Automatic Structures with Parameters

Frédéric Reinhardt
Matrikelnummer 233053

Diplomarbeit
vorgelegt der Fakultät
für Mathematik, Informatik und Naturwissenschaften
der Rheinisch-Westfälischen Technischen Hochschule Aachen

März 2013

angefertigt am
Lehr- und Forschungsgebiet
Mathematische Grundlagen der Informatik
Prof. Dr. Erich Grädel

Hiermit versichere ich, dass ich die Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Aachen, den 28. März 2013

(Frédéric Reinhardt)

Contents

Contents

Chapter 1

Introduction

This diploma thesis is a contribution to the theory of automatic structures, a research topic within the field known as algorithmic model theory. A common interest among computer scientists is the translation of structured datasets and models into a representation that enables a computer to solve algorithmic problems on the structure such as the evaluation of logical formulas and database queries. An automatic structure is a structure whose domain and relations can be represented by means of finite automata, i.e. the elements of the structure can be encoded as words and the set of its elements as well as its relations become regular languages, which have many nice algorithmic properties. Such a representation has for instance the advantage that all first order logic queries on the structure are decidable. The investigation of automatic structures was initiated by Khoussainov and Nerode [14] and was later carried foward by Blumensath [6], Rubin [18], Bárány [5], Kaiser [13] and many others. A problem within this field that had been occupying the minds of automatic structure researchers for some time was the question whether the additive rational group $(\mathbb{Q}, +)$ has an automatic presentation, until this problem was finally solved by Tsankov [21] in the negative. Later the observation was made that the rational group does have an automatic presentation with an advice automaton which means that a finite automaton can compute the sum of rational numbers if it is allowed access to an infinite advice tape during its computation. Furthermore this modification of the finite automaton model preserves many of the advantages of regular languages. This observation sparked some interest and in a recent publication [15], the model of a finite automaton with advice tape was presented and the question posed, which other structures have automatic presentations with advice automata. In this thesis a few more examples of advice automatic structures are presented, a logical formalism is developed that characterizes regular languages with a fixed advice and some of the techniques to prove the non-automaticity of structures are adapted to the new scenario.

The outline of this thesis is the following:

In **chapter 2** we revisit the model of an automaton with advice and some of its properties, that was introduced in [15]. We then define another automaton model with advice that operates on word sequences, where each word has a fixed length that is given by a length sequence. The purpose of these automata is that we can then define a corresponding logical formalism in **chapter 3** to express properties of automatic structures with parameters by quantification over regular word and number sequences. This formalism will then be used in **chapter 4** to give parametrised automatic presentations of the torsion-free rank 1 groups without the need to specify the automata. Instead we use a logical interpretation of the rational groups in the structure $\mathcal{W}(\mathbb{N})$ which will be defined in chapter 3. Chapter 5 is devoted to an investigation of non-automaticity criteria. We revisit the sum augmentation criterion of Delhommé and show how it can be modified to prove that there is no automatic presentation with parameter of any linear order with infinite VD-rank. Then we investigate the finite VD-hierarchy a bit closer and show that all linear orders on level 2 of the VD-hierarchy have an automatic presentation with parameter. The technique for the automatic presentation of scattered linear orders can be generalized to define an automatic subhierarchy of the VD hierarchy. We also revisit a non-automaticity criterion by Zaid [1] and show that it can be used when parameters are allowed in the automatic presentation.

I would like to thank Professor Grädel for giving me the opportunity to write my diploma thesis at his institute and Faried for his support and advice.

Chapter 2

Finite Advice Automata

In this chapter we consider formal languages and several types of automata with a finite set of states which can operate on objects of these languages (words, infinite words, sequences of finite words) by reading the input in a sequential manner. Finite automata are well-known formal modelling tools in computer science and mathematical logic which have been proven useful in many ways. Their utility as modelling tools is twofold. For one the class of finite automata which process finite objects constitutes a class of simple and robust sequential algorithms. It models precisely the class of algorithms that operate with a working memory of constant size (independent of the input size) which is represented by the finite state set of the automaton. There is a rich class of algorithmic problems which can be solved by finite automata. A standard example is the school method of adding two natural numbers in their decimal representation. An automatic structure is in this respect a structure whose basic relations and functions are not merely computable but computable by algorithms that have a constant space constraint. The second aspect that has made finite automata a preferred modelling tool is their utility as a finite representational device for formal languages and the connection between formal language theory and logic. Beginning with the work of Büchi, Elgot and Trakhtenbrot [8], who discovered the correspondence between relations that are definable in the monadic second order logic of one successor and (omega-)regular languages, finite automata that define formal languages are frequently used as a formal means to capture the expressivity of a logic. Together with the operational aspect of finite automata this method furthermore enables one to investigate the algorithmic properties of a logic, such as for instance the question of its decidability and its model-checking problem, which offers practical applications in the fields of software and hardware verification, database query languages and query evaluations.

In the following we are going to provide the reader with notations and standard results of formal language and automata theory that will be relevant throughout this thesis without going into more detail than necessary in so far as standard concepts are concerned. Readers who don't already have the required background knowledge in formal language theory are asked to consult the extensive literature [12], [16] on the topic for further clarification of the concepts mentioned here . The main focus of this chapter is the definition of languages that are regular with advice which have to the knowledge of the author only recently been introduced [15]. We then develop a more general automaton model that operates on word and number sequences and can also be parametrised.

2.1 Formal Languages

Alphabets An alphabet is a finite, non-empty set, that is commonly denoted by a greek capital letter $\Sigma, \Gamma, \ldots$.

(ω)words and sequences Let Σ be an alphabet. $\Sigma^* := \{a_1 \ldots a_n : n \in \mathbb{N}, a_i \in \Sigma\}$ and $\Sigma^{<\omega}$ denote the set of all finite words over Σ including the empty word ϵ. $(\Sigma^*, \cdot, \epsilon)$ is the free monoid generated by Σ with the concatenation product $\cdot$ of words. $\Sigma^\omega := \{a_0 a_1 \ldots : a_i \in \Sigma\}$ is the set of all ω-words, i.e. the set of all infinite sequences over Σ. $\Sigma^{\leq\omega} := \Sigma^{<\omega} \cup \Sigma^\omega$ is the set of all finite and ω-words. $\Sigma^{\leq\omega}$ can formally be identified with the set of all partial functions $\alpha \colon \mathbb{N} \to \Sigma$ with a domain $\mathrm{def}(\alpha) = [0, n)$, $n \in \mathbb{N} \cup \{\infty\}$ that is an initial segment of $\mathbb{N}$. We will frequently use the index-notation $\alpha(i)$ to refer to the i-th letter of the word α. For $i \notin \mathrm{Def}(\alpha)$ we set $\alpha(i) := \Box$, where $\Box$ is a padding symbol that is not in Σ. $\Sigma_\Box := \Sigma \cup \{\Box\}$ where it is assumed that $\Box \notin \Sigma$. More generally, a sequence σ over any set M is a partial function $\sigma \colon \mathbb{N} \to M$. () denotes the empty sequence and $\mathrm{Seq}(M)$, $\omega\mathrm{Seq}(M)$ the set of all finite resp. ω-sequences over M. $|\sigma| := |\mathrm{Def}(\sigma)| \in \mathbb{N} \cup \{\infty\}$ is the length of the sequene σ.

Languages A finite word-language L over alphabet Σ is a subset of Σ^*. An ω-word language over alphabet Σ is a subset of Σ^ω.

2.2 Finite Advice Automata

A finite advice automaton differs from the ordinary finite automaton model only in so far as in addition to its input tape, it has access to an infinite advice tape which holds an advice word $\alpha \in \Gamma^\omega$. The automaton reads the input word character by character from left to right in parallel with the advice α such that in the i-th step of its run it reads the i-th letter of the input and the i-th letter $\alpha(i)$ on the advice tape synchronously. Finite automata can thereby be regarded as particular advice automata with empty advice tape.

2.2.1 General properties

Since we want to provide an operational definition of finite advice automata which can be easily translated into an algorithm, we require the parameter $\alpha \in \Gamma^\omega$ to be computable as a function. Later we will relax this requirement and also allow non-computable functions as parameters in order to investigate automatic structures with parameters from a purely logical standpoint. The following definition is only a slight variation of the one given in [15].

Definition 2.2.1. A *finite advice automaton (FAA)* is a tuple $\mathcal{A} = (Q, \Sigma, \Gamma, \Delta, I, \mathcal{F})$ where

- Q is a finite set of *states*
- Σ is the *input alphabet*
- Γ is the *advice alphabet*

- $\Delta \subseteq Q \times \Sigma \times \Gamma \times Q$ is the *state transition relation*
- $I \subseteq Q$ is the set of *initial states*
- $\mathcal{F} \subseteq \mathcal{P}(Q)$ is the *acceptance component*

Given $\alpha \in \Gamma^\omega$ the pair $\mathcal{A}[\alpha] := (\mathcal{A}, \alpha)$ is called *finite automaton with advice α.*
A *run* $\rho \in Q^{\leq\omega}$ of $\mathcal{A}[\alpha]$ on input word $w \in \Sigma^{\leq\omega}$ is a a sequence of states that satisfies:

1. $\rho(0) \in I$
2. $(\rho(i), w(i), \alpha(i), \rho(i+1)) \in \Delta$ for all $i \in \mathrm{Def}(w)$

A finite word w is accepted by $\mathcal{A}[\alpha]$, if $\rho(|w|) \in \bigcup \mathcal{F}$.

An ω-word w is accepted by $\mathcal{A}[\alpha]$, if $\mathrm{Inf}(\rho) := \{q \in Q : \forall i \exists j > i(\rho(j) = q)\} \in \mathcal{F}$.

The language $L^{\leq\omega}(\mathcal{A}[\alpha])$ that is recognized by $\mathcal{A}[\alpha]$ is the set of all finite words and ω-words over input alphabet Σ that $\mathcal{A}[\alpha]$ accepts.

A language that is recognized by a finite automaton with advice α is called *regular with advice α.*

A FAA whose transition relation is a function is called *deterministic*, otherwise *non-deterministic*

The acceptance component $\mathcal{F}$ can be replaced w.l.o.g by a single acceptance set F if the FAA only operates on finite words.

Remark 2.2.1. A FAA can be operated in different modi in which it recognizes different languages.

finite words: We write $L(\mathcal{A}[\alpha]) \subseteq \Sigma^*$ for the set of finite words that $\mathcal{A}$ recognizes with advice α.

ω-words: $L^\omega(\mathcal{A}[\alpha]) \subseteq \Sigma^\omega$ denotes the ω-languages that $\mathcal{A}$ recognizes with advice α

mixed mode: $L^{\leq\omega}(\mathcal{A}[\alpha]) \subseteq \Sigma^{\leq\omega}$ with $L^{\leq\omega}(\mathcal{A}[\alpha]) := L(\mathcal{A}[\alpha]) \cup L^\omega(\mathcal{A}[\alpha])$.

Muller mode: In Muller mode the automaton is operated on input alphabet $\Sigma_\Box \times \Gamma_\Box$ without advice and recognizes the language $L(\mathcal{A})^{\leq\omega} \subseteq (\Sigma_\Box \times \Gamma_\Box)^{\leq\omega}$ with $(w, \alpha) \in L(\mathcal{A})^{\leq\omega} \Leftrightarrow w \in L(\mathcal{A}[\alpha])$ resp. $L(\mathcal{A})$ and $L^\omega(\mathcal{A})$ for finite and ω-words.

Example 2.2.1. For any $\alpha \in \Sigma^\omega$ the set of prefixes $\mathrm{Pref}(\alpha) := \{w \in \Sigma^* : w <_p \alpha\}$ of α is a regular language with advice α and the ω-language $\{\alpha\}$ is a ω-regular language with advice α. If α is not ultimately periodic then those languages cannot be recognized by any finite automaton without advice, since any ω-regular language contains an ultimately periodic word and any deterministic finite automaton which recognizes $\mathrm{Pref}(\alpha)$ recognizes $\{\alpha\}$ when run as a deterministic Büchi-automaton. A FAA that recognizes those languages is given by:

$$\mathcal{A} := (Q := \{q_0, q_1\}, \Sigma, \Sigma, q_0, \delta, \mathcal{F} := \{\{q_0\}\})$$

- $\delta(q_0, a, b) := \begin{cases} q_0 & \text{if } a = b \\ q_1 & \text{if } a \neq b \end{cases}$, for all $a, b \in \Sigma$

- $\delta(q_1, a, b) := q_1$, for all $a, b \in \Sigma$

It is easy to see that $L(\mathcal{A}[\alpha]) = \mathrm{Pref}(\alpha)$ and $L^\omega(\mathcal{A}[\alpha]) = \{\alpha\}$ for any $\alpha \in \Sigma^\omega$.

For finite deterministic automata without advice one can extend the transition function δ recursively to a function $\delta^*\colon Q \times \Sigma^* \to Q$ with

- $\delta^*(q, \epsilon) := q$
- $\delta^*(q, aw) = \delta^*(\delta(q, a), w)$ for $w \in \Sigma^*$ and $a \in \Sigma$

This doesn't work anymore for deterministic automata with advice, because there the transition function depends on the position of the automaton on the advice tape. For every position n let $\delta_n\colon Q \times \Sigma \to Q$ be the transition function of the automaton on position n of the advice tape. Formally, denote by $t_n(\alpha)$ the translated function $t_n(\alpha)(i) := \alpha(i + n)$. For a deterministic advice automaton $\mathcal{A} = (Q, \Sigma, \Gamma, \delta, q_0, \mathcal{F})$ let $t_n(\mathcal{A}) := (Q, \Sigma, \Gamma, \delta_n, q_0, \mathcal{F})$ be the advice automaton starting from position n on the advice tape. The run of $t_n(\mathcal{A})$ with advice α is then identical to the run of $\mathcal{A}$ on $t_n(\alpha)$. So that $L(t_n(\mathcal{A}))[\alpha] = L(\mathcal{A})[t_n(\alpha)]$.

The transition function can then be extended to words in the following way

- $\delta_n^*(q, \epsilon) := q$
- $\delta_n^*(q, aw) := \delta_{n+1}^*(\delta_n(q, a), w)$

The membership problem for an automaton $\mathcal{A}$ is the algorithmic problem of deciding for each finite input word $w \in \Sigma^*$ whether $\mathcal{A}$ accepts w or not.

Theorem 2.2.1. *For each computable parameter $\alpha \in \Gamma^\omega$ the membership problem for any finite automaton $\mathcal{A}[\alpha]$ with advice α is decidable.*

Proof. Let $\mathcal{A} = (Q, \Sigma, \Gamma, \Delta, I, F)$ be the given FAA. Let $w \in \Sigma^*$ be the input. Since α is computable, one can compute the sets $P_i := \{q \in Q : \exists p \in P_{i-1}(p, w(i-1), \alpha(i-1), q) \in \Delta\}$, beginning with $P_0 := I$ for $i = 1, 2, \ldots, |w|$ and test whether an accepting state is reached. □

The class of regular languages with a computable advice α constitutes thereby a subclass of the class of all recursive languages. The computational complexity of the decision procedure depends however on the computational complexity of the function α which can be non-elementary.

A finite advice automaton $\mathcal{A}$ defines not just one regular language, but a whole class of advice regular languages. For each advice $\alpha \in \Gamma^\omega$ the language $L(\mathcal{A}[\alpha])$. One might then ask which classes of languages can be defined in this way by a single advice automaton.

Definition 2.2.2. A class $\mathcal{K}$ of advice regular languages over input alphabet Σ is called *uniform regular*, if there is an advice alphabet Γ and a set $C \subseteq \Gamma^\omega$ of parameters, so that $L(\mathcal{A}[C]) := \{L(\mathcal{A}[\alpha]) : \alpha \in C\} = \mathcal{K}$.

Example 2.2.2. Let $\Sigma := \{a, b\}$.

1. For each two-valued function $f\colon \mathbb{N}^{>0} \to \{0,1\}$ define the set $U_f := \{a^n : f(n) = 1\} \subseteq \{a\}^*$. Then the class $\mathcal{U}_2 := \{U_f :$ f is a 2-valued function$\}$ is uniform regular. To prove this choose the advice alphabet $\Gamma := \{0,1\}$ and construct a finite automaton $\mathcal{A}$ that recognizes the regular language $\left(\begin{bmatrix}a\\0\end{bmatrix} + \begin{bmatrix}a\\1\end{bmatrix}\right)^* \begin{bmatrix}a\\1\end{bmatrix}$ over the alphabet $\Sigma \times \Gamma$. Let $(n_i)_{i\in\mathbb{N}}$ be a sequence of natural numbers with $0 < n_i < n_{i+1}$ for all i. Then for any parameter $\alpha = 0^{n_0-1}10^{n_1-n_0-1}10^{n_2-n_1-1}1\ldots$ it is $L(\mathcal{A}[\alpha]) = \{a^{n_i} : i \in \mathbb{N}\}$. For a 2-valued function enumerate the set $f^{-1}(\{1\}) = \{n_0 < n_1 < n_2 < \ldots\}$ in increasing order and choose the corresponding parameter.

2. As an example for class of advice regular sets that is not uniform regular consider the language family $L_i := \{w \in \{a,b\}^* : |w|_a \equiv 0 \pmod{i}\}$, for $i > 0$ where $|w|_a$ is the number of a's in w. Every L_i is regular, but the family of L_i's is not uniform regular. To see why, assume there were an advice automaton $\mathcal{A}$ wlog deterministic with $|Q|$ states that recognized the class $(L_i)_{i>0}$ uniformly. Let $C := |Q| + 1$. We show that the automaton doesn't recognize L_C for any parameter. Consider the C words $w_i := a^i b^{C-i}$ for $i = 0, \ldots, C-1$. Since the automaton has only $C-1$ states for each parameter α there must be two words $w_i \neq w_j$ so that the runs of $\mathcal{A}[\alpha]$ on w_i and w_j end in the same state q. Then the automaton cannot distinguish w_i and w_j, so that the runs on $w_i a^{C-i}$ and $w_j a^{C-i}$ also end in the same state and therefore either both words are accepted or both are rejected. Since $|w_i a^{C-i}|_a \equiv 0 \pmod{C}$ but $|w_j a^{C-i}|_a \equiv j - i \neq 0 \pmod{C}$ the automaton therefore doesn't recognize L_C.

