \mathbb{N}$. Thus

$$\begin{aligned}
A_{j_1,\dots,j_n} &= \{\, \alpha + [a_1^{\ell_1} \cdots a_n^{\ell_n}]_{F^s} : \mathcal{N} \models \chi(\ell_1, \dots, \ell_n) \,\} \\
&= (A_{j_1,\dots,j_n} \cap \{\, \alpha \,\}) \cup \bigcup_{i=1}^{n} \{\, [(\alpha + a_i) a_i^{\ell_i} a_{i+1}^{\ell_{i+1}} \cdots a_n^{\ell_n}]_{F^s} : \mathcal{N} \models \chi(0, \dots, 0, \ell_i + 1, \ell_{i+1}, \dots, \ell_n) \,\}.
\end{aligned}$$

So $A_{j_1,\dots,j_n}$ is a union of sets of the desired form, and hence as we argued at the beginning of the proof $A_{j_1,\dots,j_n}$ takes the desired form. So, since

$$A = \bigcup_{j_1=0}^{N_1-1} \cdots \bigcup_{j_n=0}^{N_n-1} A_{j_1,\dots,j_n},$$

we get that A itself takes the desired form. $\square$ Proposition 6.8

Remark 6.9. We saw in the proof of Proposition 6.7 that as long as two F-EDP sets are defined using the same power of F, their union is again F-EDP. A consequence of Proposition 6.8 is that we can prove closure under union in general. Indeed, if $A, B \subseteq \Gamma$ are F-EDP, then by Proposition 6.8 there is $s > 0$ such that we can write $A = [\mathbf{a}^{\varphi(\mathcal{N})}]_{F^s}$ and $B = [\mathbf{b}^{\psi(\mathcal{N})}]_{F^s}$. Since they use the same power F^s of F, it follows from our earlier argument that $A \cup B$ is F-EDP.

We now turn towards proving that the map $\mathbf{k} \mapsto [\mathbf{a}^{\mathbf{k}}]_F$ defines an interpretation of A_Γ in $\mathcal{N}$. We will reduce this to the following lemma:

Lemma 6.10. *Suppose Σ is any finite alphabet, and $L \subseteq (\Sigma^m)^*$ is regular. If $\mathbf{a}$ is a tuple of letters from Σ, then the relation*

$$\left\{ \begin{pmatrix} \mathbf{k}_1 \\ \vdots \\ \mathbf{k}_m \end{pmatrix} \in \mathbb{N}^{m|\mathbf{a}|} : \begin{pmatrix} \mathbf{a}^{\mathbf{k}_1} \\ \vdots \\ \mathbf{a}^{\mathbf{k}_m} \end{pmatrix} \in L \right\}$$

is definable in $\mathcal{N}$.

Here, and for the rest of the chapter, we interpret the statement $\begin{pmatrix} \mathbf{a}^{\mathbf{k}_1} \\ \vdots \\ \mathbf{a}^{\mathbf{k}_m} \end{pmatrix} \in L$ to include the assertion that $|\mathbf{a}^{\mathbf{k}_1}| = \cdots = |\mathbf{a}^{\mathbf{k}_m}|$. Indeed, otherwise we can't interpret $\begin{pmatrix} \mathbf{a}^{\mathbf{k}_1} \\ \vdots \\ \mathbf{a}^{\mathbf{k}_m} \end{pmatrix}$ as an element of $(\Sigma^m)^*$, and the statement doesn't make sense.

Proof of Lemma 6.10. Fix a DFA $M = (\Sigma, Q, q_0, \Omega, \delta)$ recognizing L. Then

$$\begin{pmatrix} \mathbf{a}^{\mathbf{k}_1} \\ \vdots \\ \mathbf{a}^{\mathbf{k}_m} \end{pmatrix} \in L \iff \bigvee_{q \in \Omega} \delta\left(q_0, \begin{pmatrix} \mathbf{a}^{\mathbf{k}_1} \\ \vdots \\ \mathbf{a}^{\mathbf{k}_m} \end{pmatrix} \right) = q,$$

where again we interpret the latter statement as including the assertion that $|\mathbf{a}^{\mathbf{k}_1}| = \cdots = |\mathbf{a}^{\mathbf{k}_m}|$, so that it is well-defined. It thus suffices to show that the relation

$$\delta\left(q_0, \begin{pmatrix} \mathbf{a}^{\mathbf{k}_1} \\ \vdots \\ \mathbf{a}^{\mathbf{k}_m} \end{pmatrix} \right) = q$$

is definable in $\mathcal{N}$ for each q. For our induction to work properly, we will instead prove the following stronger statement: for every $q, q' \in Q$ and all tuples $\mathbf{a}_1, \ldots, \mathbf{a}_m$ of elements of Σ, the relation R defined by

$$\delta\left(q, \begin{pmatrix} \mathbf{a}_1^{\mathbf{k}_1} \\ \vdots \\ \mathbf{a}_m^{\mathbf{k}_m} \end{pmatrix} \right) = q'$$

is definable in $\mathcal{N}$.

We apply induction on $|\mathbf{a}_1|\cdots|\mathbf{a}_m|$. For the base case $|\mathbf{a}_1|\cdots|\mathbf{a}_m| = 0$, we get that some $|\mathbf{a}_i| = 0$. Then since $\mathbf{a}_i^{\mathbf{k}_i} = \varepsilon$, the only way for our length condition $|\mathbf{a}_1^{\mathbf{k}_1}| = \cdots = |\mathbf{a}_m^{\mathbf{k}_m}|$ to be satisfied is if each $\mathbf{k}_j = \mathbf{0}$. So R is either empty or only contains the zero tuple, and both of these are definable in $\mathcal{N}$.

For the induction step, we may assume that no $|\mathbf{a}_i| = 0$. Write $\mathbf{a}_i = (a_{i1}, \ldots, a_{in_i})$ and $\mathbf{k}_i = (k_{i1}, \ldots, k_{in_i})$. For $1 \le i \le n$ we let

$$R_i = \left\{ \begin{pmatrix} \mathbf{k}_1 \\ \vdots \\ \mathbf{k}_m \end{pmatrix} \in R : k_{i1} = \min\{ k_{11}, \ldots, k_{m1} \} \right\}.$$

We show that R_i is definable in $\mathcal{N}$ for each i; since R is the union of the R_i, this will be sufficient to show that R is definable in $\mathcal{N}$.

For ease of notation, we check R_1. Suppose then that k_{11} is minimum among the k_{i1}. In this case, we can decompose

$$\begin{pmatrix} \mathbf{a}_1^{\mathbf{k}_1} \\ \vdots \\ \mathbf{a}_m^{\mathbf{k}_m} \end{pmatrix} = \begin{pmatrix} a_{11}^{k_{11}} \\ a_{21}^{k_{21}} \\ \vdots \\ a_{m1}^{k_{11}} \end{pmatrix} \begin{pmatrix} (\mathbf{a}_1')^{\mathbf{k}_1'} \\ a_{21}^{k_{21}-k_{11}}(\mathbf{a}_2')^{\mathbf{k}_2'} \\ \vdots \\ a_{m1}^{k_{m1}-k_{11}}(\mathbf{a}_m')^{\mathbf{k}_m'} \end{pmatrix},$$

where $\mathbf{a}_i' = (a_{i2}, \ldots, a_{in_i})$, and likewise with $\mathbf{k}_i'$. Hence if we let $q_k = \delta\left(q, \begin{pmatrix} a_{11}^k \\ \vdots \\ a_{m1}^k \end{pmatrix}\right)$ for $k \in \mathbb{N}$, then

$$\delta\left(q, \begin{pmatrix} \mathbf{a}_1^{\mathbf{k}_1} \\ \vdots \\ \mathbf{a}_m^{\mathbf{k}_m} \end{pmatrix}\right) = \delta\left(q_{k_{11}}, \begin{pmatrix} (\mathbf{a}_1')^{\mathbf{k}_1'} \\ a_{21}^{k_{21}-k_{11}}(\mathbf{a}_2')^{\mathbf{k}_2'} \\ \vdots \\ a_{m1}^{k_{m1}-k_{11}}(\mathbf{a}_m')^{\mathbf{k}_m'} \end{pmatrix}\right).$$

But since $|\mathbf{a}_1'||\mathbf{a}_2|\cdots|\mathbf{a}_m| < |\mathbf{a}_1||\mathbf{a}_2|\cdots|\mathbf{a}_m|$, and since subtraction is definable in $\mathcal{N}$, it follows from the induction hypothesis that for fixed $\theta \in Q$ the relation

$$\delta\left(\theta, \begin{pmatrix} (\mathbf{a}_1')^{\mathbf{k}_1'} \\ a_{21}^{k_{21}-k_{11}}(\mathbf{a}_2')^{\mathbf{k}_2'} \\ \vdots \\ a_{m1}^{k_{m1}-k_{11}}(\mathbf{a}_m')^{\mathbf{k}_m'} \end{pmatrix}\right) = q'$$

is definable in $\mathcal{N}$ (again assuming that $k_{11} \leq k_{i1}$ for all i). Moreover since there are only finitely many states in M we get that q_k is ultimately periodic in k, and hence that the statement $q_k = \theta$ is an $\mathcal{N}$-definable property of k for any fixed $\theta \in Q$. Putting these together, we find that R_1 is equivalent to

$$k_{11} = \min\{\, k_{11}, \ldots, k_{m1} \,\} \wedge \bigvee_{\theta \in Q} \left(q_{k_{11}} = \theta \wedge \delta \left(\theta, \begin{pmatrix} (\mathbf{a}'_1)^{\mathbf{k}'_1} \\ a_{21}^{k_{21}-k_{11}} (\mathbf{a}'_2)^{\mathbf{k}'_2} \\ \vdots \\ a_{m1}^{k_{m1}-k_{11}} (\mathbf{a}'_m)^{\mathbf{k}'_m} \end{pmatrix} \right) = q' \right),$$

and is thus definable in $\mathcal{N}$.

The case R_i for $i > 1$ is similar. So each R_i is definable in $\mathcal{N}$, and thus so is R. $\square$ Lemma 6.10

Proposition 6.11. *If $A \subseteq \Gamma$ is F-EDP then A_Γ is interpretable in $\mathcal{N}$.*

Proof. Recall from Proposition 2.17 that if Γ admits an F^s-spanning set then it admits an F^t-spanning set whenever $s \mid t$. Hence by Proposition 6.8 we may assume A takes the form $[\mathbf{a}^{\varphi(\mathcal{N})}]_{F^s}$ for some tuple $\mathbf{a} = (a_1, \ldots, a_n)$ of elements of Γ, some formula φ in Presburger arithmetic, and some $s > 0$ for which there exists an F^s-spanning set Σ. By Proposition 2.15 we may assume that each $a_i \in \Sigma$.

Consider the map $\Phi \colon \varphi(\mathcal{N}) \to A$ given by $\mathbf{k} \mapsto [\mathbf{a}^{\mathbf{k}}]_{F^s}$. It is clear that Φ is surjective. So to show that Φ induces an interpretation of A_Γ in $\mathcal{N}$, we need only show that given a basic relation (including equality) $X \cap A^m$ of A_Γ, its preimage $\Phi^{-1}(X \cap A^m)$ is definable in $\mathcal{N}$.