A difference between the two examples is that in the positive example the number of states required to recognize each language in the class was bounded, whereas in the second example it was unbounded. This is no coincidence. To get another perspective on the role that the advice alphabet plays in the automaton consider that an advice automaton can also equivalently be specified in the form $\mathcal{A} = (Q, \Sigma, \Gamma, (\Delta_\gamma)_{\gamma\in\Gamma}, q_0, \mathcal{F})$ with $\Delta_\gamma := \{(p, a, q) \in Q \times \Sigma \times Q : (p, a, \gamma, q) \in \Delta\}$ for every $\gamma \in \Gamma$, i.e. any advice character γ specifies a different transition relation and the advice informs the automaton about which transition relation to use in each step. Imagined as a graph an advice automaton is therefore just a set of states with an overlay of different edge relations. Instead of relations we can also use functions, since deterministic and non-deterministic advice automata recognize the same class of languages. The number of different transition functions on a finite state set Q and edge labels Σ is bounded from above by $|Q^{Q\times\Sigma}|$. It therefore wouldn't make sense to use an advice alphabet for an automaton with more elements than there are transition functions, i.e. the advice alphabet $\Gamma := Q^{Q\times\Sigma}$ is already enough to capture all advice regular languages that can be recognized with $|Q|$ states. As in the case of finite automata with only one transition function, the states partition the set of all words into equivalence classes. Two words that end in the same state cannot be distinguished if there is only one transition relation. In the case of advice automata however this is only true for words of equal lengths, i.e. words in Σ^n for a n. Two words of unequal lengths can end in the same state and yet be distinguished in a later step, because the automaton won't necessarily use the same transition relations from that point on. Which transition relations it uses depends on its position on the advice tape. Formally, for any advice regular language L recognized by an advice automaton with $|Q|$ states, there exists a family of equivalence relations $(\equiv_n^L)_{n\in\mathbb{N}}$, where $\equiv_n^L \subseteq \Sigma^n \times \Sigma^n$ partitions Σ^n into $|Q|$ equivalence classes and any two words $w, v \in \Sigma^n$ with $w \in L$ and $v \notin L$ must lie in different equivalence classes. Any language L can be partitioned into equivalence classes with this property via $\Sigma^n / \equiv_n^L = \{(L \cap \Sigma^n), (L^{\complement} \cap \Sigma^n)\}$, so that this property alone is by far not sufficient to characterize the advice regular languages. Additionally transitions from a $\equiv_n$-equivalence class must land in the same $\equiv_{n+1}$-class: For all $w, v \in \Sigma^n$, $w \equiv_n^L v$ must imply $wa \equiv_{n+1}^L va$ for all $a \in \Sigma$. These two properties completely characterize language families that are uniform regular and in particular single languages that are regular with advice. (Additionally we need to assume that either all languages contain the empty word ϵ or none contains it). The proof is basically the same as the one in [15] for the Myhill-Nerode theorem for advice regular languages.

Proposition 2.2.1. *Let $(L_i)_{i\in I}$ be a family of languages over alphabet Σ, so that either all L_i contain ϵ or none. $(L_i)_{\in I}$ is uniform regular if and only if there exists a constant $C \in \mathbb{N}$ such that for all $i \in I$ there is a family $(\equiv_n^{L_i})_{n\in\mathbb{N}}$ of equivalence relations $\equiv_n^{L_i}$ on Σ^n so that for all $n \in \mathbb{N}$:*

0. $|\Sigma^n / \equiv_n^{L_i}| \leq C$
1. $\forall v \in L_i \forall w \in L_i^{\complement} : v \not\equiv_n^{L_i} w$
2. $\forall v, w \in \Sigma^n \forall a \in \Sigma : v \equiv_n^{L_i} w \rightarrow va \equiv_{n+1}^{L_i} wa$

Proof. $\Rightarrow$: Assume there is a finite advice automaton $\mathcal{A} = (Q, \Sigma, \Gamma, \delta, q_0, F)$ so that for every $i \in I$ there is a parameter $\alpha_i \in \Gamma^\omega$ with $L_i = L(\mathcal{A}[\alpha_i])$. Set $C := |Q|$ and define $\equiv_n^{L_i}$ by $w \equiv_n^{L_i} v$:iff "the runs of $\mathcal{A}[\alpha_i]$ on w and v end in the same state". Then conditions 0.-2. are evidently satisfied.
$\Leftarrow$: Assume C and $(\equiv_n^{L_i})_{n\in\mathbb{N}}$ with the stated properties exist. Define a finite advice automaton $\mathcal{A} := (Q, \Sigma, \Gamma, (\delta_\gamma)_{\gamma\in\Gamma}, q_0, F)$ where

- $Q := C \times \{0, 1\}$
- $\Gamma := Q^{Q\times\Sigma}$
- $\delta_\gamma := \gamma$
- $F := C \times \{1\}$
- $q_0 \in F$ if $\epsilon \in L_i$ for all i, else $q_0 \in F^{\complement}$.

We claim that for any $i \in I$ there exists a parameter $\alpha_i \in \Gamma^\omega$ with $L_i = L(\mathcal{A}[\alpha_i])$.
Let $i \in I$. First choose for every n an injective assignment $f_n \colon \Sigma^n / \equiv_n^{L_i} \rightarrow Q$ of equivalence classes to states, such that $f_n([w]_{\equiv_n^{L_i}}) \in F \Leftrightarrow [w]_{\equiv_n^{L_i}} \subseteq L_i$, where $[w]_{\equiv_n^{L_i}}$ denotes the $\equiv_n^{L_i}$-equivalence class that contains the word $w \in \Sigma^n$ and $f_0([\epsilon]_0^{L_i}) := q_0$. Such an assignment always exists, because for every n there are no more than C equivalence classes, whereas there are C non-accepting and C accepting states to choose from. Define the parameter α_i by $\alpha_i(n) := \gamma_n$, where $\gamma_n \in Q^{Q\times\Sigma}$ is the function that is defined by $\gamma_n(f_n([w]_{\equiv_n^{L_i}}), a) := f_{n+1}([wa]_{\equiv_{n+1}^{L_i}})$ for all n and $w \in \Sigma^n$, $a \in \Sigma$. Choose an

arbitrary value for $\gamma_n(q,a)$ for those (q,a) that haven't been assigned yet. Note that γ_n is well-defined, because $f_n([w]_{\equiv_n^{L_i}}) = f_n([v]_{\equiv_n^{L_i}}) \Rightarrow w \equiv_n^{L_i} v \Rightarrow wa \equiv_{n+1}^{L_i} va \Rightarrow f_{n+1}([wa]_{\equiv_{n+1}^{L_i}}) = f_{n+1}([va]_{\equiv_{n+1}^{L_i}})$ due to property 2. Now the unique run on a word $w = a_1 \dots a_k \in \Sigma^+$ is given by

- $\gamma_0(f_0([\epsilon]_{\equiv_0^{L_i}}, a_1)) = f_1([a_1]_{\equiv_1^{L_i}})$
- $\gamma_j(f_j([a_1 \dots a_j]_{\equiv_j^{L_i}}), a_{j+1}) = f_{j+1}([a_1 \dots a_{j+1}]_{\equiv_{j+1}^{L_i}})$ for $j = 1, \dots, k-1$
- $f_k([a_1 \dots a_k]_{\equiv_k^{L_i}}) \in F$ iff $a_1 \dots a_k \in L_i$.

$\square$

Note that for the proof the complete advice alphabet $\Gamma = Q^{Q \times \Sigma}$ was needed. One could ask for a characterization of those language families that are uniform regular if the size of the advice alphabet is fixed to a value $k = |\Gamma|$. As [15] show for every k there are languages that are regular with an advice over an advice alphabet with $k+1$ elements that cannot be recognized with an advice over an alphabet with only k elements.

Another equivalent characterization of the languages that are regular with advice can be given in terms of the colourings of the successor tree $\mathfrak{T}_\Sigma := (\Sigma^*, (\sigma_a)_{a \in \Sigma})$ with $\sigma_a(w) := wa$. A colouring of $\mathfrak{T}$ is a finite valued function $c : \Sigma^* \to C$. A subtree of $\mathfrak{T}_\Sigma$ is a substructure of $\mathfrak{T}$ induced by a subset of the form $w\{a,b\}^*$ for a $w \in \{a,b\}^*$, where w is the root of the subtree. A regular tree is a coloured tree such that it has up to isomorphism (i.e. identically coloured) only finitely many subtrees. For any such colouring and subset $S \subseteq C$ of colours $c^{-1}(S)$ is a regular language and vice versa for any regular language there is a colouring of $\mathfrak{T}_\Sigma$ with this property.

For advice regular languages the number of non-isomorphic subtrees is not necessarily finite but the number of non-isomorphic subtrees on each level is uniformly bounded.

2.2.2 Closure Properties

In the following we will establish that the class of ω-regular languages with advice α is closed under all operations that correspond to the semantical, set-theoretic interpretation of the connectives that occur in the inductive structure of first order logic formulae, which makes them suited as a formalism to represent the first order definable relations of structures that are automatic presentable with a parameter. In order to facilitate the proof and avoid reinventing the wheel, we will show how it can be reduced to the case of ordinary ω-regular languages without advice, for which closure under aforementioned operations is already a well-established fact.

Notice that finite automata with advice $\mathcal{A} = (Q, \Sigma, \Gamma, q_0, \delta, \mathcal{F})$ can also be interpreted as ordinary deterministic Muller-automata over extended alphabet $\Sigma \times \Gamma$.

Definition 2.2.3. Let $\alpha \in \Gamma^\omega$ be an advice. The *α-projection* is defined as follows

$$\cdot[\alpha] \colon \mathcal{P}(\Sigma^\omega \times \Gamma^\omega) \to \mathcal{P}(\Sigma^\omega)$$

$$L[\alpha] := \{\beta \in \Sigma^\omega : (\beta, \alpha) \in L\}$$

for all $L \subseteq \Sigma^\omega \times \Gamma^\omega$

Lemma 2.2.1. *For any finite advice automaton $\mathcal{A}$ and any parameter $\alpha \in \Gamma^\omega$ it holds that*

$$L^\omega(\mathcal{A}[\alpha]) = L^\omega(\mathcal{A})[\alpha] \subseteq \Sigma^\omega$$

Proof. Since the run of $\mathcal{A}[\alpha]$ on an input word $\beta \in \Sigma^\omega$ is identical to the run of of $\mathcal{A}$ on input (β, α) interpreted as deterministic Muller-automaton over extended alphabet $\Sigma \times \Gamma$ it holds that $w \in L(\mathcal{A}[\alpha]) \Leftrightarrow (\beta, \alpha) \in L(\mathcal{A}) \Leftrightarrow \beta \in L(\mathcal{A})[\alpha]$. □

Corollary 2.2.1. *For any $\alpha \in \Sigma^\omega$ the class of ω-regular languages with advice α is a Borel class.*

Proof. It is a well-known fact, that any ω-regular language $L \subseteq \Sigma^\omega$ is a Borel set [16]. Since furthermore $\Sigma^\omega \times \{\alpha\}$ is a closed set, $L[\alpha] = L \cap \Sigma^\omega \times \{\alpha\}$ is Borel, too. □

Theorem 2.2.2. *For any $\alpha \in \Sigma^\omega$ the class of ω-regular languages with advice α is effectively closed under union, complement, projection and cylindrification.*

Proof. We use Lemma 2.2.1 and the effective closure of ordinary ω-regular languages under the same operations [16]. Consider the projection $\pi\colon (\Sigma_1 \times \Sigma_2 \times \Gamma)^\omega \to (\Sigma_2 \times \Gamma)^\omega$, $(\beta_1, \beta_2, \alpha) \mapsto (\beta_2, \alpha)$. Let $\mathcal{A}$ be a finite advice automaton with advice alphabet Γ and input alphabet $\Sigma_1 \times \Sigma_2$. For $\mathcal{A}$ interpreted as a deterministic Muller-automaton over $\Sigma_1 \times \Sigma_2 \times \Gamma$ a deterministic Muller-automaton $\mathcal{A}_\pi$ over $\Sigma_2 \times \Gamma$ can be effectively constructed that recognizes the projection language $\pi(L^\omega(\mathcal{A})) = L^\omega(\mathcal{A}_\pi)$. Due to Lemma 2.2.1 the same automaton interpreted as a finite advice automaton recognizes the language $\pi(L^\omega(\mathcal{A}[\alpha])) = \pi(L^\omega(\mathcal{A})[\alpha]) = L^\omega(\mathcal{A}_\pi)[\alpha] = L^\omega(\mathcal{A}_\pi[\alpha])$. Closure under the remaining operations is proved analogously. □

Note that while the constructions of the finite advice automata in the previous proof are effective in the case of ω-regular languages with advice, the same doesn't hold true in general for finite word regular languages with advice. Though we can apply the constructions of the proof also to finite word languages, by embedding Σ^* into $\Sigma_\Box^\omega$ via $w \mapsto w\Box^\omega$ and treating the finite advice automaton as an automaton that reads ω-words, we cannot in general transform the resulting automaton back into an automaton over finite words over the same advice. In [15] a distinction is made between non-terminating and terminating automata with advice. Non-terminating automata are those that read an infinite number of □'s after the finite input words w has been read and accept according to the Muller-acceptance condition. The distinction is warranted, because as they demonstrated, there is a language L that can be recognized by a non-terminating automaton with advice α, but not by a terminating automaton with the same advice α. As a consequence finite word regular languages with a fixed advice α are not necessarily closed under projection, though they are closed under boolean operations. We use the example of the language L that was given in [15] to show non-closure under projection over a fixed advice.

Proposition 2.2.2. *There exists an advice α, so that the class of regular languages over finite words with advice α is not closed under projection.*

Proof. Take for α any non-ultimately periodic word, for example $\alpha = \alpha(0)\alpha(1)\ldots = 1101001000\ldots$. Let $R := \{(0^n, 0^{n+1}) \in 0^* \times 0^* : \alpha(n) = 1\}$. Then the convolution of R can be recognized by a finite automaton with advice α. The automaton merely has to check that the input $\begin{bmatrix} 0^n\Box \\ 0^{n+1} \end{bmatrix}$ aligns with a 1 on the advice tape. The projection of R on the first component is the language $R_1 := \{0^n : \alpha(n) = 1\}$ and is not regular with advice α. Intuitively the automaton would have to guess whether the next character on the advice tape is a 1 or a 0. Suppose for a contradiction that there were a finite advice automaton $\mathcal{A} = (Q, \{0\}, \{0,1\}, \delta, q_0, F)$ that recognizes L with advice α. Then one could also construct an automaton without advice, that recognizes $Pref(\alpha)$. The automaton can simulate $\mathcal{A}$ using its own input as the advice for $\mathcal{A}$ and whether the next character that it is reading must be a 1 or a 0, depending on whether $\mathcal{A}$ is in a final state or not. $\Box$

On the other side closure under projection for the class of regular languages with advice (any advice) can be guaranteed, because of this proposition:

Proposition 2.2.3 ([15])**.** *For every language L, if L is non-terminating regular with advice α then there exists an advice α', such that L is terminating regular with advice α'*

Proof. Let $\mathcal{A}$ be a finite advice automaton that recognizes L in its non-terminating mode, i.e. $w \in L$ if and only if $\mathcal{A}$ accepts $w\Box^\omega$. The advice α' simply encodes at each position n the information whether $\mathcal{A}$ accepts the word $\Box^\omega$ beginning in state q and position n on the advice tape. $\Box$

2.2.3 Advice automata over sequences

It is sometimes more convenient to tokenize the input stream of a finite advice automaton into a list of seperate words and to treat the automaton as processing a list of finite words, where each word is read together with a seperate finite parameter, instead of just as a sequence of characters. Frequently we will use parameters $\alpha \in \Gamma^\omega$, that consist of an infinite list $\alpha = w_1\#w_2\#\ldots$ of finite parameters $w_i \in (\Gamma \setminus \{\#\})^*$, seperated by some delimiter symbol $\# \in \Gamma$. Such a parameter naturally induces a tokenization of the input word $\beta = v_1a_1v_2a_2\ldots$, with $a_i \in \Sigma$,$v_i \in \Sigma^*$ and $|v_ia_i| = |w_i\#|$. Any symbol in the advice alphabet that occurs infinitely often in the parameter can serve as a delimiter symbol that determines a tokenization of the input word. Since the automaton reads input and advice tape synchronously in the form $\beta \otimes \alpha = \begin{bmatrix} v_1a_1 \\ w_1\# \end{bmatrix} \begin{bmatrix} v_2a_2 \\ w_2\# \end{bmatrix} \begin{bmatrix} v_3a_3 \\ w_3\# \end{bmatrix} \ldots$ it can recognize with the help of the delimiter symbol where a token v_ia_i ends and where the next token $v_{i+1}a_{i+1}$ begins. By eventually padding the input word with $\Box$'s, we can always guarantee that its end aligns with a $\#$ on the advice tape.

Definition 2.2.4. A *finite sequence automaton with advice (FSAA)* is a tuple $\mathcal{G} = (Q, \Sigma, \Gamma, (\mathcal{A}_{ij})_{i,j\in Q}, I, \mathcal{F})$ where

- Q is a finite set of *states*
- Σ is the *input alphabet*
- Γ is the *advice alphabet*
- $(\mathcal{A}_{ij})_{i,j\in Q}$ are *finite advice automata* over Σ,Γ
- $I \subseteq Q$ is the set of *initial states*
- $\mathcal{F} \subseteq \mathcal{P}(Q)$ is the *acceptance component*

Given an advice sequence $\gamma \in (\omega)\mathrm{Seq}(\Gamma^*)$ The pair $\mathcal{G}[\gamma] := (\mathcal{G}, \gamma)$ is called *finite sequence automaton with advice sequence* γ.

A *run* $\rho \in Q^{\leq\omega}$ of $\mathcal{G}[\gamma]$ on a sequence of input words $\varsigma \in \mathrm{Seq}(\Sigma^*)$, is a sequence of states that satisfies:

1. $\rho(0) \in I$
2. $\varsigma(i) \in L(\mathcal{A}_{\rho(i),\rho(i+1)}[\gamma(i)])$ for all $i \in \mathrm{Def}(\varsigma)$

The acceptance condition is the same as for FAA's.

The sequence language that is recognized by $\mathcal{G}[\gamma]$ is the set $S^{\leq\omega}(\mathcal{G}[\gamma]) := \{\varsigma \in (\omega)\mathrm{Seq}(\Sigma^*) : \mathcal{G}[\gamma] \text{ has an accepting run on } \varsigma\}$ A set $S \subseteq (\omega)\mathrm{Seq}(\Sigma^*)$ is called (ω)*regular with advice sequence* γ iff it is recognized by a sequence automaton with advice sequence γ.

Example 2.2.3. Let $\Sigma := \{a, b\}, \Gamma := \{0, \#\}$ and

$$\begin{aligned} S :=& \{\varsigma \in \mathrm{Seq}(\Sigma^*) : \forall i \in \mathrm{Def}(\varsigma)(\varsigma(i) \in a^*ba^* \wedge |\varsigma(i)| = i+1)\} \\ =& \{(), (b), (b, ab), (b, ba), (b, ab, aab), (b, ab, aba), (b, ab, baa), (b, ba, aab), \ldots\}. \end{aligned}$$

Then S is regular with advice sequence $\gamma = (\#, 0\#, 00\#, 000\#, \ldots)$. A sequence automaton that recognizes S is given by

$$\mathcal{G} := (Q := \{q_0\}, \Sigma, \Gamma, \mathcal{A}_{q_0,q_0}, q_0, F := \{q_0\})$$

, where $\mathcal{A}_{q_0,q_0}$ is an advice automaton that recognizes the regular language

$$\begin{bmatrix} a \\ 0 \end{bmatrix}^* \left(\begin{bmatrix} b \\ \# \end{bmatrix} + \begin{bmatrix} b \\ 0 \end{bmatrix} \begin{bmatrix} a \\ 0 \end{bmatrix}^* \begin{bmatrix} a \\ \# \end{bmatrix} \right)$$

in Muller-mode. Then $S(\mathcal{G}[\gamma]) = S \subseteq \mathrm{Seq}(\Sigma^*)$

A sequence automaton can also be operated in word mode, in which it tokenizes an input word into a sequence and then reads the sequence token for token. There are a priori several different ways in which a word can be tokenized.

Definition 2.2.5. Let $\mathcal{G}$ be a sequence automaton with advice sequence γ. Every sequence of natural numbers $\bar{l} \in \omega\mathrm{Seq}(\mathbb{N}^{>0})$ determines a *word mode* of $\mathcal{G}$. The $\bar{l}$-*tokenization* of a word $w \in \Sigma^{\leq\omega}$ is the word sequence $\varsigma := (w)_{\bar{l}}$ with

$$\varsigma(i) := w[\bar{l}(i), \bar{l}(i+1))$$

. The language $L^{\bar{l}}(\mathcal{G}[\gamma])$ is defined by

$$w \in L^{\bar{l}}(\mathcal{G}[\gamma]) :\Leftrightarrow (w)_{\bar{l}} \in S(\mathcal{G}[\gamma])$$

Let

$$\bar{l}\text{-Seq}(\Sigma^*) := \{\sigma \in \text{Seq}(\Sigma^*) : \forall i(i \in \text{Def}(\sigma) \to |\gamma(i)| = l(i)\}$$

Example 2.2.4. Consider the language

$$\begin{aligned} L :=& \{x_0 x_1 \dots x_k \in \{a, b\}^+ : k \geq 0 \wedge \forall i(|x_i| = i + 1 \wedge x_{2i} \in a^* \wedge x_{2i+1} \in b^*)\} \\ =& \{a, abb, abbaaa, abbaaabbbb, \dots\} \end{aligned}$$

Choose the length sequence $\bar{l} = (1, 2, 3, \dots)$. Then a sequence automaton that recognizes L in its $\bar{l}$-word mode can be easily constructed. The sequence automaton $\mathcal{G}$ has two states q_0, q_1 that are both accepting and regular transition languages $L_{q_0,q_1} = \begin{bmatrix} a \\ 0 \end{bmatrix}^*$ and $L_{q_1,q_0} := \begin{bmatrix} b \\ 0 \end{bmatrix}^*$. Then $L^{\bar{l}}(\mathcal{G}[\cdot]) = L$ since $\bar{l}$ determines that words in L get tokenized into the sequence language $S := \{(a), (a, bb), (a, bb, aaa), \dots\}$ and $S(\mathcal{G}[\cdot]) \cap \bar{l}\text{-Seq}(\Sigma^*) = S$. An additional advice besides $\bar{l}$ which is already implicit in the definition of its word mode is not needed for $\mathcal{G}$ to recognize the language.

In general, sequence automata that operate in word mode and word automata with advice are not equivalent, because a word automaton cannot recognize where a new token begins. If the advice has however a delimiter symbol # it determines a tokenization of the input word and a fixed length sequence.

Definition 2.2.6. The *#-tokenization* $(\alpha)_\#$ of a parameter $\alpha = w_0 \# w_1 \# \dots \in \Gamma^\omega$ with $w_i \in (\Gamma \setminus \{\#\})^*$ for all $i \in \mathbb{N}$ is defined as the sequence

$$(\alpha)_\# := (w_0\#, w_1\#, \dots) \in \omega\text{Seq}((\Gamma \setminus \{\#\})^*\#)$$

$l(\alpha)$ is the *length sequence* determined by α, given by $l(\alpha)(i) := |w_i\#|$.

Proposition 2.2.4. *For any sequence automaton $\mathcal{G}$ an advice automaton $\mathcal{A}$ can be effectively constructed, so that for all $\alpha \in \Gamma^\omega$ with $\# \in \text{Inf}(\alpha)$ it holds that*

$$L^{l(\alpha)}(\mathcal{G}[(\alpha)_\#]) = L(\mathcal{A}[\alpha])$$

Proof. Let $\mathcal{G} = (Q, \Sigma, \Gamma, (\mathcal{A}_{i,j})_{i,j \in Q}, q_0, \mathcal{F})$ be an FSSA with transition automata $\mathcal{A}_{i,j} = (Q_{i,j}, \Sigma, \Gamma, \delta, q^0_{i,j}, \mathcal{F}_{i,j})$. The transition automata can be thought of as "subprograms" that the automaton calls on every token of the input sequence and whose computation determines the next state. First substitute every transition automaton $\mathcal{A}_{i,j}$ by a FAA $\mathcal{A}'_{i,j}$ with

$$L(\mathcal{A}'_{i,j}) = L(\mathcal{A}_{i,j}) \cap (\Sigma \times (\Gamma \setminus \{\#\}))^*(\Sigma \times \{\#\})$$

, i.e. $\mathcal{A}_{i,j}$ and $\mathcal{A}'_{i,j}$ recognize the same tokens $\begin{bmatrix} va \\ w\# \end{bmatrix}$, but $\mathcal{A}'_{i,j}$ doesn't accept any word that doesn't have the form $\begin{bmatrix} va \\ w\# \end{bmatrix}$. Obviously this substitution doesn't change the behaviour of $\mathcal{G}$ on #-tokenizations. Now integrate everything into one big automaton.

$$\mathcal{A} := (P, \Sigma, \Gamma, \Delta, I, \mathcal{F})$$

- wlog assume that the sets $Q_{i,j}$ and Q are pairwise disjoint and set $P := \bigcup_{i,j} Q_{i,j} \cup Q$
- Now connect the "main states" $i \in Q$ to the initial states of the transition automata $\mathcal{A}_{i,j}$ and all final states of the $\mathcal{A}'_{i,j}$ to $j \in Q$ via ϵ-transitions.

$$\Delta := \bigcup_{i,j \in Q} \delta_{i,j} \cup \bigcup_{i,j \in Q} \{(i, \epsilon, q^0_{i,j})\} \cup \bigcup_{i,j \in Q} \{(q, \epsilon, j) : q \in \bigcup \mathcal{F}_{i,j}\}$$

- An FAA with ϵ-transitions can easily be transformed into an equivalent one without ϵ-transitions by computing its ϵ-hull.
- It remains to show that $\mathcal{A}$ and $\mathcal{G}$ have the same behaviour on #-tokenizations. Every accepting run of $\mathcal{G}$ has the form $i_0 \xrightarrow{\begin{bmatrix} v_0 a_0 \\ w_0\# \end{bmatrix}} i_1 \xrightarrow{\begin{bmatrix} v_1 a_1 \\ w_1\# \end{bmatrix}} i_2 \ldots$ with $\forall j : i_j \in Q$ and $\forall i : \begin{bmatrix} v_i a_i \\ w_i\# \end{bmatrix} \in L(\mathcal{A}'_{i,j})$. $\mathcal{A}$ can produce a run of the same form by taking the appropriate ϵ-transitions and vice versa $\mathcal{A}$ has only accepting runs of the above form due to our modification of the transition automata.

$\square$

Just as finite automata can operate on tuples of words and thereby define regular word relations, sequence automata can operate on tuples of sequences and thereby define regular sequence relations with advice. For this matter k-relations over Σ are reduced to languages by convoluting a tuple of words into a a single word over an extended alphabet $\Sigma^k_\square := \Sigma_\square \times \Sigma_\square \ldots \times \Sigma_\square$,where $\square$ is a padding symbol that is being used to right align words of different lengths. In the same way tuples of sequences can be transformed into a single word over an extended alphabet.

Definition 2.2.7. Let $(w_1, \ldots, w_k) \in \underbrace{\Sigma^* \times \ldots \times \Sigma^*}_{\text{k times}}$. To convolute a word-tuple, the words $w_1, \ldots, w_k$ are written below each other and then padded with $\square$ to the length $l := \max_{i=1}^k |w_i|$ of the longest word. I.e. $\otimes_l(w_1, \ldots, w_k) := \begin{bmatrix} w_1(0) \\ \vdots \\ w_k(0) \end{bmatrix} \ldots \begin{bmatrix} w_1(l-1) \\ \vdots \\ w_k(l-1) \end{bmatrix} \in (\Sigma^k_\square)^*$, where $w_i(j) = \square$, if $j \notin \mathrm{Def}(w_i)$. $\otimes$ is extended to sets of k-tuples $R \subseteq \Sigma^* \times \ldots \times \Sigma^*$ by convoluting ever element of the set: $\otimes R := \{\otimes_l(w_1, \ldots, w_k) : (w_1, \ldots, w_k) \in R \wedge l = \max_{i=1}^k |w_i|\}$.

- A k-ary relation $R \subseteq \Sigma^* \times \ldots \Sigma^*$ is called *regular with advice* $\alpha \in \Gamma^\omega$ iff the language $\otimes R \subseteq (\Sigma_\Box^k)^*$ is regular with advice α.

Similarily we define the convolution of k-tuples of sequences $(\varsigma_1, \ldots, \varsigma_k) \in \bar{l}$-$(\omega)\mathrm{Seq}(\Sigma^*)$, where each component has a fixed length $|\varsigma(i)| = \bar{l}(i)$ for all $i \in \mathrm{Def}(\varsigma)$. $\otimes_{\bar{l}}(\varsigma_1, \ldots, \varsigma_k) \in \bar{l}$-$(\omega)\mathrm{Seq}(\Sigma^*)$ with $\otimes_{\bar{l}}(\varsigma_1, \ldots, \varsigma_k)(i) := \otimes_{l(i)}(\varsigma_1(i), \ldots, \varsigma_k(i))$ for all $i \in \bigcup_{j=1}^k \mathrm{Def}(\varsigma_j)$. $\otimes_{\bar{l}}$ is also extended to sets by applying it to every element of the set.

- A k-ary relation $R \subseteq \bar{l}$-$\mathrm{Seq}(\Sigma^*) \times \ldots \times \bar{l}$-$\mathrm{Seq}(\Sigma^*)$ is called *regular with advice sequence* γ iff there is a sequence automaton $\mathcal{G}$ with $S(\mathcal{G}[\gamma]) = \otimes_{\bar{l}} R$

Example 2.2.5.

$$\otimes_{(3,2,4,\ldots)}((aaa, ba), (baa, aa, bbbb)) = \left(\begin{bmatrix} aaa \\ baa \end{bmatrix}, \begin{bmatrix} ba \\ aa \end{bmatrix}, \begin{bmatrix} bbbb \\ \Box\Box\Box\Box \end{bmatrix} \right)$$

The following proposition establishes the correspondence between regular sequence relations with advice and regular word relations with advice.

Proposition 2.2.5. *Let* $\alpha \in \Gamma^\omega$ *with* $\# \in \mathrm{Inf}(\alpha)$.

$$S \subseteq l(\alpha)\text{-}\mathrm{Seq}(\Sigma^*)^k \textit{ is regular with advice } (\alpha)_\#$$
$$\Leftrightarrow$$
$$[\otimes_{l(\alpha)} S]_{\mathrm{concat}} \subseteq (\Sigma_\Box^k)^* \textit{ is regular with advice } \alpha$$

Proof. Let S be regular with advice $(\alpha)_\#$. According to definition there exists then a sequence automaton with $S(\mathcal{G}[(\alpha)_\#] = \otimes_{l(\alpha)} S$. According to proposition 2.2.4 there exists an advice automaton with $L(\mathcal{A}[\alpha]) = L^{l(\alpha)}(\mathcal{G}[(\alpha)_\#]) = [\otimes_{l(\alpha)} S]_{\mathrm{concat}}$. $\Box$

Our next goal is to define an automaton model that operates on natural number sequences which can then later be used to define arithmetic operations on the generalized "digits" of a number in other representation systems than the usual p-adic ones. For this matter we first need to define a proper coding of number sequences as word sequences.

Definition 2.2.8. Interpret each word $w \in \{0,1\}^{l-1}\#$ as the binary representation of a natural number $\mathrm{num}(w) := \sum_{i<l-1} w(i)2^i$ of length $\mathrm{len}(w) := l$. Conversely every pair $(n,l) \in \mathbb{N} \times \mathbb{N}^{>0}$, with $n < 2^{l-1}$, has a unique representation of the form $\mathrm{bin}(n,l) = a_1 a_2 \ldots a_{l-1}\# \in \{0,1\}^{l-1}\#$, so that $\mathrm{bin}(\mathrm{num}(w), \mathrm{len}(w)) = w$. Let $\mathcal{N} := \{(n,l) \in \mathbb{N} \times \mathbb{N}^{>0} : n < 2^l\}$. The coding $\mathrm{bin}\colon \mathcal{N} \to \{0,1\}^*\#$ is bijective with inverse function $\mathrm{bin}^{-1}(w) = (\mathrm{num}(w), \mathrm{len}(w))$. Extend bin to a one-to-one correspondence between sequences: $\overline{\mathrm{bin}}\colon (\omega)\mathrm{Seq}(\mathcal{N}) \to (\omega)\mathrm{Seq}(\{0,1\}^*\#)$.

Definition 2.2.9. Let $l \in \mathbb{N}^{>0}$ and $(m_1, \ldots, m_k) \in \mathbb{N}_\Box^k$ and $\Sigma_2 := \{0, 1, \#\}$ The l-*convolution* of a k-tuple of numbers (and $\Box$'s) is the word

$$\otimes_l(m_1, \ldots, m_k) := \begin{bmatrix} \mathrm{bin}(m_1, l) \\ \mathrm{bin}(m_2, l) \\ \vdots \\ \mathrm{bin}(m_k, l) \end{bmatrix} \in \Sigma_{2,\Box}^l, \text{where } \mathrm{bin}(\Box, l) := \Box^l$$

$\otimes$ is extended to sets of k-tuples $R \subseteq \mathbb{N}^k$ by convoluting ever element of the set: $\otimes R := \{\otimes_l(m_1, \ldots, m_k) : (m_1, \ldots, m_k) \in R \wedge l = \max_{i=1}^{k} \lfloor \log_2(m_i) \rfloor + 2\}$.

- A k-ary relation $R \subseteq \mathbb{N}^k$ is called *regular with advice* $\alpha \in \Gamma^\omega$ iff the language $\otimes R \subseteq (\Sigma_{2,\square}^k)^*$ is regular with advice α.

Let $\bar{l} \in \omega\mathrm{Seq}(\mathbb{N}^{>0})$ and $\mathfrak{m} \in (\omega)\mathrm{Seq}(\mathbb{N})^k$. The $\bar{l}$-*convolution* of $\mathfrak{m} = (\mathfrak{m}_1, \ldots, \mathfrak{m}_k)$ is the sequence of words $\otimes_{\bar{l}}\mathfrak{m} \in (\omega)\mathrm{Seq}((\Sigma_{2,\square}^k)^*)$ with $\otimes_{\bar{l}}\mathfrak{m}(i) := \otimes_{l(i)}(\mathfrak{m}_1(i), \ldots, \mathfrak{m}_k(i))$ for all $i \in \bigcup_{j=1}^{k} \mathrm{Def}(\mathfrak{m}_j)$.