Since $X \subseteq \Gamma^m$ is definable in $(\Gamma, +)$, Corollary 2.35 yields that X is F-automatic. Moreover we saw in Example 2.13 (2) that Σ^m is an F^s-spanning set for Γ^m. So by Corollary 2.27 we get that X is (F^s, Σ^m)-automatic, and thus

$$L := \left\{ \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_m \end{pmatrix} \in (\Sigma^m)^* : \begin{pmatrix} [\sigma_1]_{F^s} \\ \vdots \\ [\sigma_m]_{F^s} \end{pmatrix} \in X \right\}$$

is regular. So Lemma 6.10 applies, and we can conclude that the relation $\begin{pmatrix} \mathbf{a}^{\mathbf{k}_1} \\ \vdots \\ \mathbf{a}^{\mathbf{k}_m} \end{pmatrix} \in L$ is definable in $\mathcal{N}$.

We would like to use this to show that $\Phi^{-1}(X \cap A^m)$ is definable in $\mathcal{N}$. Unfortunately, given $\mathbf{k}_1, \ldots, \mathbf{k}_m \in \varphi(\mathcal{N})$, it doesn't necessarily hold that $|\mathbf{a}^{\mathbf{k}_1}| = \cdots = |\mathbf{a}^{\mathbf{k}_m}|$; so $\begin{pmatrix} \mathbf{a}^{\mathbf{k}_1} \\ \vdots \\ \mathbf{a}^{\mathbf{k}_m} \end{pmatrix} \in L$ may fail simply because the lengths don't match up.

We can remedy this by adding trailing zeroes. If we choose $\ell_1, \ldots, \ell_m \in \mathbb{N}$ such that $|\mathbf{a}^{\mathbf{k}_1} 0^{\ell_1}| = \cdots = |\mathbf{a}^{\mathbf{k}_m} 0^{\ell_m}|$, then we can regard

$$\begin{pmatrix} \mathbf{a}^{\mathbf{k}_1} 0^{\ell_1} \\ \vdots \\ \mathbf{a}^{\mathbf{k}_m} 0^{\ell_m} \end{pmatrix} = \begin{pmatrix} (\mathbf{a}0)^{\mathbf{k}_1 \ell_1} \\ \vdots \\ (\mathbf{a}0)^{\mathbf{k}_m \ell_m} \end{pmatrix}$$

as a bona fide element of $(\Sigma^m)^*$, where $\mathbf{a}0$ denotes the tuple whose first entries are the entries of $\mathbf{a}_i$ and whose last entry is 0, and likewise with $\mathbf{k}_i \ell_i$. Hence

$$\begin{pmatrix} \Phi(\mathbf{k}_1) \\ \vdots \\ \Phi(\mathbf{k}_m) \end{pmatrix} \in X \iff \begin{pmatrix} [(\mathbf{a}0)^{\mathbf{k}_1 \ell_1}]_{F^s} \\ \vdots \\ [(\mathbf{a}0)^{\mathbf{k}_m \ell_m}]_{F^s} \end{pmatrix} \in X \iff \begin{pmatrix} (\mathbf{a}0)^{\mathbf{k}_1 \ell_1} \\ \vdots \\ (\mathbf{a}0)^{\mathbf{k}_m \ell_m} \end{pmatrix} \in L.$$

Of course, this only holds if we choose $\ell_1, \ldots, \ell_n$ to be such that $|(\mathbf{a}0)^{\mathbf{k}_1 \ell}| = \cdots = |(\mathbf{a}0)^{\mathbf{k}_m \ell_m}|$. If we existentially quantify the $\ell_1, \ldots, \ell_n$, however, we find that

$$\begin{pmatrix} \Phi(\mathbf{k}_1) \\ \vdots \\ \Phi(\mathbf{k}_m) \end{pmatrix} \in X \iff \exists \ell_1, \ldots, \ell_m \left(\begin{pmatrix} (\mathbf{a}0)^{\mathbf{k}_1 \ell_1} \\ \vdots \\ (\mathbf{a}0)^{\mathbf{k}_m \ell_m} \end{pmatrix} \in L \right).$$

Hence applying Lemma 6.10 to $\mathbf{a}0$ and L, it follows that $\Phi^{-1}(X \cap A^m)$ is indeed definable in $\mathcal{N}$. □ Proposition 6.11

Remark 6.12. Let $\mathscr{G}$ be the structure with domain Γ and a basic relation for every F-automatic set in every Γ^m. The above proof in fact shows the following stronger result: that the map $\Phi \colon \varphi(\mathcal{N}) \to A$ given by $\mathbf{k} \mapsto [\mathbf{a}^{\mathbf{k}}]_{F^s}$ induces an interpretation of the induced structure $A_{\mathscr{G}}$ in $\mathcal{N}$.

Our theorem now quickly follows:

Theorem 6.13. *If* Γ *is weakly minimal and* $A \subseteq \Gamma$ *is* F*-EDP then* $\mathrm{Th}(\Gamma, +, A)$ *is NIP.*

Proof. By Fact 6.5 it suffices to show that A_Γ is NIP. But $\mathrm{Th}(\mathcal{N})$ is NIP, and Proposition 6.11 yields that A_Γ is interpretable in $\mathcal{N}$. So, since NIP is defined by the absence of a definable relation with the independence property, it follows that A_Γ is indeed NIP. □ Theorem 6.13

Corollary 6.14. *If Γ is weakly minimal and $A \subseteq \Gamma$ is F-sparse then $\mathrm{Th}(\Gamma, +, A)$ is NIP.*

6.2 Examples of expansions by F-EDP sets

In this section we give two applications of Theorem 6.13.