$\otimes_{\bar{l}}$ is also extended to sets by applying it to every element of the set. Let $\bar{l}$-$(\omega)\mathrm{Seq}(\mathbb{N}) := \{\eta \in (\omega)\mathrm{Seq}(\mathbb{N}) : \forall i \in \mathrm{Def}(\eta)(\eta(i) < 2^{l(i)-1})\}$

- A k-ary relation $R \subseteq \bar{l}$-$\mathrm{Seq}(\mathbb{N})^k$ is called *regular with advice sequence* γ iff there is a sequence automaton $\mathcal{G}$ with $S(\mathcal{G}[\gamma]) = \otimes_{\bar{l}} R$

Example 2.2.6.

$$\otimes_{(5,5,4,3)}((13, 9, 0, 1), (4, 6)) = \left(\begin{bmatrix} 1011\# \\ 0010\# \end{bmatrix}, \begin{bmatrix} 1001\# \\ 0110\# \end{bmatrix}, \begin{bmatrix} 000\# \\ \square\square\square\square \end{bmatrix} \begin{bmatrix} 10\# \\ \square\square\square \end{bmatrix} \right)$$

In analogy to the case of word sequences we have:

Proposition 2.2.6. *Let $\alpha \in \Sigma_2^\omega$ with $\# \in \mathrm{Inf}(\alpha)$.*

$$S \subseteq l(\alpha)\text{-}(\omega)\mathrm{Seq}(\mathbb{N})^k \textit{ is regular with advice } (\alpha)_\#$$
$$\Leftrightarrow$$
$$[\otimes_{l(\alpha)} S]_{\mathrm{concat}} \subseteq (\Sigma_{2,\square}^k)^* \textit{ is regular with advice } \alpha$$

We will make use of the fact that finite automaton recognizable relations over the natural numbers can also be defined in an extension of presburger arithmetic. FO-formulae can thereby be used to specify the transition relations of sequence advice automata in a more informative way.

Theorem 2.2.3 (Bruyère [7])**.** *$R \subseteq \mathbb{N}^k$ is FO-definable in $(\mathbb{N}, +, |_2)$ if and only if $\otimes R$ is regular.*

Example 2.2.7. We describe a sequence advice automaton that reads the coefficient list of two natural numbers in their factorial representation and computes the coefficient list of their sum. Every natural number $n \in \mathbb{N}$ has a representation of the form $n = \sum_{i=0}^{k} a_i(i+1)!$ with coefficients $0 \leq a_i < i+2$ which is unique if it is additionally required that the coefficient list ends with a number a_k that is not zero, unless $n = 0$. An important difference to other representations of numbers in positional numeration systems, such as the base-k representations, is that the digits a_i of the factorial representation are unbounded whereas the digits of the better known base-k representations of natural numbers are members of a finite alphabet. The factorial representation can therefore better be thought of not as a single finite word but a list of finite words and it cannot be recognized by a finite automaton without advice. To add two numbers in factorial representation one merely computes for $i = 0, \ldots$ the modulo $i + 2$ sum of the i-th coefficient pair and the previous carry and saves a carry for the next step if and only if their sum is greater or equal $i + 2$. Since an ordinary finite automaton cannot count indefinitely the automaton is adviced with the number $i + 2$ in the i-th position.

The automaton has two state q_0 and q_1 which remember whether a carry was generated in the previous operation or not and its transition relation is given by FO-formulae that define the operation on the coefficients. Since the coefficient operations can be defined in $(\mathbb{N}, +)$, it is a regular relation over the naturals. The FO-formulae φ_{q_i,q_j} define the transition relation. The variables x,y,z stand for the two input and the one output coefficients and the variable i denotes the advice that is read in that step. Δ is a domain formula that restricts the range of the coefficients that are allowed as inputs. A finite advice automaton over words with the parameter $\alpha = \text{bin}(2)\#\text{bin}(3)\#\text{bin}(4)\#\ldots = 01\#11\#001\#101\#011\#,\ldots$ can compute this function.

$\Delta(x, y, i) := x < i + 2 \wedge y < i + 2$

$\varphi_{q_0,q_0}(x, y, z, i) := x + y < i + 2 \wedge x + y = z$

$\varphi_{q_0,q_1}(x, y, z, i) := x + y \geq i + 2 \wedge x + y = z + i + 2$

$\varphi_{q_1,q_0}(x, y, z, i) := x + y + 1 < i + 2 \wedge x + y + 1 = z$

$\varphi_{q_1,q_1}(x, y, z, i) := x + y + 1 \geq i + 2 \wedge x + y + 1 = z + i + 2$

Chapter 3

Automatic Presentations with Parameters

In this chapter we define the notion of an automatic presentation of a structure with an additional parameter. Furthermore we develop a logical formalism that allows quantification over sequences that are regular with advice.

3.1 Preliminaries

We recall some standard notions of logic and model theory in order to fix notations and give the reader an overview of the concepts that are presupposed in this and subsequent chapters.

Structures *Signatures* specify the predicate $(R_0, R_1, \ldots)$,function $(f_0, f_1, \ldots)$ and constant $(c_0, c_1, \ldots)$ -symbols which occur in logic formulae and are denoted by τ. τ-structures $\mathfrak{A} = (A, R_0^{\mathfrak{A}}, R_1^{\mathfrak{A}}, \ldots, f_0^{\mathfrak{A}}, f_1^{\mathfrak{A}}, \ldots, c_0^{\mathfrak{A}}, c_1^{\mathfrak{A}}, \ldots)$ interpret those symbols by set-theoretical relations, functions and constants over a universe A of elements. The arity of a relation R_i is usually denoted by $\mathrm{ar}(R_i)$ or r_i. Example: $(\mathbb{N}, +, \cdot, <, 0)$ is a $\tau = \{+, \cdot, <, 0\}$-structure with universe $\mathbb{N}$, two 2-ary functions $+, \cdot$, one 2-ary relation $<$ and one constant 0. When we consider automatic presentations or interpretations of structures the functions are usually represented by their graphs and constants by singleton sets. A *relational signature* is a signature that consists only of relations.

Logic Given a signature τ and a logic $\mathcal{L}$, we denote the class of $\mathcal{L}$ formulae over τ by $\mathcal{L}(\tau)$. This chapter is mainly concerned with first order logic (FO). First order logic can be extended by the counting quantifiers $\exists^{\infty} x$ for "there exist infinitely many x.." and $\exists^{(t,k)} x$ for "there exist t modulo k many x.." and cardinality quantifiers $\exists^{\kappa}$ for "there exists κ many x..", where κ is a cardinal number. This extension is denoted by FOC.

3.2 Automatic Presentations

Definition 3.2.1. Let τ be a relational signature and $\mathfrak{A} = \big(A, (R^{\mathfrak{A}})_{R \in \tau}\big)$ a τ-structure.

$$\mathfrak{d} := (\nu, \Sigma, \Gamma, L_{\Delta}, L_{\approx}, (L_R)_{R \in \tau}, \alpha)$$

is called *(ω)automatic presentation of $\mathfrak{A}$ with parameter α* iff

- $\alpha \in \Gamma^\omega$
- $L_\Delta[\alpha] \subseteq \Sigma^{\leq\omega}$, $L_\approx[\alpha] \subseteq (\Sigma^{\leq\omega})^2$, $L_R[\alpha] \subseteq (\Sigma^{\leq\omega})^{\mathrm{ar}(R)}$ are regular with advice α
- $\nu \colon L_\Delta[\alpha] \to \mathrm{A}$ is surjective
- $\forall (x,y) \in L_\Delta[\alpha]^2 : (x,y) \in L_\approx[\alpha]^2 \Leftrightarrow \nu(x) = \nu(y)$
- $\forall (x_1, \ldots, x_{\mathrm{ar}(R)}) \in L_\Delta[\alpha]^{\mathrm{ar}(R)}$:
  $(x_1, \ldots, x_{\mathrm{ar}(R)}) \in L_{\mathrm{R}}[\alpha]^{\mathrm{ar}(R)} \Leftrightarrow (\nu(x_1), \ldots, \nu(x_{\mathrm{ar}(R)})) \in R$

In this case $L_\approx[\alpha]$ is a congruence relation on the τ-structure

$$\mathfrak{A}_\mathfrak{d} = (L_\Delta[\alpha], L_\approx[\alpha], (L_R[\alpha])_{R\in\tau})$$

and the quotient structure $\mathfrak{A}_\mathfrak{d}/\approx$ and $\mathfrak{A}$ are ismorphic.

(ω)AutStr$[\alpha]$ denotes the class of all structures that have an (ω)automatic presentation with parameter α and for a set of parameters $M \subseteq \Gamma^\omega$, $(\omega)\text{AutStr}[M] := \bigcup_{\alpha\in M}(\omega)\text{AutStr}[\alpha]$. By AutStr[all] is meant the class of structures that are automatic presentable with some parameter.

Example 3.2.1. A ω-word $\alpha \in \Gamma^\omega$ can be represented as a word structure $W_\alpha = (\mathbb{N}, <, (P_a)_{a\in\Gamma})$ over the linear ordering of the natural numbers with monadic predicates $P_a = \{i \in \mathbb{N} : \alpha(i) = a\}$. W_α has always an automatic presentation with α itself as a parameter.

$$\mathfrak{d} = (\nu, \Sigma := \{b\}, \Gamma := \{a_0, \ldots, a_r\}, L_\Delta, L_\approx, L_<, (L_a)_{a\in\Gamma}, \alpha)$$

with $L_\Delta = b^*$, $\nu(b^i) := i$, $L_\approx := \begin{bmatrix} b \\ b \end{bmatrix}^*$, $L_< = \begin{bmatrix} b \\ b \end{bmatrix}^* \begin{bmatrix} \square \\ b \end{bmatrix}^+$, $L_a := L'_a[\alpha]$ with $L'_a := \left(\begin{bmatrix} b \\ a_0 \end{bmatrix} + \ldots + \begin{bmatrix} b \\ a_r \end{bmatrix}\right)^* \begin{bmatrix} b \\ a \end{bmatrix}$.

Definition 3.2.2. Let τ,σ be relational signatures. A k-dimensional (σ, τ)-Interpretation with p parameters is given by a sequence of σ-formulae

$$\mathcal{I}(\overline{z}) := (\Delta, \psi_\approx, (\psi_R)_{R\in\tau})$$

, where

- $\Delta(\overline{x}, \overline{z})$ is the domain formula with $\overline{x} := x_1, \ldots, x_k$ and $\overline{z} := z_1, \ldots, z_p$
- $\psi_\approx(\overline{x}_1, \overline{x}_2, \overline{z})$ is the equality formula
- $\psi_R(\overline{x}_1, \ldots, \overline{x}_{ar(R)}, \overline{z})$ are the relation formulae for every $R \in \sigma$ with arity $ar(R)$.

Let $\mathfrak{B}$ be a σ-structure. For ever l-tuple of parameters $\overline{b}$ in $\mathfrak{B}$, $\mathcal{I}$ defines a $\tau \cup \{\approx\}$-structure $\mathcal{I}(\mathfrak{B}, \overline{b}) = (\Delta^{\mathfrak{B},\overline{b}}, \psi_\approx^{\mathfrak{B},\overline{b}}, (\psi_R^{\mathfrak{B},\overline{b}})_{R\in\tau})$ in $\mathfrak{B}$.

- $\mathcal{I}(\overline{z})$ interprets the τ-structure $\mathfrak{A}$ in $\mathfrak{B}$ with parameters $\overline{b}$ if and only if $\mathcal{I}(\mathfrak{B}, \overline{b})/\psi_{\approx}^{\mathfrak{B},\overline{b}} \cong \mathfrak{A}$, which is denoted by $\mathfrak{A} \leq_{FO}^{\mathcal{I}(\overline{b})} \mathfrak{B}$ or $\mathfrak{A} \leq_{FO}^{\mathcal{I}} (\mathfrak{B}, \overline{b})$

The significance of interpretations is that they can be used to effectively translate formulae from the signature of the interpreted structure into the signature of the interpreting structure. This enables one in particular to reduce the model checking problem for a structure $\mathfrak{A}$ and a logic to the corresponding model checking of the interpreting structure. The formulae translation is accomplished by substituting atomic formulae by their corresponding defining formulae in the interpreting signature and relativizing the quantors to the universe that is defined by Δ. $\mathcal{I}(\overline{z})$ induces a formulae translation $\phi(\overline{x}) \mapsto \phi^{\mathcal{I}(\overline{z})}(\overline{x}_1, \ldots, \overline{x}_k, \overline{z})$ from τ- to σ-formulae, such that for any epimorphism $\delta\colon \mathcal{I}(\mathfrak{B}, \overline{b}) \to \mathfrak{A} := (A, =, (R^{\mathfrak{A}})_{R\in\tau})$

$$\mathfrak{A} \models \phi(\delta(\overline{c})) \Leftrightarrow \mathfrak{B} \models \phi^{\mathcal{I}(\overline{z})}(\overline{c}, \overline{b})$$

Note that it is sufficient to use only a single parameter α in the definition of automatic presentations with parameter, because a presentation with several finitely many parameters $\alpha_i \in \Gamma_i^\omega$ for $i \in I$ can always be transformed into an equivalent one that uses only one parameter $\otimes_{i\in I}\alpha_i$ which is considered to be a single ω-word over the alphabet $\Gamma := \prod_{i\in I} \Gamma_i$.

An elegant method to characterize classes of structures is to choose a host structure $\mathfrak{A}$ and consider structures that are $\mathcal{L}$-interpretable in the host for a logic $\mathcal{L}$. A structure $\mathfrak{C}$ of a class $\mathcal{K}$ is called *complete in* $\mathcal{K}$ under the type of logical interpretation under consideration, if it is

a) itself a member of $\mathcal{K}$: $\mathfrak{C} \in \mathcal{K}$

b) every $\mathfrak{A} \in \mathcal{K}$ is $\mathcal{L}$-interpretable in $\mathfrak{A}$: $\mathfrak{A} \leq_{\mathcal{L}} \mathfrak{C}$

For the class (ω)AutStr of (ω)automatic presentable structures two interesting structures that are known to be complete under FO-Interpretations are $\mathfrak{N}_p = (\mathbb{N}, +, |_p)$ which is the extension of Presburger arithmetic by the binary predicate $|_p$ that expresses divisibility by a power of p and its analogoue $\mathfrak{R}_p$ for uncountable structures as defined below.

Definition 3.2.3. Let Σ be a finite alphabet with $|\Sigma| \geq 2$ and $p \geq 2$.
$\mathfrak{N}_p := (\mathbb{N}, +, |_p)$:

$$|_p : \forall x \forall y (x \mid_p y :\Leftrightarrow \exists n \in \mathbb{N}(x = p^n \wedge p^n \mid y))$$

$\mathfrak{W}^{\leq\omega}(\Sigma) := (\Sigma^{\leq\omega}, (\sigma_a)_{a\in\Sigma}, \leq_p, \mathrm{el})$:

$$\begin{aligned} \sigma_a &: \forall w \forall v (\sigma_a(w, v) :\Leftrightarrow v = wa) \\ \leq_p &: \forall w \forall v (w \leq_p v :\Leftrightarrow \exists x (v = wx)) \\ \mathrm{el} &: \forall w \forall v (\mathrm{el}(w, v) :\Leftrightarrow |w| = |v|) \end{aligned}$$

$\mathfrak{R}_p := (\mathbb{R}, +, |_p, \leq, 1)$:

$$|_p : \forall x \forall y (x \mid_p y :\Leftrightarrow \exists z, k \in \mathbb{Z}(x = p^z \wedge kp^z = y))$$

Theorem 3.2.1. *[6]*

1. *A k-ary relation $R \subseteq (\Sigma^{\leq\omega})^k$ is regular if and only if R is FO-definable in $\mathfrak{W}^{\leq\omega}(\Sigma)$*
2. *A k-ary relation $R \subseteq \mathbb{N}^k$ is regular in its p-adic representation if and only if R is FO-definable in $\mathfrak{N}_p$*
3. *$\mathfrak{N}_p$ and $\mathfrak{W}(\Sigma)$ are complete under FO-interpretations for* AutStr
4. *$\mathfrak{R}_p$ and $\mathfrak{W}^{\leq\omega}(\Sigma)$ are complete under FO-interpretations for ω*AutStr

Those structures can also be used to characterize automatic structures with parameters.