Theorem 6.15. *If $d \geq 2$ then $\mathrm{Th}(\mathbb{Z}, +, d^{\mathbb{N}}, \times\restriction d^{\mathbb{N}})$ is NIP, where $\times\restriction d^{\mathbb{N}}$ is the graph of multiplication restricted to $d^{\mathbb{N}}$.*

Our primary difficulty is that we wish to expand by a subset of $\mathbb{Z}^3$, whereas Theorem 6.13 only applies to subsets of $\mathbb{Z}$. We will need to encode our expansion in a subset of $\mathbb{Z}$.

Proof. We first do the case $d \geq 8$. We let $F \colon \mathbb{Z} \to \mathbb{Z}$ be multiplication by d.

We choose F-EDP $A \subseteq \mathbb{Z}$ such that both $d^{\mathbb{N}}$ and $B := \{ d^i + 2d^j + 4d^{i+j} : 0 < i < j \}$ are definable in $(\mathbb{Z}, +, A)$. We let

$$A = \{ da : a \in d^{\mathbb{N}} \} \cup \{ 1 + db : b \in B \} = \{ [0^{n+1}1]_d : n \in \mathbb{N} \} \cup \{ [10^k 10^\ell 20^{k-1}4]_d : k > 0, \ell \geq 0 \}.$$

So A is a union of F-EDP sets, and is thus F-EDP. Moreover $d^{\mathbb{N}}$ and B are indeed definable in A: they are defined by $dx \in A$ and $1 + dx \in A$, respectively.

It is clear from Fact 6.4 that $(\mathbb{Z}, +)$ is weakly minimal. Theorem 6.13 then yields that $\mathrm{Th}(\mathbb{Z}, +, A)$ is NIP, and hence so is $\mathrm{Th}(\mathbb{Z}, +, d^{\mathbb{N}}, B)$. It then suffices to show that $\times\restriction d^{\mathbb{N}}$ is definable in $(\mathbb{Z}, +, d^{\mathbb{N}}, B)$.

I claim that if $d^i, d^j, d^k \in d^{\mathbb{N}}$ satisfy $1 < d^i < d^j$ then $d^i \cdot d^j = d^k$ if and only if $d^i + 2d^j + 4d^k \in B$. The left-to-right direction is clear; for the right-to-left, suppose we have $s > r > 0$ such that $d^i + 2d^j + 4d^k = d^r + 2d^s + 4d^{r+s}$. The usual base-$d$ representation for $d^r + 2d^s + 4d^{r+s}$ is $0^r 10^{s-r-1}20^{r-1}4$. At this point, examining the possible base-d representations of $d^i + 2d^j + 4d^k$, and using the uniqueness of base-d representations up to trailing zeroes, we find that $i = r$, $j = s$, and $k = r + s = i + j$; so $d^i \cdot d^j = d^k$, as desired.

It follows that for $x \neq y$ both $\neq 1$, we can define $\times\restriction d^{\mathbb{N}}$ by

$$x, y, z \in d^{\mathbb{N}} \wedge (x + 2y + 4z \in B \vee y + 2x + 4z \in B).$$

It remains to deal with the case where one of x, y is 1, and the case where $x = y \neq 1$. In the former, $\times\restriction d^{\mathbb{N}}$ is definable by $x, y, z \in d^{\mathbb{N}} \wedge z = y$ in the case $x = 1$, and by $x, y, z \in d^{\mathbb{N}} \wedge z = x$ in the case $y = 1$. In the latter case, we observe that $x^2 = z$ for $x \in d^{\mathbb{N}} \setminus \{ 1 \}$ if and only if $(d^{-1}x)(dx) = z$, and $d^{-1}x < dx$; so in this case $\times\restriction d^{\mathbb{N}}$ is definable by $x, y, z \in d^{\mathbb{N}} \wedge d^{-1}x + 2dx + 4z \in B$.

Putting these together, we find that $\times\restriction d^{\mathbb{N}}$ is definable with parameters in $(\mathbb{Z}, +, d^{\mathbb{N}}, B)$ by the formula

$$\begin{aligned} x, y, z \in d^{\mathbb{N}} \wedge \Big(& x \neq y \wedge (x, y \neq 1) \wedge (x + 2y + 4z \in B \vee y + 2x + 4z \in B) \vee \\ & (x = 1 \wedge z = y) \vee \\ & (y = 1 \wedge z = x) \vee \\ & (x = y \neq 1 \wedge d^{-1}x + 2dx + 4z \in B) \Big). \end{aligned}$$

It follows that $(\mathbb{Z}, +, d^{\mathbb{N}}, \times\restriction d^{\mathbb{N}})$ has an NIP theory.

The case $d < 8$ follows from the case $d \geq 8$, since $(\mathbb{Z}, +, d^{\mathbb{N}}, \times\restriction d^{\mathbb{N}})$ is a reduct of $(\mathbb{Z}, +, (d')^{\mathbb{N}}, \times\restriction (d')^{\mathbb{N}})$ whenever d' is a power of d. □ Theorem 6.15

Theorem 6.16. *If $p \geq 9$ is prime then $\mathrm{Th}(\mathbb{F}_p[t], +, t^{\mathbb{N}}, \times\restriction t^{\mathbb{N}})$ is NIP.*

Proof. We let $F \colon \mathbb{F}_p[t] \to \mathbb{F}_p[t]$ be $f \mapsto tf$. We enumerate the elements of $\mathbb{F}_p[t] \cong \mathbb{Z}/p\mathbb{Z}$ as $\bar{i}$ for $0 \leq i < p$. It follows from Example 2.13 (4) that $\mathbb{F}_p[t]$ admits an F-spanning set; so Theorem 6.13 applies.

As in Theorem 6.15, we choose F-EDP $A \subseteq \mathbb{F}[t]$ such that both $t^{\mathbb{N}}$ and $B := \{ t^i + \bar{2}t^j + \bar{4}t^{i+j} : 0 < i < j \}$ are definable in $(\mathbb{F}_p[t], +, A)$. In this context, however, we can no longer use congruences to encode two sets into one, since congruence modulo any fixed d is either vacuous or degenerate; so our amalgamation will have to be a bit more subtle.

We let

$$\begin{aligned} A &= t^{\mathbb{N}} \cup \bar{2}t^{\mathbb{N}} \cup B \\ &= \{ [(\bar{0})^n \bar{1}]_F : n \in \mathbb{N} \} \cup \{ [(\bar{0})^n \bar{2}]_F : n \in \mathbb{N} \} \cup \{ [(\bar{0})^k \bar{1} (\bar{0})^\ell \bar{2} (\bar{0})^{k-1} \bar{4}]_F : k > 0, \ell \geq 0 \}. \end{aligned}$$

So A is a union of F-EDP sets, and is thus F-EDP. Moreover $t^{\mathbb{N}}$ and B are indeed definable in A: the former is defined by $x \in A \wedge 2x \in A$, and the latter by $x \in A \wedge x \notin t^{\mathbb{N}} \wedge x \notin 2t^{\mathbb{N}}$. (Note since $p \geq 9$ that if $f \in B$ then $2f \notin A$; this can be seen by noting that the leading coefficient of $2f$ is $\bar{8}$, which cannot appear as a leading coefficient in A.)

From here, it follows as in Theorem 6.15 that outside of the case $x = y \neq \overline{1}$ we can define $\times\restriction t^{\mathbb{N}}$ in $(\mathbb{F}_p[t], +, t^{\mathbb{N}}, B)$ (with parameters). A subtlety arises in the case $x = y \neq \overline{1}$: unlike in Theorem 6.15, the map $t^i \mapsto t^{i+1}$ isn't definable in the group structure, so we can't simply use the formula

$$x, y, z \in t^{\mathbb{N}} \wedge t^{-1}x + 2tx + 4z \in B$$

as we did before. However, we can make use of the fact that $\times\restriction t^{\mathbb{N}}$ is definable outside of this case: note if $t^i, t^j, t^k \in t^{\mathbb{N}}$ and $t^i = t^j \neq t$ then $t^i \cdot t^j = t^k$ if and only if $t^k \neq t$ and $t^i \cdot (t \cdot t^i) = t \cdot t^k$. Since $t^i, t^k \neq t$, it follows that $t \cdot t^i, t \cdot t^k$ fall into the already-defined case of $\times\restriction t^{\mathbb{N}}$, as does $t^i \cdot (t \cdot t^i)$; so $\times\restriction t^{\mathbb{N}}$ is definable in $(\mathbb{F}_p[t], +, t^{\mathbb{N}}, B)$ in the case $x = y \neq t$. This only leaves the case $x = y = t$; for this we can use the formula $x = y = t \wedge z = t^2$.

It thus follows that $(\mathbb{F}_p[t], +, t^{\mathbb{N}}, \times\restriction t^{\mathbb{N}})$ is definable with parameters in $(\mathbb{F}_p[t], +, t^{\mathbb{N}}, B)$, and thus in $(\mathbb{F}_p[t], +, A)$. But $(\mathbb{F}_p[t], +)$ is weakly minimal: one can check that

$$n\mathbb{F}_p[t] = \begin{cases} \mathbb{F}_p[t] & \text{if } p \nmid n \\ \{\,0\,\} & \text{else} \end{cases}$$
$$\{\, a \in \mathbb{F}_p[t] : na = 0 \,\} = \begin{cases} \{\,0\,\} & \text{if } p \nmid n \\ \mathbb{F}_p[t] & \text{else.} \end{cases}$$

So Theorem 6.13 yields that $\mathrm{Th}(\mathbb{F}_p[t], +, A)$ is NIP, and hence that $\mathrm{Th}(\mathbb{F}_p[t], +, t^{\mathbb{N}}, \times\restriction t^{\mathbb{N}})$ is NIP. □ Theorem 6.16

6.3 $(\mathbb{Z}, +, <, d^{\mathbb{N}})$ is NIP

The following theorem is another example of an NIP expansion of $(\mathbb{Z}, +)$. Unlike the previous two examples, this one doesn't directly follow from Theorem 6.13, though the argument is similar.

Theorem 6.17. *If $d \geq 2$ then $\mathrm{Th}(\mathbb{Z}, +, <, d^{\mathbb{N}})$ is NIP.*

This was proven by Lambotte and Point in [17, Corollary 2.34]; we use automata to give an alternate proof.

The difficulty here is that it seems unlikely that we'll be able to encode $<$ in an F-EDP set. So rather than view $(\mathbb{Z}, +, d^{\mathbb{N}}, <)$ as the expansion of $(\mathbb{Z}, +)$ by $d^{\mathbb{N}}$ and $<$, we instead view it as the expansion of $(\mathbb{Z}, +, <)$ by $d^{\mathbb{N}}$. Of course, $\mathrm{Th}(\mathbb{Z}, +, <)$ isn't even stable, let alone weakly minimal; so the methods of Conant and Laskowski don't apply. We instead make use of a more general result of Chernikov and Simon. This result replaces the weak minimality hypothesis with the assumption that A is *bounded* in $(\Gamma, +, <)$:

Definition 6.18. Suppose L is a first-order signature; let $L_P = L \cup \{ P \}$ for some unary predicate symbol P not appearing in L. A *bounded* L_P-formula is a formula of the form $(Q_1 x_1 \in P) \cdots (Q_n x_n \in P)\varphi$ for some quantifiers $Q_1, \ldots, Q_n$ and some L-formula φ. If M is an L-structure and $A \subseteq M$, we say that A is *bounded* in M if every L_P-formula is equivalent in $\mathrm{Th}(M, A)$ to a bounded L_P-formula.