Corollary 3.2.1. 1. *$R \subseteq (\Sigma^\omega)^k$ is regular with advice $\otimes(\alpha_1, \ldots, \alpha_n) \in \Sigma^\omega$ if and only if R is FO-definable in $(\mathfrak{W}(\Sigma)^{\leq\omega}, \alpha_1, \ldots, \alpha_n)$*

2. *$\mathfrak{W}^{\leq\omega}(\Sigma)$ is complete under FO-interpretations with parameters for the class* ωAutStr[all].
3. *$(\mathfrak{W}^{\leq\omega}(\Sigma), \alpha_1, \ldots, \alpha_n)$ is complete under FO-interpretations for the class* ωAutStr$[\otimes(\alpha_1, \ldots, \alpha_n)]$ *for all $\alpha_1, \ldots, \alpha_n \in \Sigma^\omega$*

Proof. 1. A k-ary relation $R \subseteq (\Sigma^\omega)^k$ is regular with advice $\otimes(\alpha_1, \ldots, \alpha_n)$ if and only if there is a regular $k+n$-ary relation $R' \subseteq (\Sigma^\omega)^{k+n}$ with $R = R'[\alpha_1, \ldots, \alpha_n]$ if and only if there is a FO-formula $\varphi(x_1, \ldots, x_k, z_1, \ldots, z_p)$ such that

$$\mathfrak{W}^{\leq\omega}(\Sigma) \models \varphi(a_1, \ldots, a_k, \alpha_1, \ldots, \alpha_p) \Leftrightarrow (a_1, \ldots, a_k, \alpha_1, \ldots, \alpha_p) \in R'$$

Substitute $z_1, \ldots, z_p$ in $\varphi(x_1, \ldots, x_k, z_1, \ldots, z_p)$ by the constant symbols $\alpha_1, \ldots, \alpha_p$. Then:

$$(\mathfrak{W}^{\leq\omega}(\Sigma), \alpha_1, \ldots, \alpha_p) \models \varphi(a_1, \ldots, a_k) \Leftrightarrow (a_1, \ldots, a_k) \in R'[\alpha_1, \ldots, \alpha_n] = R$$

2.,3. To see that $(\mathfrak{W}^{\leq\omega}(\Sigma), \alpha_1, \ldots, \alpha_n) \in \omega$AutStr$[\alpha_1 \otimes \ldots \otimes \alpha_n]$ take the canonical automatic presentation of $\mathfrak{W}^{\leq\omega}(\Sigma)$ where every word w is coded by itself. The singleton sets $\{\alpha_i\}$ are regular with advice α_i and therefore also with advice $\alpha_1 \otimes \ldots \otimes \alpha_n$. If $\mathfrak{A} \in \omega$AutStr$[\alpha_1 \otimes \ldots \otimes \alpha_n]$ then for every k-ary relation R of an automatic presentation with parameter $\otimes(\alpha_1, \ldots, \alpha_n)$ there is a $k+n$-ary regular relation R' with $R'[\alpha_1, \ldots, \alpha_n] = R$, which is definable by a FO-formula $\varphi_R(x_1, \ldots, x_k, z_1, \ldots, z_n)$ in $\mathfrak{W}^{\leq\omega}(\Sigma)$. This gives us an FO-interpretation $\mathcal{I}$ with n parameters of $\mathfrak{A}$ in $\mathfrak{W}^{\leq\omega}(\Sigma)$.

□

While regular relations with advice are precisely those that are definable in the $\mathfrak{W}(\Sigma)^{\leq\omega}$ with the advice being used as a constant in the formula, it is more convenient for the specification of regular arithmetical relationships to use another structure that allows direct quantification over number sequences and access to the elements of the sequence as generalized digits.

Definition 3.2.4. Let Σ be a finite alphabet with $|\Sigma| \geq 2$ and $\textsc{Reg}(\Sigma)$ ($\textsc{Reg}(\mathbb{N})$ the class of all regular relations over Σ ($\mathbb{N}$)). Define the structures:
$\mathfrak{W}^{\leq\omega}(\Sigma^*) := ((\omega)\mathrm{Seq}(\Sigma^*), (\mathrm{dig}_R)_{R\in\textsc{Reg}(\Sigma)}, \leq_p, \mathrm{el})$
$\mathfrak{W}^{\leq\omega}(\mathbb{N}) := ((\omega)\mathrm{Seq}(\mathbb{N}), (\mathrm{dig}_R)_{R\in\textsc{Reg}(\mathbb{N})}, \leq_p, \mathrm{el})$

$$\begin{array}{ll} \mathrm{dig}_R & : \forall i \forall w_1 \ldots \forall w_k (\mathrm{dig_R}(i, w_1, \ldots, w_k) :\Leftrightarrow \\ & \begin{cases} (w_1(|i|-1), \ldots, w_k(|i|-1)) \in R & \text{if } 0 < |i| < \infty \\ \text{false} & \text{if } |i| = 0 \vee |i| = \infty \end{cases} \\ \mathrm{el} & : \forall w \forall v (\mathrm{el}(w,v) :\Leftrightarrow |w| = |v|) \\ \leq_\mathrm{p} & : \forall w \forall v (w \leq_p v :\Leftrightarrow \exists x (v = wx)) \end{array}$$

The universe of the structure $\mathfrak{W}^{\leq\omega}(\Sigma^*)$ is the set of all finite and infinite word sequences over the alphabet Σ. In addition to the prefix and equal length relations on sequences, that are defined analogously to the case of words and ω-words, $\mathfrak{W}^{\leq\omega}(\Sigma^*)$ has the $k+1$-ary relation $\mathrm{dig_R}(i, x_1, \ldots, x_k)$ for every k-ary regular relation $R \subseteq (\Sigma^*)^k$. $\mathrm{dig_R}(i, x_1, \ldots, x_k)$ expresses that the k-tuple of the i-th components of the sequences $(x_1, \ldots, x_k)$ is an element of R. i is hereby used as an index variable, which is however interpreted as a finite sequence in the structure. The last component of i is thereby used to indicate the position in the tuple that is being adressed. For infinite sequences i or the empty sequence, the expression $\mathrm{dig_R}(i, x_1, \ldots, x_k)$ evaluates to 'false'. By convention a formula of the form $\forall i \phi(i, x_1, \ldots, x_k)$, $\exists i \phi(i, x_1, \ldots, x_k)$ where i occurs as an index variable in a subformula of type $\mathrm{dig_R}(i, x_1, \ldots, x_k)$ are therefore interpreted as their relativized versions $\forall i(\phi_{\mathrm{fin}}(i) \rightarrow \phi(i, x_1, \ldots, x_k))$ and $\exists i(\phi_{\mathrm{fin}}(i) \wedge \phi(i, x_1, \ldots, x_n))$, where $\phi_{\mathrm{fin}}(i) := \exists x(x \neq i \wedge x \leq_p i \wedge \forall y(y \leq_p i \rightarrow y \leq_p x))$ defines the set of finite non-empty sequences in $\mathfrak{W}(\Sigma^*)$. Furthermore formulas of the type $(x_1, \ldots, x_k) \in R$, where R is a k-ary regular relation over words as well as other formulas that define word relations may be used in place of the predicate $\mathrm{dig_R}$. Thus for instance the formula $\phi(x, i) := x(i) \neq y(i)$ may be used, because the set $\{(w, v) \in \Sigma^* \times \Sigma^* : w \neq v\}$ is a regular relation. It is more intuitive to use the index-notation $x(i)$ as an abbrevation for $x(|i|-1)$, which indices the $|i|-1$-th component of the sequence x, where $|i|$ is the length of the sequence i. $x(i)$ is also defined, if x is only a finite sequence and i longer than x . In this case the value of $x(i)$ consists of a sequence of padding symbols $\Box$. Here are a few additional macros:

Definition 3.2.5. "$|x| = |y|$" $\equiv \mathrm{el}(x, y)$

"$|x| \leq |y|$" $\equiv \exists z(\mathrm{el}(z, x) \wedge z \leq_p x)$

"$y <_p x$" $\equiv (\neg(y = x) \wedge y \leq_p x)$

"$\mathrm{pred}(x, y)$" $\equiv (y <_p x \wedge \forall z(z <_p x \rightarrow z \leq_p y))$

$"\sigma_L(x,y)" \equiv \mathrm{pred}(y,x) \wedge y(y) \in L$

$"|y| = 0" \equiv \neg\exists z \mathrm{pred}(y,z)$

$"|y| = |x| + n" \equiv \exists x_1 \exists x_2 \ldots \exists x_{n+1}(\bigwedge_{i=1}^{n} \mathrm{pred}(x_i, x_{i+1}) \wedge x_1 = y \wedge x_{n+1} = x)$, for $n > 0$

$"|y| = |x| - n" \equiv |x| = |y| + n$, for $n > 0$

$"x(i) = w" \equiv \mathrm{dig}_{\{w\}}(i,x)$

$"\varphi(x(i \pm n), \overline{y})" \equiv \exists j(|j| = |i| \pm n \wedge \varphi(x(j), \overline{y}))"$

$"i = \max(i_0, \ldots, i_k)" \equiv \bigwedge_{j=0}^{k} |i_j| \leq |i| \wedge \bigvee_{j=1}^{k} |i| = |i_j|$

$\mathfrak{W}^{\leq\omega}(\Sigma^*)$ and $\mathfrak{W}^{\leq\omega}(\mathbb{N})$ themselves are not word-automatic presentable with parameter which will be shown in chapter 4. We are mainly interested in a certain class of its substructures.

Definition 3.2.6. Let $\bar{l} \in \omega\mathrm{Seq}(\mathbb{N}^{>0})$ be a length-sequence. The $\bar{l}$-restrictions

$$\bar{l}\text{-}\mathfrak{W}^{\leq\omega}(\Sigma^*) \text{ and } \bar{l}\text{-}\mathfrak{W}^{\leq\omega}(\mathbb{N})$$

are the substructures of

$$\mathfrak{W}^{\leq\omega}(\Sigma^*) \text{ and } \mathfrak{W}^{\leq\omega}(\mathbb{N})$$

that are generated by the subsets

$$\bar{l}\text{-}(\omega)\mathrm{Seq}(\Sigma^*) = \{\sigma \in \mathrm{Seq}(\Sigma^*) : \forall i \in \mathrm{Def}(\varsigma)(|\varsigma(i)| = \bar{l}(i))\}$$

and $\bar{l}\text{-}(\omega)\mathrm{Seq}(\mathbb{N}) = \{\eta \in (\omega)\mathrm{Seq}\mathbb{N} : \forall i \in \mathrm{Def}(\eta)(\eta(i) < 2^{l(i)-1})\}$.

Theorem 3.2.2. *Let $\alpha \in \Sigma^\omega$ with $\# \in \mathrm{Inf}(\alpha)$ and $(\alpha)_\# \in \bar{l}\text{-}\mathrm{Seq}(\Sigma^*)$, $\bar{l} = l(\alpha)$*

1. *$R \subseteq \bar{l}\text{-}\mathrm{Seq}(\Sigma^*)^k$ is regular with advice $(\alpha)_\#$ iff R is FO-definable in $\bar{l}\text{-}\mathfrak{W}(\Sigma^*)$ with parameter $(\alpha)_\#$.*
2. *$R \subseteq \bar{l}\text{-}\omega\mathrm{Seq}(\Sigma^*)^k$ is regular with advice $(\alpha)_\#$ iff R is FO-definable in $\bar{l}\text{-}\mathfrak{W}^\omega(\Sigma^*)$ with parameter $(\alpha)_\#$.*
3. *$\mathfrak{A} \in \mathrm{AutStr}[\alpha]$ iff $\mathfrak{A} \leq_{FO} (\bar{l}\text{-}\mathfrak{W}(\Sigma^*), (\alpha)_\#)$*
4. *$\mathfrak{A} \in \omega\mathrm{AutStr}[\alpha]$ iff $\mathfrak{A} \leq_{FO} (\bar{l}\text{-}\mathfrak{W}^\omega(\Sigma^*), (\alpha)_\#)$*

Proof. 1. $\Leftarrow$: Using the closure properties of the class of regular relations over $\bar{l}$-Seq(Σ^*) it is enough to show that atomic formulae define regular relations with advice $(\alpha)_\#$ in $(\bar{l}\text{-}\mathfrak{W}(\Sigma^*), (\alpha)_\#)$.For the specification of the transition relations of the sequence automata, we can use FO-formulae over $\mathfrak{W}(\Sigma_\Box)$ with a finite parameter, i.e. and additional variable z.

$x = y$**:** $\mathcal{G}_= := (\{q_0\}, \Sigma_\Box, \Sigma, \varphi_{q_0,q_0}, q_0, \{q_0\})$ where $\varphi_{q_0,q_0}(x,y,z) := x = y$

$x \leq_p y$: $\mathcal{G}_{\leq_p} := (\{q_0, q_1\}, \Sigma_\Box, \Sigma, (\varphi_{q_i,q_j})_{i,j\in Q}, q_0, \{q_0, q_1\})$ where
$\varphi_{q_0,q_0}(x, y, z) := x = y \wedge y \neq \Box^{|z|}$, $\varphi_{q_1,q_1} = \varphi_{q_0,q_1}(x, y, z) := x = \Box^{|z|}$

el(x, y): $\mathcal{G}_{\mathrm{el}} := (\{q_0\}, \Sigma_\Box, \Sigma, \varphi_{q_0,q_0}, q_0, \{q_0\})$ where
$\varphi_{q_0,q_0}(x, y, z) := x \neq \Box^{|z|} \wedge y \neq \Box^{|z|}$

dig$_R$: $\mathcal{G}_{\mathrm{dig}_R} := (\{q_0, q_1, q_2\}, \Sigma_\Box, \Sigma, (\varphi_{q_i,q_j})_{i,j\in Q}, q_0, \{q_2\})$ where
$\varphi_{q_0,q_0}(i, x_1, \ldots, x_k, z) := \bigwedge_{i=1}^k x_i \neq \Box^{|z|} \wedge i \neq \Box^{|z|}$,
$\varphi_{q_0,q_1}(i, x_1, \ldots, x_k, z) := i \neq \Box^{|z|} \wedge (x_1, \ldots, x_k, z) \in R$
$\varphi_{q_1,q_2}(i, x_1 \ldots, x_k, z) := i = \Box^{|z|}$
$\varphi_{q_2,q_2}(i, x_1 \ldots, x_k, z) :=$ true

1. $\Rightarrow$: Let $\mathcal{G} = (Q = \{q_0, \ldots, q_r\}, \Sigma_\Box^k, \Sigma, (\varphi_{p,q\in Q})_{p,q\in Q}, q_0, F)$ be a sequence automaton that recognizes R with advice $(\alpha)_\#$, where $\varphi_{p,q}(x_1, \ldots, x_k, z)$ are FO-formula over $\mathfrak{W}(\Sigma_\Box)$. We construct an FO-formula $\psi_\mathcal{G}(x_1, \ldots, x_k, z)$ so that for all $(\omega_1, \ldots, \omega_k) \in \bar{l}\text{-Seq}(\Sigma^*)$:

$$\bar{l}\text{-}\mathfrak{W}(\Sigma^*) \models \psi_\mathcal{G}(\omega_1, \ldots, \omega_k, (\alpha)_\#) \Leftrightarrow (\omega_1, \ldots, \omega_k) \in L(\mathcal{G}[(\alpha)_\#])$$

$$\psi_\mathcal{G}(x_1, \ldots, x_k, z) := \exists y_{q_0} \exists y_{q_1} \ldots \exists y_{q_r} (\mathrm{INIT} \wedge \mathrm{RUN} \wedge \mathrm{ACCEPT})$$

The formula expresses that there is an accepting run of $\mathcal{G}$ on $(x_1, \ldots, x_k, z)$. Let $a_1 \in \Sigma_\Box$. The tuple $(y_0, \ldots, y_r)$ codes a sequence of states with $y_{q_j}(i) = a_1$ iff the automaton can reach state q_j in its i-th step.

$$\mathrm{INIT} := y_{q_0}(0) = a_1 \wedge \bigwedge_{j=1}^r y_{q_j}(0) \neq a_1$$

Expresses that the automaton starts in its initial state q_0.

$$\begin{aligned}\mathrm{RUN} := \forall i(\quad & i < \max(x_1, \ldots, x_k) \rightarrow \\ & \bigwedge_{p,q\in Q} (y_p(i) = a_1 \wedge \varphi_{p,q}(x_1(i), \ldots, x_k(i), z(i)) \leftrightarrow y_q(i+1) = a_1))\end{aligned}$$

Defines the transitions at each position i.

$$\mathrm{ACCEPT} := \exists e(e = \max(x_1, \ldots, x_r) \wedge \bigvee_{q\in F} y_q(e) = a_1)$$

Expresses that an accepting state can be reached in the last step.

2. For ω-sequences the proof is analogous and simpler since the index variable i doesn't need to be restricted. One merely has to specify a different acceptance condition:

$$\begin{aligned}\omega\mathrm{ACCEPT} := \bigvee_{F\in\mathcal{F}} (\quad & \bigwedge_{q\in\mathcal{F}} \forall i \exists j (|i| < |j| \wedge y_q(j) = a_1) \wedge \\ & \bigwedge_{q\notin\mathcal{F}} \neg\forall i \exists j (|i| < |j| \wedge y_q(j) \neq a_1))\end{aligned}$$

3.,4. As a consequence of proposition 2.2.4 for every automatic presentation with parameter

$$\mathfrak{d} = (\nu, \Sigma, \Sigma, L_\Delta, L_\approx, (L_R)_{R\in\tau}, \alpha)$$

of $\mathfrak{A}$ there is an equivalent presentation with sequence automata and due to 1. and 2. there are formulae Δ, $\psi_\approx$, $(\psi_R)_{R\in\tau}$ that define the regular relations of the presentation in $\bar{l}$-$\mathfrak{W}^{\leq}(\Sigma^*)$, so that one can construct a 1-dimensional interpretation with 1 parameter

$$\mathcal{I}(z) := (\Delta, \psi_\approx, (\psi_R)_{R\in\tau})$$

that interprets $\mathfrak{A}$ in $\bar{l}$-$\mathfrak{W}^{\leq\omega}(\Sigma^*)$ with parameter $(\alpha)_\#$ and vice versa.