Fact 6.19 ([8, Corollary 2.5]). *Suppose M is an L-structure and $A \subseteq M$. If $\mathrm{Th}(M)$ is NIP, A is bounded in M, and the theory of the induced structure A_M is NIP, then $\mathrm{Th}(M, A)$ is NIP.*[1]

We look to apply this with $M = (\mathbb{Z}, +, <)$ and $A = d^{\mathbb{N}}$. Checking that A is bounded in $(\mathbb{Z}, +, <)$ looks intimidating at first glance; happily, we have the following result of Point to fall back on.

Fact 6.20 ([19, Propositions 9 and 11]). $\mathrm{Th}\Big(\mathbb{N}, +, \dot{-}, <, \big(\frac{\cdot}{n}\big)_{n>0}, \lambda, S, S^{-1}\Big)$ *admits quantifier elimination, where*

- $a \dot{-} b = \max\{ 0, a - b \}$;
- $\frac{\cdot}{n}$ *denotes integer division rounded down;*
- $\lambda(0) = 0$, *and if* $a > 0$ *then* $\lambda(a) = \max\{ b \in d^{\mathbb{N}} : b \leq a \}$;
- $S(d^n) = d^{n+1}$, *and* $S(a) = a$ *for all other* a; *and*
- $S^{-1}(d^{n+1}) = d^n$, *and* $S^{-1}(a) = a$ *for all other* a.

Proof of Theorem 6.17. To more easily make use of Fact 6.20, we instead show that $\mathrm{Th}(\mathbb{N}, +, d^{\mathbb{N}})$ is NIP; since $(\mathbb{Z}, +, <, d^{\mathbb{N}})$ is interpretable in $(\mathbb{N}, +, d^{\mathbb{N}})$, this will suffice.

We look to apply Fact 6.19. We have seen that $\mathrm{Th}(\mathbb{N}, +) = \mathrm{Th}(\mathcal{N})$ is NIP.

[1] In fact the theorem is stronger than stated here; rather than requiring that the full induced structure of M on A be NIP, they only require that the structure on A induced by the 0-definable sets in M be NIP.

We need to show that the induced structure $(d^{\mathbb{N}})_{\mathcal{N}}$ is NIP. Since $\mathbb{N}$ is definable in $(\mathbb{Z},+,<)$, we get that $(d^{\mathbb{N}})_{\mathcal{N}}$ is a reduct of $(d^{\mathbb{N}})_{(\mathbb{Z},+,<)}$; so it suffices to show that $(d^{\mathbb{N}})_{(\mathbb{Z},+,<)}$ is NIP. Note that

$$\left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{Z}^2 : a < b \right\}$$

is d-automatic in $\mathbb{Z}$. Indeed, one can check that its $(d, \{ 0, \ldots, d-1 \})$-kernel contains only

$$\left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{Z}^2 : a < b \right\} \text{ and } \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{Z}^2 : a \leq b \right\},$$

and is thus finite, at which point Corollary 2.27 yields d-automaticity. Hence if we let $\mathscr{Z}$ be the structure with domain $\mathbb{Z}$ and a basic relation for every d-automatic subset of every $\mathbb{Z}^m$, we get from Corollary 2.35 that $(\mathbb{Z},+,<)$ is a reduct of $\mathscr{Z}$. Hence $(d^{\mathbb{N}})_{(\mathbb{Z},+,<)}$ is a reduct of $(d^{\mathbb{N}})_{\mathscr{Z}}$, and it suffices to show that $(d^{\mathbb{N}})_{\mathscr{Z}}$ is NIP. But this follows from Remark 6.12, since $d^{\mathbb{N}} = \{ [0^i 1]_d : i \in \mathbb{N} \}$ is F-EDP.

So the induced structure $(d^{\mathbb{N}})_{\mathcal{N}}$ is NIP. It remains to show that $d^{\mathbb{N}}$ is bounded in $\mathcal{N}$. Let $L = \{ + \}$ and $L' = \left\{ +, \dot{-}, \left(\frac{\cdot}{n}\right)_{n>0}, \lambda, S, S^{-1} \right\}$; so by Fact 6.20 every L_P formula is equivalent to a quantifier-free L'-formula (since $d^{\mathbb{N}}$ is definable in $\left(\mathbb{N}, +, \dot{-}, <, \left(\frac{\cdot}{n}\right)_{n>0}, \lambda, S, S^{-1}\right)$ by the formula $x \neq 0 \wedge \lambda(x) = x$). We show that any quantifier-free L'-formula φ is equivalent to a bounded L_P-formula. We do so by induction on the number of occurrences of λ, S, or S^{-1} in φ. For the base case, if φ contains no such occurrences, then φ is a quantifier-free formula in the signature $\left\{ +, \dot{-}, \left(\frac{\cdot}{n}\right)_{n>0} \right\}$; so φ is equivalent to an L-formula, which is a bounded L_P-formula. For the induction step, suppose φ contains a term of the form $\lambda(t)$ for some L'-term t that contains no occurrences of λ, S, or S'; so φ takes the form $\varphi'(\lambda(t), \mathbf{x})$ for some quantifier-free L'-formula φ'. Then recalling the definition of λ we find that φ is equivalent to

$$\begin{aligned} &(t = 0 \wedge \varphi'(0, \mathbf{x})) \vee (t \neq 0 \wedge (\exists y \in d^{\mathbb{N}})(y \leq t \wedge (\forall z \in d^{\mathbb{N}})(z \leq t \to z \leq y) \wedge \varphi'(y, \mathbf{x}))) \\ \equiv\ &(\exists y \in d^{\mathbb{N}})(\forall z \in d^{\mathbb{N}})\Big(\underbrace{(t = 0 \wedge \varphi'(0, \mathbf{x}))}_{\chi_1} \vee \underbrace{(t \neq 0 \wedge (y \leq t \wedge (z \leq t \to z \leq y) \wedge \varphi'(y, \mathbf{x})))}_{\chi_2}\Big) \end{aligned}$$

(where by possibly changing variables we have assumed that $\mathbf{x}, y, z$ are all distinct). But each of χ_1, χ_2 are both quantifier-free L'-formulas that contain one fewer occurrence of λ, S, or S^{-1} than φ does; so by the induction hypothesis they are equivalent to bounded L_P-formulas. So φ is as well.

Similarly if φ takes the form $\varphi'(S(t), \mathbf{x})$ then φ is equivalent to

$$(\exists w \in d^{\mathbb{N}})(t = w \wedge \varphi'(dw, \mathbf{x})) \vee (\forall w \in d^{\mathbb{N}})(t \neq w \wedge \varphi'(t, \mathbf{x})),$$

and if φ takes the form $\varphi'(S^{-1}(t), \mathbf{x})$ then φ is equivalent to

$$(\exists w \in d^{\mathbb{N}})(t = dw \wedge \varphi'(w, \mathbf{x})) \vee (\forall w \in d^{\mathbb{N}})(t \neq dw \wedge \varphi'(t, \mathbf{x})).$$

In both of these cases we can argue as we did in the first case that φ is equivalent to a bounded L_P-formula. Note moreover that if φ contains an occurrence of $\lambda(t)$, $S(t)$, or $S^{-1}(t)$ then it must contain one in which t does not have an occurrence of λ, S, or S^{-1}; i.e., we fall into one of the above three cases. So any such φ is equivalent to a bounded L_P-formula, and by induction we're done.

So $\mathrm{Th}(\mathcal{N})$ is NIP, $\mathrm{Th}((d^{\mathbb{N}})_{\mathcal{N}})$ is NIP, and $d^{\mathbb{N}}$ is bounded in $(\mathbb{N}, +)$. So Fact 6.19 yields that $\mathrm{Th}(\mathbb{N}, +, d^{\mathbb{N}})$ is NIP. □ Theorem 6.17

6.4 Bibliographical notes

NIP is a standard notion; see [21]. Theorem 6.17 was conjectured in [4], and was proven by Lambotte and Point in [17] using different methods from the ones appearing here. All other results in this chapter are original.

Chapter 7

Future research

We list some possible avenues of future research.

- In classical automata theory, Christol's theorem (see e.g. [3, Theorem 12.2.5]) tells us that if p is prime then a power series $a_0 + a_1 t + \cdots$ with coefficients in $\mathbb{Z}/p\mathbb{Z}$ is algebraic over $(\mathbb{Z}/p\mathbb{Z})(t)$ if and only if $\{ i \in \mathbb{N} : a_i = j \}$ is p-automatic for all $j \in \mathbb{Z}/p\mathbb{Z}$. Is there a generalization of this to the setting of F-automatic sets?

- Another result of classical automata theory is Cobham's theorem ([3, Theorem 11.2.1]): that if $A \subseteq \mathbb{N}$ is both d- and d'-automatic for d, d' multiplicatively independent (i.e., $d^i \neq (d')^j$ for any $i, j > 0$) then the characteristic function of A is ultimately periodic. Is there some analogue for sets $A \subseteq \Gamma$ that are both F- and F'-automatic, under some assumptions on F and F'?

- In the classical context, we say that a sequence $(a_n)_{n\in\mathbb{N}}$ with values in some finite set Δ is *d-automatic* if $\{ i \in \mathbb{N} : a_i = \delta \}$ is a d-automatic set for every $\delta \in \Delta$. In [2], Allouche and Shallit use kernels to give a general notion of d-automaticity for sequences with values in a (possibly infinite) ring, which they call *d-regularity*; this can be further extended to sequences taking values in an abelian group, as described in [3, Chapter 16]. Can we make a similar (coherent) definition in our setting? That is, if G is an abelian group, can we sensibly define a notion of *F-regularity* of a map $\Gamma \to G$?

- A link between automata theory and logic is that $A \subseteq \mathbb{N}$ is d-automatic if and only if it is definable in $(\mathbb{N}, +, V_d)$, where $V_d(a)$ is the largest power of d dividing a (and $V_d(0) = 1$); for more details on this see [7, Theorem 4.1]. Can we make a similar statement about F-automatic sets being the definable sets in some natural expansion of $(\Gamma, +)$?

- Can Theorem 5.17 be extended to a characterization of the stable d-automatic $A \subseteq \mathbb{Z}^m$ (rather than just $A \subseteq \mathbb{Z}$)? More generally, can Theorem 4.11 be extended to a description of the stable F-automatic $A \subseteq \Gamma$? A first conjecture is that they are precisely the F-sets.