□

By using the coding $\overline{\text{bin}}$ of number sequence as word sequences over the alphabet $\{0, 1, \#, \square\}$ we get in the same way:

Theorem 3.2.3. *Let* $\bar{l} \in \omega\text{Seq}(\mathbb{N}^{>0})$ *and* $\eta \in \omega\text{Seq}(\mathbb{N})$ *with* $\eta(i) < 2^{l(i)-1}$ *for all* $i \in \mathbb{N}$

1. $R \subseteq \bar{l}\text{-Seq}(\mathbb{N})^k$ *is regular with advice sequence* η *iff* R *is FO-definable in* $\bar{l}$-$\mathfrak{W}(\mathbb{N})$ *with parameter* η.
2. $R \subseteq \bar{l}\text{-}\omega\text{Seq}(\mathbb{N})^k$ *is regular with advice* η *iff* R *is FO-definable in* $\bar{l}$-$\mathfrak{W}^\omega(\mathbb{N})$ *with parameter* η.
3. $\mathfrak{A} \in \text{AutStr}[\otimes_{\bar{l}}\eta]$ *iff* $\mathfrak{A} \leq_{FO} (\bar{l}\text{-}\mathfrak{W}(\mathbb{N}), \eta)$
4. $\mathfrak{A} \in \omega\text{AutStr}[\otimes_{\bar{l}}\eta]$ *iff* $\mathfrak{A} \leq_{FO} (\bar{l}\text{-}\mathfrak{W}^\omega(\mathbb{N}), \eta)$

Chapter 4

Torsion-free Abelian Groups

The integer ring $(\mathbb{Z}, +, \cdot)$ acts on the elements a of an abelian group $(A, +)$ via the operation $n \cdot a := \underbrace{a + a + \cdots + a}_{n \text{ times}}$ and $-n \cdot a := n \cdot -a$ for $n \geq 0$, whereby abelian groups can be treated as $\mathbb{Z}$-modules. A subset X of an abelian group is called *linearly independent* (over $\mathbb{Z}$), if it has no non-trivial finite linear combination of 0, i.e. for any $x_1, \cdots, x_k \in X$ $z_1 \cdot x_1 + \cdots + z_k \cdot x_k = 0$ implies $z_1 = \cdots = z_k = 0$. It is an elementary result of group theory that all maximal linearly independent subsets of an abelian group have the same cardinality. This unambigous cardinal number of a maximal linearly independent subset is called the *rank* of A and denoted by rank(A). An abelian group is called *torsion-free* if all its non-zero elements have non-finite order, i.e. $z \cdot a = 0$ holds only for $z = 0$ or $a = 0$. We say that an integer $z \neq 0$ *divides* a and write $z \mid a$ if there exists an element $x \in A$ with $z \cdot x = a$. In torsion-free groups there is at most one such x, which will be denoted by $x := \frac{a}{z}$ if it exists. For any prime number $p \in \mathbb{P}$ the *p-height of* a is defined as $h_p(a) := \sup\{k \in \mathbb{N} : p^k \mid a\} \in \mathbb{N} \cup \{\infty\}$. The characteristic of an element a is the sequence $\chi(a) := (h_p(a))_{p \in \mathbb{P}}$ of its p-heights. In the additive group $(\mathbb{Q}, +)$ for instance every element has the same characteristic $(\infty, \infty, \cdots)$, because every element is arbitrarily often divisible by any prime number.

Using the notion of the characteristic of an element it is easy to classify all subgroups of $(\mathbb{Q}, +)$, known as the *rational groups.* The rational groups are up to isomorphism precisely the torsion-free abelian groups of rank 1. We show that the rational groups are automatic presentable with parameters.

4.1 Rational groups

Definition 4.1.1. Every characteristic sequence $\overline{c} = (c_p)_{p \in \mathbb{P}} \in (\mathbb{N} \cup \{\infty\})^{\mathbb{P}}$ induces a *rational group*

$$\mathbb{Q}_{\overline{c}} := \left\langle \frac{1}{p^e} \mid p \in \mathbb{P}, e \leq c_p \right\rangle$$

Here are some elementary results of abelian group theory.

Lemma 4.1.1. *Let $(A, +)$ be a torsion-free abelian group. For all $a, b \in A$,$z, z' \in \mathbb{Z} \setminus \{0\}$ and primes $p \in \mathbb{N}$*

1. $p \mid z \cdot a \implies p \mid z \vee p \mid a$

2. $h_p(z \cdot a) = h_p(z) + h_p(a)$

Proof. 1. Let $p \mid z \cdot a$, i.e. $z \cdot a = p \cdot x$,for a $x \in A$. Suppose $p \nmid z$. According to Euclid's theorem there are then integers u, v with $up + vz = 1$. Thus $p \cdot (u \cdot a + \cdot a) = up \cdot a + vz \cdot a = a$, i.e. $p \mid a$.

2. From 1. can be easily inferred that $p^e \mid z \cdot a \Leftrightarrow \exists e_1 \exists e_2 (e_1 + e_2 = e \wedge p^{e_1} \mid z \wedge p^{e_1} \mid a)$, which is equivalent to $h_p(z \cdot a) = h_p(z) + h_p(a)$.

$\square$

Lemma 4.1.2. $\left\langle \frac{1}{p^e} \mid p \in \mathbb{P}, e \leq c_p \right\rangle = \left\{ \frac{z}{p_1^{e_1} \ldots p_k^{e_k}} : e_i \leq c_{p_i}, p_i \in \mathbb{P}, z \in \mathbb{Z} \right\}$

Proof. $\subseteq$: Let $r := \frac{z_1}{p_1^{e_1}} + \ldots + \frac{z_s}{p_s^{e_s}}$ for some $z_i \in \mathbb{Z}$ with primes p_i and exponents $e_i \leq c_{p_i}$. Then $r = \frac{z_1 d_1 + \ldots + z_s d_s}{p_1^{e_1} \ldots p_s^{e_s}}$ with $d_i = \frac{p_1^{e_1} \ldots p_s^{e_s}}{p_i^{e_i}}$ for $i = 1, \ldots, s$.

$\supseteq$: Let $r := \frac{z}{p_1^{e_1} \ldots p_s^{e_s}}$ with $z \in \mathbb{Z}$ and $1 \leq e_i \leq c_{p_i}$ for $i = 1, \ldots, s$. Then the numbers $d_i = \frac{p_1^{e_1} \ldots p_s^{e_s}}{p_i^{e_i}}$ are pairwise co-prime. According to Euclid's theorem there are integers $z_1, \ldots, z_s$ with $z_1 d_1 + \ldots z_s d_s = 1$. Thus $\frac{z}{p_1^{e_1} \ldots p_s^{e_s}} = \frac{z(z_1 d_1 + \ldots z_s d_s)}{p_1^{e_1} \ldots p_s^{e_s}} = \frac{z z_1 d_1}{p_1^{e_1}} + \ldots + \frac{z z_s d_s}{p_s^{e_s}}$. $\square$

The characterisation of the rational groups was first discovered by Reinhold Baer in [4].

Proposition 4.1.1. *Every subgroup A of $\mathbb{Q}$ has the form $n\mathbb{Q}_{\bar{c}}$ for some characteristic sequence $\bar{c}$ and $n \in \mathbb{N}$.*

Proof. Let A be a subgroup of $\mathbb{Q}$. Since $A \cap \mathbb{Z}$ is a subgroup of $\mathbb{Z}$ there exists a number $n \in \mathbb{N}$, such that $A \cap \mathbb{Z} = n\mathbb{Z}$. We show that $A = n\mathbb{Q}_{\chi(n)}$.

$n\mathbb{Q}_{\chi(n)} \subseteq A$: For any prime number p and $e \leq h_p(n)$ the quotient $\frac{n}{p^e}$ is in A and therefore $n\mathbb{Q}_{\chi(n)} = \left\langle \frac{n}{p^e} \mid p \in \mathbb{P}, e \leq h_p(n) \right\rangle \subseteq A$.

$A \subseteq n\mathbb{Q}_{\chi(n)}$: Let $a = \frac{z}{b} \in A$ for co-prime integers $z \neq 0$ and $b > 0$. Then $b \cdot a = z \in A \cap \mathbb{Z} = n\mathbb{Z} \Rightarrow z = z'n$ for an integer z'. Let $b = p_1^{e_1} \cdots p_k^{e_k}$ be the prime factorization of b. Then $p_i^{e_i}$ divides $z = z'n$ for every i, i.e. $h_{p_i}(z'n) = h_{p_i}(z') + h_{p_i}(n) \geq e_i$ for all i according to Lemma 4.1.1. Thus $e_i \leq h_{p_i}(n)$ and $a = \frac{z'n}{p_1^{e_1} \ldots p_k^{e_k}} \in n\mathbb{Q}_\chi(n)$ according to Lemma 4.1.2. $\square$

Note that 1 in the group $\mathbb{Q}_{\bar{c}}$ has the characteristic $\chi(1) = \bar{c}$. We can also determine the isomorphism classes of the rational groups.

Proposition 4.1.2. *Two rational groups $n\mathbb{Q}_{\bar{c}}$ and $m\mathbb{Q}_{\bar{d}}$ are isomorphic if and only if $\bar{c}$ and $\bar{d}$ differ in only finitely many places and $\forall p (c_p = \infty \leftrightarrow d_p = \infty)$*

Proof. $\Leftarrow$: First notice that $r\mathbb{Q}_{\bar{e}}$ and $\mathbb{Q}_{\bar{e}}$ are isomorphic via $q \mapsto rq$ for all $r \in \mathbb{Q}$, so we can assume wlog $n = m = 1$ and that therefore $\bar{c}$,$\bar{d}$ are equal to the characteristic of 1 in both groups. Now let $p_1, \ldots, p_k$ be the primes on which c and d differ and $\Delta_i := d_{p_i} - c_{p_i}$ for $i = 1, \ldots, k$. Then $p_1^{\Delta_1} \ldots p_k^{\Delta_k} \left\langle \frac{1}{p^e} \mid p \in \mathbb{P}, e \leq d_p \right\rangle = \left\langle \frac{1}{p^e} \mid p \in \mathbb{P}, e \leq c_p \right\rangle$,

i.e. $p_1^{\Delta_1} \dots p_k^{\Delta_k} \mathbb{Q}_{\overline{d}} = \mathbb{Q}_{\overline{c}}$.
$\Rightarrow$: If $\overline{c}$ and $\overline{d}$ differ in infinitely many places, then there are infinitely many prime powers $p_1^{e_1}, p_2^{e_2}, \dots$ that divide 1 in one group, but not in the other. Towards a contradiction let us assume there were an isomorphism ϕ between the groups with $\phi(1) = r_0$. Then $\phi(p_i^{e_i}) = p_i^{e_i}\phi(1)$ and so all of those prime powers divide r_0. Since the groups have rank 1 , there exist integers $z_0, z_1 \neq 0$ with $z_0 \cdot r_0 = z_1 \cdot 1$. According to Lemma 4.1.1 $h_p(z_0 \cdot r_0) = h_p(z_0) + h_p(r_0) = h_p(z_1 \cdot 1) = h_p(z_1) + h_p(1)$ and for the infinitely many primes that don't divide z_0 or z_1 $h_{p_i}(r_0) = h_{p_i}(1)$ which means that infinitely many of those prime powers also divide 1 in the other group. Contradiction! $\square$

The classification problem for torsion-free abelian groups of higher ranks than 1 appears in contrast to the rank 1 case to be much more intricate. As in the rank 1 case it can be shown that the torsion-free abelian groups of rank n coincide up to isomorphism with the subgroups of $\mathbb{Q}^n$, but a characterization of the isomorphism classes of those subgroups is one of the main research programs and an open problem in the field of infinite abelian group theory [20], [11].

We now consider automatic representations of the rational groups. The idea is to find a representation of rational numbers $r < 1$ in the form $r = \sum_{i=1}^{k} a_i \frac{1}{b_i}$, where the a_i are natural numbers so that the summation of sequences $(a_0, a_1, \dots, a_k)$ can be computed by a finite advice automaton. The b_i are referred to as a *numeration base*. There is a uniform way to find a suitable base for all rational groups, which only depends on their characteristic sequence.

Proposition 4.1.3. *Let $\overline{c}$ be a characteristic sequence and $(n_i)_{i\geq 0}$ a sequence of natural numbers with*

a) $\forall i : n_i \geq 2$

c) $\forall p \in \mathbb{P} : \sum_{i=0}^{\infty} h_p(n_i) = c_p$, *where* $h_p(n_i) := \max\{e : p^e \mid n_i\}$.

Let $(b_i)_{i\geq 0}$ be the sequence $b_i := \prod_{j\leq i} n_j$.
Then for every $r \in \mathbb{Q}_{\overline{c}}$ with $r \geq 0$ there exists a unique pair of sequences $((a_i)_{0\leq i\leq k}, (d_j)_{0\leq j\leq l})$ such that

1. $\lfloor r \rfloor = d_0 + \sum_{i=1}^{l} d_i b_{i-1}$

2. $\{r\} = \sum_{i=0}^{k} a_i \frac{1}{b_i}$

3. $\forall i : 0 \leq a_i, d_i < n_i$

4. *k and l are minimal*

Proof. Let $r \in \mathbb{Q}_c$ with $r \geq 0$. Decompose r into its integral and fractional part $r = m+f$, with $m \in \mathbb{N}$ and $0 \leq f < 1$.

(Existence) :

First we show that m has a representation of the desired form. We prove it by induction over m. If $m < n_0 = b_0$, then set $l := 0$ and $d_0 := m$, which gives us the desired representation. Otherwise $m \geq n_0$ and we choose the largest s with $m \geq b_s$, i.e. $b_{s+1} > m \geq b_s$. Set $d_{s+1} := \left\lfloor \frac{m}{b_s} \right\rfloor$. Then $1 \leq d_{s+1} \leq \frac{m}{b_s} < \frac{b_{s+1}}{b_s} = n_{s+1}$ and $m' := m - d_{s+1}b_s < b_s$. By the induction hypothesis m' has a representation $m' = d_0 + \sum_{i=1}^{l'} d_i b_{i-1}$ with $0 \leq l' < s+1$. Then $m = m' + d_{s+1}b_s = d_0 + \sum_{i=1}^{l'} d_i b_{i-1} + d_{s+1}b_s$ is a representation with $0 \leq d_i < n_i$ and $l := s+1$ as requested.

Now we prove that also f has a representation in this base. According to Lemma 4.1.2 $f = \frac{z}{p_1^{e_1} \ldots p_s^{e_s}}$ with $e_i \leq c_i$. Since $\sum_{i=0}^{\infty} h_{p_j}(n_i) = c_{p_j} \geq e_j$ for $j = 1, \ldots, s$, there is a $K \geq 0$ so that $h_{p_j}(n_0 \ldots n_K) = h_{p_j}(n_0) + \ldots + h_{p_j}(n_K) \geq e_j$ for $j = 1, \ldots, s$. Choose the smallest such K. Then $p_1^{e_1} \ldots p_s^{e_s}$ divides $n_0 \ldots n_K$, i.e. $n_0 \ldots n_K = p_1^{e_1} \ldots p_s^{e_s} d$ for some natural $d > 0$. Thus $f = \frac{zd}{n_0 \ldots n_K}$. We prove by induction over K that any such number has a representation in base $(b_i)_{0 \leq i}$. If $K = 0$, then set $k := 0$ and $a_0 := zd < n_0$. Let $K > 0$. If $zd < n_K$, then we are done, because then $f = a_K \frac{1}{b_K}$ is a representation of the desired form with $a_K := zd < n_K$, $a_i := 0$ for $i < K$ and $k := K$. Let $zd \geq n_K$. Euclidian division of zd by n_K gives us $zd = qn_K + a_K$ for some $0 \leq a_K < n_K$ and $qn_K \geq 0$. So $f = \frac{zd}{n_0 \ldots n_K} = \frac{qn_K + a_K}{n_0 \ldots n_K} = \frac{q}{n_0 \ldots n_{K-1}} + a_K \frac{1}{n_0 \ldots n_K}$. Apply the induction hypothesis to $\frac{q}{n_0 \ldots n_{K-1}}$ to get $\frac{q}{n_0 \ldots n_{K-1}} = \sum_{i=0}^{k'} a_i \frac{1}{b_i}$, so that $(a_i)_{0 \leq i \leq k'}$ satisfies the conditions and $k' \leq K-1$. Then $(a_i)_{0 \leq i \leq k}$ with $k := K$ and $a_i := 0 \; \forall k' < i < k$ satisfies the conditions.

(Uniqueness) :

Suppose there were two different representations $(d_i)_{0 \leq i \leq l}$ and $(d_i')_{0 \leq i \leq l}$ of m with minimal l. Let j be the first index, where they differ.
Since $d_0 = d_0 + \sum_{i=1}^{l} d_i b_{i-1} \pmod{n_0} = d_0' + \sum_{i=1}^{l} d_i' b_{i-1} \pmod{n_0} = d_0'$, j is larger than 0. Then
$\frac{(d_0 + \sum_{i=1}^{l} d_i b_{i-1}) - (d_0' + \sum_{i=1}^{l} d_i' b_{i-1})}{n_0 \ldots n_{j-1}} \pmod{n} = d_j - d_j' + \sum_{i=j+1}^{l} (d_i - d_i') b_{i-1} \pmod{n_j} = d_j - d_j' = 0 \Rightarrow d_j = d_j'$, because $d_j, d_j' < n_j$. Contradiction!

Suppose $0 \leq f < 1$ had two different representations $(a_i)_{0 \leq i \leq k}$, $(a_i')_{0 \leq i \leq k}$ with minimal k. Let r bet the smallest number with $a_r \neq a_r'$. Then $0 = \sum_{i=0}^{k} a_i \frac{1}{b_i} - \sum_{i=0}^{k} a_i' \frac{1}{b_i} = a_r - a_r' + \Delta$, where $|\Delta| \leq \sum_{i=r+1}^{k} \frac{|a_i - a_i'|}{b_i} \leq \sum_{i=r+1}^{k} \frac{n_i}{b_i} = \sum_{i=r+1}^{k} \frac{1}{b_{i-1}} \leq \sum_{i=r+1}^{k} \frac{1}{2^{i-1}} < 1$. Since $\Delta = a_r' - a_r \in \mathbb{Z} \Rightarrow \Delta = 0$ and $a_r = a_r'$. Contradiction!

□

The previous proof in the case of natural numbers is based on what is known as the "‘greedy" algorithm in the theory of numeration system [3] which investigates the variety of possibilities to represent natural and complex numbers.