- We saw in Theorem 6.13 that if $A \subseteq \Gamma$ is F-sparse (or more generally F-EDP) and Γ is weakly minimal then $\mathrm{Th}(\Gamma, +, A)$ is NIP. Can we extend this to a characterization of the F-automatic $A \subseteq \Gamma$ such that $\mathrm{Th}(\Gamma, +, A)$ is NIP? Are there any examples of an F-automatic set A that isn't F-sparse or the complement of an F-sparse set but nonetheless satisfies $\mathrm{Th}(\Gamma, +, A)$ is NIP? (Note from the definition that every F-EDP set is contained in an F-sparse set; thus if an F-EDP set is F-automatic then it's F-sparse by Proposition 3.6 (3). So the F-EDP sets aren't the sets we're asking for.)

  In a different direction, can we refine Theorem 6.13 to say something about whether $\mathrm{Th}(\Gamma, +, A)$ satisfies more fine-grained tameness properties within NIP? For example, will it be distal? What are the possible dp-ranks of types in $\mathrm{Th}(\Gamma, +, A)$? Is $\mathrm{Th}(\Gamma, +, A)$ strongly dependent? Dp-minimal?

References

[1] Boris Adamczewski and Jason P. Bell. On vanishing coefficients of algebraic power series over fields of positive characteristic. *Inventiones mathematicae*, 187(2):343–393, 2012. 1, 24, 43

[2] Jean-Paul Allouche and Jeffrey Shallit. The ring of k-regular sequences. *Theoretical Computer Science*, 98(2):163–197, 1992. 119

[3] Jean-Paul Allouche and Jeffrey Shallit. *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press, 2003. 1, 8, 119

[4] Matthias Aschenbrenner, Alf Dolich, Deirdre Haskell, Dugald Macpherson, and Sergei Starchenko. Vapnik-Chervonenkis density in some theories without the independence property, I. *Transactions of the American Mathematical Society*, 368(8):5889–5949, 2016. 118

[5] Jason Bell, Dragos Ghioca, and Rahim Moosa. Effective isotrivial Mordell-Lang in positive characteristic. *arXiv preprint arXiv:2010.08579*, 2020. 2, 10, 41, 43, 61

[6] Jason Bell and Rahim Moosa. F-sets and finite automata. *Journal de théorie des nombres de Bordeaux*, 31(1):101–130, 2019. 1, 2, 10, 31, 41, 43, 44, 45, 61

[7] Véronique Bruyere, Georges Hansel, Christian Michaux, and Roger Villemaire. Logic and p-recognizable sets of integers. *Bulletin of the Belgian Mathematical Society Simon Stevin*, 1994. 44, 119

[8] Artem Chernikov and Pierre Simon. Externally definable sets and dependent pairs. *Israel Journal of Mathematics*, 194(1):409–425, 2013. 116

[9] G. Conant, A. Pillay, and C. Terry. A group version of stable regularity. *Mathematical Proceedings of the Cambridge Philosophical Society*, 168(2):405–413, 2020. 81, 101, 103

[10] Gabriel Conant and Michael C. Laskowski. Weakly minimal groups with a new predicate. *Journal of Mathematical Logic*, 20(02):2050011, 2020. 105, 106

[11] Harm Derksen. A Skolem–Mahler–Lech theorem in positive characteristic and finite automata. *Inventiones mathematicae*, 168(1):175–224, 2007. 41

[12] Yuri Gurevich and Peter H. Schmitt. The theory of ordered abelian groups does not have the independence property. *Transactions of the American Mathematical Society*, 284(1):171–182, 1984. 104

[13] Christopher Hawthorne. Contributions to the theory of F-automatic sets. *arXiv preprint arXiv:2011.12405*, to appear in the Journal of Symbolic Logic. 3

[14] Christopher DC Hawthorne. Automata and tame expansions of $(\mathbb{Z}, +)$. *arXiv preprint arXiv:2007.00070*, to appear in the Israel Journal of Mathematics. 3

[15] Wilfrid Hodges. *Model theory*. Cambridge University Press, 2008. 81

[16] Oscar H. Ibarra and Bala Ravikumar. On sparseness, ambiguity and other decision problems for acceptors and transducers. In *Annual Symposium on Theoretical Aspects of Computer Science*, pages 171–179. Springer, 1986. 47

[17] Quentin Lambotte and Françoise Point. On expansions of $(\mathbb{Z}, +, 0)$. *Annals of Pure and Applied Logic*, page 102809, 2020. 3, 115, 118

[18] Rahim Moosa and Thomas Scanlon. F-structures and integral points on semiabelian varieties over finite fields. *American Journal of Mathematics*, 126(3):473–522, 2004. 10, 19, 41, 44, 45, 63, 74, 82

[19] Françoise Point. On decidable extensions of Presburger arithmetic: from A. Bertrand numeration sytems to Pisot numbers. *The Journal of Symbolic Logic*, 65(3):1347–1374, 2000. 116

[20] Olivier Salon. Suites automatiques à multi-indices et algébricité. *Comptes rendus de l'Académie des Sciences*, 305(12):501–504, 1987. 1, 43

[21] Pierre Simon. *A guide to NIP theories*. Cambridge University Press, 2015. 104, 118

[22] Andrew Szilard, Sheng Yu, Kaizhong Zhang, and Jeffrey Shallit. Characterizing regular languages with polynomial densities. In *International Symposium on Mathematical Foundations of Computer Science*, pages 494–503. Springer, 1992. 47

[23] K. Tent and M. Ziegler. *A Course in Model Theory*. Lecture Notes in Logic. Cambridge University Press, 2012. 62, 81, 97

[24] Sheng Yu. Regular languages. In Grzegorz Rozenberg and Arto Salomaa, editors, *Handbook of Formal Languages: Volume 1 Word, Language, Grammar*, pages 41–110. Springer Berlin Heidelberg, Berlin, Heidelberg, 1997. 4, 6, 86, 97, 99