Example 4.1.1. 1. In chapter 1 2.2.7 the factorial representation of a natural number has already been mentioned. It corresponds to the base sequence $n_i := i + 2$ for all

$i \geq 0$, so that $b_k = n_0 n_1 \ldots n_k = 2 \cdot 3 \ldots \cdot (k+2) = (k+2)!$ for all $i, k \geq 0$. Since the sums $h_p(1) + h_p(2) + \ldots$ tend to ∞ for all $p \in \mathbb{P}$, the corresponding characteristic sequence is the sequence $\overline{c} = (\infty, \infty, \ldots)$ that characterizes the torsion-free rank 1 group $(\mathbb{Q}, +)$ up to isomorphism. The base can therefore be used to represent all rational numbers. To calculate the representation of the number $\frac{153}{20}$ in this base, first decompose the number into integral and fractional part: $\frac{153}{20} = 7 + \frac{13}{20}$. Now expand the fraction so that the denominator becomes a factorial (which is always possible): $\frac{13}{20} = \frac{13}{4 \cdot 5} = \frac{78}{2 \cdot 3 \cdot 4 \cdot 5} = \frac{78}{5!}$. Then divide the numerator by $5 = n_3$ with remainder: $78 = 5 \cdot 15 + 3 \Rightarrow \frac{78}{5!} = \frac{5 \cdot 15 + 3}{5!} = \frac{5 \cdot 15}{5!} + \frac{3}{5!} = \frac{15}{4!} + \frac{3}{5!}$, so $a_3 := 3$ and repeat the procedure recursively for $\frac{15}{4!}$.

$$\frac{153}{20} = \begin{array}{ccccccc} 1 & + & 0 \cdot 2! & + & 1 \cdot 3! & + & \\ 1 \cdot \frac{1}{2!} & + & 0 \cdot \frac{1}{3!} & + & 3 \cdot \frac{1}{4!} & + & 3 \cdot \frac{1}{5!} \end{array}$$

which gives the sequence representation $((1,0,1),(1,0,3,3))$. Notice that 0 always has the unique representation $((0),(0))$ in every base.

2. Now consider the rational group that is generated by the prime reciprocals, i.e. $\mathbb{Q}_{\overline{c}} = \left\langle \frac{1}{p} \mid p \in \mathbb{P} \right\rangle = \left\{ \frac{z}{n} : z \in \mathbb{Z}, n \in \mathbb{N}, n \text{ is squarefree} \right\}$. It has the characteristic sequence $c = (1, 1, \ldots)$. One possible base is $\forall i \geq 0 : n_i := p_i$, where p_i is the i-th prime number of an enumeration of the primes. We could however also take any other sequence of pairwise co-prime square-free integers n_i. The only condition that the sequence n_i needs to satisfy in order to qualify as a base is $\sum_{i=0}^{\infty} h_p(n_i) = 1$ for all primes p. In other words: every prime must occur in exactly one n_i as a prime factor with exponent 1. The number $\frac{153}{20}$ has no (at least no finite) representation in this base, because the denominator of $\frac{153}{20} = \frac{153}{2^2 \cdot 5}$ isn't square-free. The number $\frac{1}{7 \cdot 11}$ has the representation $\frac{1}{7 \cdot 11} = \frac{2}{2 \cdot 3 \cdot 5 \cdot 7} + \frac{8}{2 \cdot 3 \cdot 5 \cdot 7 \cdot 11}$ in this base.

Theorem 4.1.1. *There is a parametrised FO-Interpretation*

$$\mathcal{I}[n] := (\Delta(x_s, x_d, x_a, n), \psi_+(x_s, x_d, x_a, y_s, y_a, y_d, z_s, z_a, z_d, n))$$

such that for any characteristic sequence $\overline{c}$ and any parameter $\overline{n} \in \mathbb{N}^{\mathbb{N}}$ with $\forall j \geq 0$ $\sum_{i=0}^{\infty} h_{p_j}(n_i) = c_j$ and $n_j \geq 2$ it holds that

$$\mathcal{I}(\mathfrak{W}(\mathbb{N}), \overline{n}) = (\mathbb{Q}_{\overline{c}}, +)$$

.

In particular $(\mathbb{Q}_{\overline{c}}, +) \in$ AUTSTR$[\otimes_{\#}\overline{n}]$

Proof. Non-negative numbers $r \in \mathbb{Q}_{\overline{c}}$ can be represented in $\mathfrak{W}(\mathbb{N})$ as the two coefficient sequences $(a_i)_{0 \leq i \leq k}$ and $(d_i)_{0 \leq i \leq l}$ of their base $\overline{n}$-representation for which the variables that are sub-indexed $*_a$ resp $*_d$ serve as placeholders. Additionally a sign variable $*_s$ is used that is interpreted as the singleton sequence (0) for positive or (1) for negative numbers.

Definition of the summation formula:

$$\begin{aligned} \phi_{\overrightarrow{c}}(x, y, c, i, n) &:= (c(i+1) = 1 \leftrightarrow x(i) + y(i) + c(i) \geq n(i)) \\ \phi_{\overleftarrow{c}}(x, y, c, i, n) &:= (c(i) = 1 \leftrightarrow x(i+1) + y(i+1) + c(i+1) \geq n(i+1)) \end{aligned}$$

The two variables $c_\rightarrow$ and $c_\leftarrow$ are used to represent the carry at each position. For each coefficient sequence one. The carry of the coefficient sequence that represents the fractional part of the number runs from right to left, the other one in the opposite direction.

$$\begin{aligned}\phi_{\text{init}}(x_d, y_d, c_\rightarrow, c_\leftarrow, i, n) := \quad & c_\rightarrow(i) \in \{0,1\} \wedge c_\leftarrow(i) \in \{0,1\} \wedge c_\leftarrow(max(x_d, y_d)) = 0 \wedge \\ & c_\rightarrow(0) = 1 \leftrightarrow x_d(0) + y_d(0) + c_\leftarrow(0) \geq n(0)\end{aligned}$$

Initializes the carry variables with 0 or 1.
When the carry $c_\leftarrow$ of the fractional coefficient list reaches the 0-th position it "wraps around" to the integer list and initializes the carry $c_\rightarrow$ of the integer part. $c_\leftarrow$ needs to be initialized with 0 at the rightmost position of x_d and y_d.

$$\begin{aligned}\phi_s(x, y, z, c, i, n) \quad := \quad & ((x(i) + y(i) + c(i) \geq n(i) \rightarrow x(i) + y(i) + c(i) = z(i) + n(i)) \wedge \\ & (x(i) + y(i) + c(i) < n(i) \rightarrow x(i) + y(i) + c(i) = z(i)))\end{aligned}$$

Expresses that $z(i)$ is the sum modulo $n(i)$ of the inputs $x(i)$,$y(i)$ plus carry $c(i)$.

$$\begin{aligned}\phi_+(x_d, x_a, y_a, y_d, z_a, z_d, n) := \exists c_\rightarrow \exists c_\leftarrow \forall i(\quad & \phi_{\text{init}}(x_d, y_d, c_\rightarrow, c_\leftarrow, i, n) \wedge \\ & \phi_{\underrightarrow{c}}(x_a, y_a, c_\rightarrow, i, n) \wedge \phi_{\underleftarrow{c}}(x_d, y_d, c_\leftarrow, i, n) \wedge \\ & \phi_s(x_a, y_a, z_a, c_\rightarrow, i, n) \wedge \phi_s(x_d, y_d, z_d, c_\leftarrow, i, n))\end{aligned}$$

Defines the summation of two non-negative numbers represented by (x_d, x_a) and (y_d, y_a), whereby (z_d, z_a) holds the sum and n is a parameter,i.e. the base sequence $(n_i)_{i \geq 0}$ in this case.

$$\begin{aligned}\psi_+(\overline{x}, \overline{y}, \overline{z}, n) := \quad & (x_s(0) = y_s(0) = z_s(0) \quad && \wedge \phi_+(\overline{x}, \overline{y}, \overline{z}, n)) \\ & && \vee \\ & (x_s(0) \neq y_s(0) = z_s(0) && \wedge \phi_+(\overline{z}, \overline{x}, \overline{y}, n)) \\ & && \vee \\ & (x_s(0) \neq y_s(0) \neq z_s(0) && \wedge \phi_+(\overline{y}, \overline{z}, \overline{x}, n))\end{aligned}$$

The complete summation formula.
Definition of the domain formula:

$$\phi_{(0,0)}(x_a, x_d) := |x_a| = 1 \wedge |x_d| = 1 \wedge x_a(x_a) = 0 \wedge x_d(x_d) = 0$$

Defines the sequence pair $((0),(0))$ which represents the 0:

$$\Delta_{\text{res}}(x_d, x_a, n) := \forall i(x_d(i) < n(i) \wedge x_a(i) < n(i)) \wedge |x_d| > 0 \wedge |x_a| > 0$$

Restricts the coefficient lists $(a_i)_{0 \leq i \leq k}$, $(d_i)_{0 \leq i \leq l}$ to $a_i, d_i < n_i$ and makes sure the lists are non-empty:

$$\Delta_{\text{unique}}(x_d, x_a) := (|x_d| > 1 \rightarrow x_d(x_d) \neq 0) \wedge (|x_a| > 1 \rightarrow x_a(x_a) \neq 0)$$

Ensures that the highest coefficients are non-zero, unless 0 is represented (so that the representation is unique):

$$\Delta_s(x_s, x_d, x_a) := |x_s| = 1 \wedge x_s(0) \in \{0,1\} \wedge (\phi_{(0,0)}(x_a, x_d) \rightarrow x_s(0) = 0)$$

Defines the "sign"-bit, which is set to 0 if the number 0 is represented in order to have a unique representation of 0:

$$\Delta(x_s, x_d, x_a, n) := \Delta_s(x_s, x_d, x_a) \wedge \Delta_{\text{res}}(x_d, x_a, n) \wedge \Delta_{\text{unique}}(x_d, x_a)$$

The complete domain formula: □

Corollary 4.1.1. *The elmentary theory* $\mathrm{T}_{\text{tfag1}}$ *of the torsion-free abelian groups of rank* 1 *is FO-interpretable in* $\mathfrak{W}(\mathbb{N})$ *and* $\mathfrak{R}_p := (\mathbb{R}, +, \leq, |_p, 1)$

Chapter 5

Criteria for Nonautomaticity

We will revisit some of the general methods and criteria that were developed to prove that a structure does not have an automatic presentation, see whether they still hold true when parameters are present in the presentation and adapt them if possible. An obvious necessary condition for automatic presentability of a structure is its computability and the decidability of its FOC-theory. This is a property that automatic structures with parameters don't necessarily have, since non-recursive languages can be encoded into the parameter. Combinatorial techniques that exploit specific properties of the representing automata can however be considered independently from decidability question.

5.1 Growth Arguments and Trees

A method that has been shown to be pretty versatile looks at the growth rate of finite words under regular functions or more generally locally finite relations. A relation $R \subseteq A^{k+l}$ is said to be *locally finite* if $\overline{x}R := \{\overline{y} \in A^l : (\overline{x}, \overline{y}) \in R\}$ is only a finite set for any k-tuple $\overline{x} \in A^k$. If $R \subseteq (\Sigma^*)^{k+l}$ is a regular, locally finite relation over words, then the length of any tuple $|\overline{y}|$ (i.e. the length of the convolution $\otimes\overline{y} \in (\Sigma^l_\Box)^*$ with $(\overline{x}, \overline{y}) \in R$) cannot be greater than $|\overline{x}| + C$ for a fixed constant $C > 0$ that is independent of $\overline{x}$ and $\overline{y}$. If $l = 1$ then the constant is the number of states of the automaton that recognizes R. If there were more letters than the automaton has states in the segment between the end of $\overline{x}$ and the end of y then by the pigeon hole principle for any accepting run on $\overline{x} \otimes y$ there is a state q that occurs two times on the segment between the end of the $\overline{x}$-track of the word and the end of the whole word. The subword of y that the automaton reads beginning when it enters q for the first time and ending when it reenters q can then be repeated or "pumped" arbitrarily often without changing $\overline{x}$. So that infinitely many y existed for which the automaton recognizes $(\overline{x}, y)$ contrary to the local finiteness of R. The case $l > 1$ can be reduced to $l = 1$ by applying the criteria to the relations $f_i^R := R \circ \pi_i$, where $\pi_i(y_1, \ldots, y_k) := y_i$ is the projection on the i-th component.

Lemma 5.1.1 (Blumensath [6]). *Let $\mathfrak{A} \in$ AutStr, $\mathfrak{d}$ an injective presentation of $\mathfrak{A}$ with coding function λ and $R \subseteq A^{k+l}$ a locally finite relation of $\mathfrak{A}$. Then there exists a constant $C > 0$, such that for all $(\overline{x}, \overline{y}) \in R$:*

$$|\lambda^{-1}(\overline{y})| \leq |\lambda^{-1}(\overline{x})| + C$$

This lemma doesn't hold true anymore for advice automata however, because in this case the transition relation does not just depend on state and input word but also on the position of the automaton on the advice tape and so loops cannot always be pumped without changing the acceptance behaviour of the automaton on the pumped word.

Example 5.1.1. The relation $R := \{(w, v) : w = \prod_{i=0}^{n} a^i b \wedge v = \prod_{i=0}^{n+1} a^i b\}$ is regular with advice $\alpha = \prod_{i=0}^{\infty} a^i b = bbabaabaaab\ldots$ but doesn't have the pumping property. Since $|\prod_{i=0}^{n+1} a^i b| - |\prod_{i=0}^{n} a^i b| = |a^{n+1}b| = n + 2 \to \infty$ the length difference between w and v is unbounded even though R is locally finite. In fact the gap between w and v can be made arbitrary large. Let $L := \left(\begin{bmatrix} a \\ a \\ a \end{bmatrix} + \begin{bmatrix} b \\ b \\ b \end{bmatrix} \right)^* \left(\begin{bmatrix} \square \\ a \\ a \end{bmatrix} \right)^* \begin{bmatrix} \square \\ b \\ b \end{bmatrix}$, then R=$L[\alpha]$ and for any advice $\alpha_f := \prod_{i=0}^{\infty}(a^{f(i)}b)$ where f is a function of arbitrary growth the gap between $(w, v) \in L[\alpha_f]$ has the same growth as f.

Let $R \subseteq A^2$ be a binary relation. Define recursively the i-th iteration of R via $xR^{i+1}y :\Leftrightarrow xR^iy \vee \exists z(xR^iz \wedge zRy)$ By repeatedly applying the pumping argument to a locally finite relation R the growth of the i-th stage R^i in the transitive closure of R can be bounded from above. We only state a special case of the more general lemma [6]

Lemma 5.1.2. *Let $\mathfrak{d}$ be an injective automatic presentation of a structure $\mathfrak{A}$ with coding function λ and $R \subseteq A^2$ a locally finite relation of $\mathfrak{A}$. Then for any $x \in A$:*

$$|xR^i| = 2^{O(i)}$$

Proof. According to Lemma 5.1.1 there is a constant $C > 0$ so that $|\lambda^{-1}(y)| \leq |\lambda^{-1}(x)| + C$ for all $(x, y) \in R$. By induction over i one shows that $|\lambda^{-1}(y)| \leq |\lambda^{-1}(x)| + iC$ for all $y \in xR^i$. Let Σ be the alphabet of the automatic presentation, then the number of words of length less or equal than $|\lambda^{-1}(x)| + iC$ is bounded by $|\Sigma^{|\lambda^{-1}(x)|+iC+1}| = 2^{O(i)}$. $\square$

Using this lemma we can show that a certain class of trees that are automatic presentable with parameter are not automatic presentable without parameters. This demonstrates in particular that Lemma 5.1.2 fails when parameters are allowed in the automatic presentation. For every sequence $\eta \in \omega\mathrm{Seq}(\mathbb{N})$ let $\mathfrak{T}_\eta$ denote the tree $\mathfrak{T}_\eta := (\{\varsigma \in \mathrm{Seq}(\mathbb{N}) : \forall i \in \mathrm{Def}(\varsigma)\ \varsigma(i) < \eta(i)\}, \leq_p)$ which is a substructure of the $\leq_p$-reduct of $\mathfrak{W}(\mathbb{N})$. $\mathfrak{T}_\eta$ is not automatic presentable without parameter if η grows too fast:

Proposition 5.1.1. *Let $\eta \in \omega\mathrm{Seq}(\mathbb{N}^{>0})$ and $p(i) := \prod_{j<i} \eta(j)$.*

1. $\mathfrak{T}_\eta \in \textsc{AutStr}[\otimes\eta]$
2. $\limsup_{i\to\infty} \frac{\log p(i)}{i} = \infty \Rightarrow \mathfrak{T}_\eta \notin \textsc{AutStr}$

Proof. 1. The domain of $\mathfrak{T}_\eta$ is FO-definable in $l(\eta)$-$\mathfrak{W}(\mathbb{N})$ with parameter η by the formula $\Delta(x, \eta) := \forall i(x(i) < \eta(i))$. Thus $\mathfrak{T}_\eta \leq_{FO} (l(\eta)\text{-}\mathfrak{W}(\mathbb{N}), \eta)$ and $\mathfrak{T} \in \textsc{AutStr}[\otimes\eta]$.

2. Suppose $\mathfrak{T}_\eta$ had an automatic presentation that is wlog injective. Then the successor relation S on $\mathfrak{T}_\eta$ defined by $(x,y) \in S :\Leftrightarrow x <_p y \wedge \forall z(z <_p y \rightarrow z \leq_p x)$ is regular and locally finite, since any node x with distance i from the root of the tree has exactly $\eta(i)$ successors $(x,0),\ldots,(x,\eta(i)-1)$. Let $r := ()$ be the root of $\mathfrak{T}_\eta$, i.e. the empty sequence. Then $rS^i = \{\varsigma \in \mathfrak{T}_\eta : |\varsigma| \leq i\}$ and it is $|rS^{i+1} \setminus rS^i| = |rS^i \setminus rS^{i-1}| \cdot \eta(i)$ for all $i \geq 1$. So that we get $|rS^i| = \sum_{j=1}^{i} |rS^j \setminus rS^{j-1}| + |rS^0| \geq \eta(0) \cdot \ldots \cdot \eta(i-1) = p(i)$ for all $i \geq 1$. According to Lemma 5.1.2 there exists an $i_0 \in \mathbb{N}$ and a constant $c > 0$ such that $|rS^i| \leq 2^{ci} \Rightarrow \frac{\log p(i)}{i} \leq \frac{\log |rS^i|}{i} \leq \log 2^c$ for all $i \geq i_0$ and thus $\limsup\limits_{i\to\infty} \frac{\log p(i)}{i} \leq \log 2^c < \infty$. Contradiction!

□

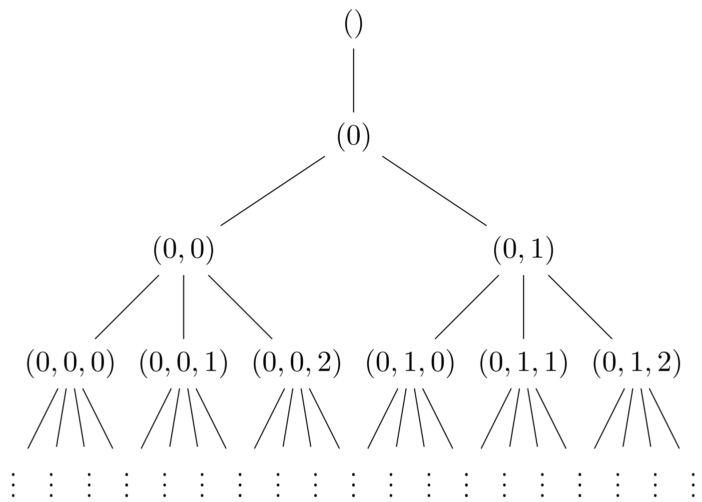

Figure 5.1: The tree $\mathfrak{T}_{(1,2,3,\ldots)}$

The failure of the growth argument for regular languages with advice is also the reason, why Tsankov's proof for the non-automaticity of $(\mathbb{Q},+)$ doesn't work anymore with paramters. For Tsankov's proof a simplified version of Freiman's theorem in additive number theory and combinatorics can be used. The field of additive combinatorics investigates the additive structures in abelian groups. One particular problem is the sum set problem: Given a constant $C > 0$, what can be said about the subsets $A \subset G$ of an abelian group, that have a subset sum $A + A := \{a + b : a, b \in A\}$ which is bounded by $|A + A| \leq C|A|$? It turns out that for torsion-free groups those sets can be contained within a generalized arithmetic progression of size not much larger than A whose rank only depends on C. Let G be a torsion-free abelian group. An arithmetic progression P of rank $d \geq 1$ in G is a set of the form $P := \{a_0 + \sum\limits_{i=1}^{d} d_i \cdot v_i : 0 \leq d_i < N_i, d_i \in \mathbb{Z}\}$ for $a_0, v_1, \ldots, v_d \in G$ and $N_1, \ldots, N_d \in \mathbb{N}$. Rank d arithmetic progressions generalize the better known rank 1 arithmetic progressions $a_0, a_0 + v_1, a_0 + 2 \cdot v_1, \ldots, a_0 + (N_1 - 1) \cdot v_1$, i.e. number sequences of equal distance, by permitting more than one distance.

Theorem 5.1.1 (Freiman [19]). *Let G be a torsion-free abelian group. For every constant $C > 0$ there are $K, d \in \mathbb{N}$ such that for all subsets $A \subseteq G$ with $|A + A| \leq C|A|$ there is a generalized arithmetic progression P of rank d with $|P| \leq |A|K$.*

The finite automata properties on which Tsankov's proof for the non-automaticity of $(\mathbb{Q}, +)$ relies are summarized in the following lemma.

Lemma 5.1.3 (Tsankov [21]). *Let $G = (D, +)$ be an automatic presentation of an abelian torsion-free group. Let $D^{\leq n} := \{w \in D : |w| \leq n\}$. There exist constants $l_0, r \in \mathbb{N}$ so that the sequence of sets $A_n := D^{\leq l_0+nr}$ has the following properties:*
There exists a constant C_1 and a function $h \colon \mathbb{N} \to \mathbb{N}$ such that:

1. $0 \in A_0 \wedge |A_0| \geq 2$
2. $A_n + A_n \subseteq A_{n+1}$
3. $|A_{n+1}| \leq C_1|A_n|$
4. $\frac{1}{p}A_n \subseteq A_{n+h(p)}$

Tsankov proceeds then by showing with the help of Freiman's theorem, that no such subset sequence exists in $(\mathbb{Q}, +)$. The finite automata property that enable those properties and which fail for regular languages with advice is the bounded growth of the word lengths under functions. Property 2. is a consequence of the fact that for any automatic presentation $(D, +)$ of a group, there is a constant C such that $D^{\leq n} + D^{\leq n} \subseteq D^{\leq n+C}$ for all n. This doesn't hold true anymore for the automatic presentation of the rationals with the parameter bin(2)#bin(3)#bin(4)#bin(5)$\ldots$ = 01#11#001#101#$\ldots$. When two numbers in factorial representation are added and a carry is generated their coefficient list can grow only by one new coefficient 1 in the last position position n, but the new coefficient must be padded to the length of the whole cell bin(n + 2)# for the automaton to be able to calculate the modulo $n + 2$-sum. So that growth of the word lengths is unbounded.

5.2 Equal Ends and Pairing Functions

In [1] a different growth argument was developed which can be used to prove that a structure doesn't have an ω-automatic presentation and which still works with parameters. Two ω-words $\alpha, \beta \in \Sigma^\omega$ are said to have *equal ends* if there is a position $n \in \omega$ from which on the words are equal. Formally one defines for every n the equivalence relation $\sim_e^n$ on all ω-words with $\alpha \sim_e^n \beta :\Leftrightarrow \forall m \geq n\, \alpha(m) = \beta(m)$. Call a set B of words *$\sim_e^n$-equivalent* if it is contained in an $\sim_e^n$-equivalence class. For finite words w, v there is always an n from which on the words are equal which is simply the maximum of their lenghts $|w \otimes v|$. For ω-words one can only find such an n, if α and β are in the same equal ends equivalence class. Let $\alpha \sim_e \beta \colon \Leftrightarrow \exists n \alpha \sim_e^n \beta$ be the equal ends equivalence relation on ω-words. Consider now a function $f \colon A^2 \to A$ whose graph $R_f \subseteq A^2 \times A$ is recognized by a finite advice automaton on ω-words with advice tape. Let $\alpha \sim_e^n \beta$ for an n. Then both words have the same suffix γ starting from position n, i.e. there are finite prefixes v, w, x with $|v| = |w| = |x| = n$, so that $\alpha = v\gamma$ and $\beta = w\gamma$ and $f(v\gamma, w\gamma) = x\gamma_q$. The suffix of the function value $f(v\gamma, w\gamma)$ from position n on depends only on the state q that the automaton attains after it has read the length n prefix $\otimes(v, w, x)$ since for any state q there can be only one suffix γ_q so that $\otimes(\gamma, \gamma, \gamma_q)$ is accepted beginning from position n, because otherwise f would map the same input to two different values $f(v\gamma, w\gamma) = x\gamma_q = x\gamma'_q$ for $\gamma_q \neq \gamma'_q$ and wouldn't be a function. As a consequence the image of a set $B \subseteq A$ of ω-words of the form $B = B'\gamma$ where $B' \subseteq \Sigma^n$ is the union of at most $r := |Q|$ different $\sim_e^n$-classes identified by their suffixes from position n on, i.e. $f(B'\gamma, B'\gamma) = B'_{q_0}\gamma_{q_0} \cup B'_{q_1}\gamma_{q_1} \ldots \cup B'_{q_{r-1}}\gamma_{q_{r-1}}$ for sets $B_{q_i} \subseteq \Sigma^n$ and different ω-suffixes γ_{q_i}. If $B'\gamma$ is the largest $\sim_e^n$ equivalence class then we can furthermore bound the size of its image under f by $|f(B'\gamma, B'\gamma)| \leq |\bigcup_{i=0}^{r-1} B'_{q_i}\gamma_{q_i}| \leq \sum_{i=0}^{r-1} |B'_{q_i}\gamma_{q_i}| \leq r|B'\gamma| = r|B'|$. In particular the bound is independent of the function's arity.

Lemma 5.2.1.

For every $\sim_e^n$-equivalent $B \subseteq \Sigma^\omega$ there exist unique $\gamma_B \in \Sigma^\omega$ and $B' \subseteq \Sigma^n$ with

$$B = B'\gamma_B$$

Proof. Let $B' := \{w \in \Sigma^* : |w| = n \wedge \exists \alpha \in B\, w \leq_p \alpha\}$ be the set of length n prefixes of B and $\gamma_B \in \Sigma^\omega$ the n-suffix that all elements in B have in common, which exists, because B is $\sim_e^n$-equivalent. □

Lemma 5.2.2 ([2]). *Let $A \subseteq \Sigma^\omega$ and $f\colon A^k \to A$ be a ω-regular function with advice $\alpha \in \Gamma^\omega$. Then there exists a constant $r \in \mathbb{N}^{>0}$ so that for all $n \in \mathbb{N}$: If $B_1, \ldots, B_k \subseteq A$ are each $\sim_e^n$-equivalent, then $f(B_1, \ldots, B_k)$ is the union of at most r sets $C_0, \ldots, C_{r-1} \subseteq A$ that are each $\sim_e^n$-equivalent, i.e.*

$$f(B_1, \ldots, B_k) = \bigcup_{i=0}^{r-1} C_i$$

Proof. Since each B_i is $\sim_e^n$-equivalent there is a k-tuple of Σ^ω-words $\overline{\gamma} = \gamma_1, \ldots, \gamma_k$ and $B'_i \subseteq \Sigma^n$ with $B_i = B'_i\gamma_i$ for $i = 1, \ldots, k$. Let $\mathcal{A} = (Q, \Sigma^k, \Gamma, \delta, q_0, \mathcal{F})$ be a finite advice automaton that recognizes the graph R_f of f with advice α, i.e. $L^\omega(\mathcal{A}[\alpha]) = \otimes R_f$. The constant we are looking for is $r := |Q|$. Towards a contradiction assume there were $r+1$ tuples $\overline{v_i}\overline{\gamma_i} \in B_1\gamma_1 \times \ldots \times B_k\gamma_k$ with pairwise non-$\sim_e^n$-equivalent function values $f(\overline{v}_i\overline{\gamma}_i) = w_i\beta_i$ and $\beta_i \neq \beta_j$ for $i \neq j$. According to the pigeon hole prinicple and since the automaton has only r states there must be $i < j$, so that $\delta^*(q_0, \otimes(\overline{v}_i, w_i)) = \delta^*(q_0, \otimes(\overline{v}_j, w_j))$. But then $f(\overline{v_i}\overline{\gamma_i}) = w_i\beta_j$, since $\delta^*(q_0, \otimes(\overline{v_i}\overline{\gamma_i}, w_i\beta_j)) = \delta_n^*(\delta^*(q_0, \otimes(\overline{v}_i, w_i)), \otimes(\overline{\gamma}_i, \beta_j)) = \delta_n^*(\delta^*(q_0, \otimes(\overline{v}_i, w_j)), \otimes(\overline{\gamma}_i, \beta_j)) = \delta^*(q_0, \otimes(\overline{v_i}\overline{\gamma_i}, w_j\beta_j)) \in \mathcal{F}$. So $f(\overline{v_i}\overline{\gamma_i}) = w_i\beta_i = w_i\beta_j \Rightarrow \beta_i = \beta_j$. Contradiction! □

As an application of the lemma the minimum image size of a ω-regular function can be bounded. This was proven by Zaid in [1].

Definition 5.2.1. Let $f\colon A^k \to A$ be a function over an infinite set A. The *minimum image size* of f is defined as

$$\mathrm{MIS}_f(n) := \min\{|f(X^k)| : X \subseteq A \wedge |X| = n\}$$

Lemma 5.2.3. *Let $\mathfrak{A}$ be an infinite structure with injective ω-automatic presentation with parameter. Then for every FOC-definable function f it holds that* $\mathrm{MIS}_f(n) = O(n)$

Since the original proof relies on the existence of an infinite $\sim_e$-equivalence class in any infinite ω-regular language, we first need to establish that also infinite ω-regular languages with advice contain an infinite $\sim_e$-equivalence class.

Proposition 5.2.1. *Let $L \subseteq \Sigma^\omega$ be an infinite ω-regular language with advice $\alpha \in \Gamma^\omega$. Then there exists an infinite $\sim_e$-class in L.*

Proof. Let $\mathcal{A} = (Q, \Sigma, \Gamma, \delta, q_0, \mathcal{F})$ be a finite advice automaton with $L([\alpha]) = L$ and let $q = |Q|$ be the number of states of the automaton. If there are only finitely many $\sim_e$-classes, then one of those must be infinite, since the language L is infinite and the finite union of finite sets would be only finite. Thus assume that there are infinitely many $\sim_e$-classes. For every $\gamma \in L$ let $[\gamma]_{\sim_e^m} := \{\beta \in L : \beta \sim_e^m \gamma\}$ the $\sim_e^m$-equivalence class of γ and $\mathrm{in}_m(\gamma) = |[\gamma]_{\sim_e^m}|$ its size. If there is a $\gamma \in L$ with $\sup\limits_{m\to\infty} \mathrm{in}_m(\gamma) = \infty$, then $[\gamma]_{\sim_e}$ is an infinite $\sim_e$-class. Otherwise $\mathrm{in}_m(\gamma)$ is bounded for every $\gamma \in L$. Let n_γ be the smallest number with $\mathrm{in}_{n_\gamma}(\gamma) = \mathrm{in}_t(\gamma)$ for all $t \geq n_\gamma$. Since there are infinitely many equivalence classes there are also $q+1$ pairwise non-$\sim_e$-equivalent words $\gamma_1, \ldots, \gamma_{q+1} \in L$. Choose a number $K' \geq \max\{n_{\gamma_1}, \ldots, n_{\gamma_{q+1}}\}$. Let $K > K'$ be a number, so that the segments of the γ_i between K' and K are pairwise different. By the pigeonhole principle there are $i < j$, so that $\delta^*(q_0, \gamma_i[0,K]) = \delta^*(q_0, \gamma_j[0,K])$. Then $\delta^*(q_0, \gamma_i[0,K]\gamma_j(K,\infty)) = \delta^*_{K+1}(\delta^*(q_0, \gamma_i[0,K]), \gamma_j(K,\infty))$
$= \delta^*_{K+1}(\delta^*(q_0, \gamma_j[0,K]), \gamma_j(K,\infty)) = \delta^*(q_0, \gamma_j) \in \mathcal{F}$. But then the word $\gamma_i[0,K]\gamma_j(K,\infty)$ is also in L and it is $\sim_e^{K+1}$-equivalent to γ_j, but not $\sim_e^{n_{\gamma_j}}$-equivalent. Contradiction to the choice of n_{γ_j}! □

The proof of Lemma 5.2.3 now works also for automatic presentations with parameters.

Proof of Lemma 5.2.3. If $\mathfrak{A}$ has an injective ω-automatic presentation with a parameter, then every FOC-definable function $f\colon A^k \to A$ is ω-regular over an alphabet Σ with advice. So Lemma 5.2.2 applies to f and there is a constant r with the properties stated in Lemma 5.2.2. Towards a contradiction assume that $\mathrm{MIS}_f(n) \neq O(n)$. Then there must be an n with $\mathrm{MIS}_f(n) > r|\Sigma|n$. According to proposition 5.2.1 there is an infinite $\sim_e$-equivalent set in A and in particular there must be a $\sim_e^m$-equivalent set of size at least n for some $m \geq 1$. Choose the smallest such m and let X be the largest $\sim_e^m$-equivalent set. Then $X = X'\gamma$ for a m-suffix γ and $X = \bigcup_{a\in\Sigma}(X'a^{-1})a\gamma$, where $X'a^{-1} := \{w : wa \in X'\}$. The sets $(X'a^{-1})a\gamma$ are $\sim_e^{m-1}$-equivalent. According to the minimal choice of m, they are therefore smaller than n. Then $|X| \leq \sum_{a\in\Sigma}|(X'a^{-1})a\gamma| < n|\Sigma|$. Lemma 5.2.2 tells us that $f(X^k) = \bigcup_{i=0}^{r-1} C_i$ for some $\sim_e^m$-equivalent sets C_i. Due to the maximality of X it is $|C_i| \leq |X|$ and we get the contradiction $|f(X^k)| \leq \sum_{i=0}^{r-1}|C_i| \leq r|X| < r|\Sigma|n$. □

A direct application of the MIS_f-Lemma are pairing funcions. A pairing function is an injective function $f\colon A^2 \to A$.

Corollary 5.2.1. *No infinite structure with a FOC-definable pairing function admits an injective ω-automatic presentation with parameter.*

Proof. For pairing functions $|f(X,X)| = |X|^2$ and therefore $\text{MIS}_f(n) = n^2 \neq O(n)$. □

The MIS-Lemma can also be used to show that the Rado graph doesn't have an automatic presentation with parameters. This has already been proven by Löding and Colcombet [9] with a different method.

Proposition 5.2.2. *The Rado graph doesn't have an automatic presentation with parameters.*

Proof. Assume towards a contradiction that (V, E) is an injective automatic presentation of the Rado graph so that $V \subseteq \Sigma^*$, $E \subseteq V \times V$ are regular with advice. We define a function $f\colon V^3 \to V$ via $\varphi_f(x,y,u,z)$ in FO with superlinear MIS growth rate. On finite words the length-lexicographical well ordering $<_{llex}$ is defined and regular. The idea is to map $x <_{llex} y <_{llex} u$ to the length-lexicographical smallest node $z \in V$ that is connected to x and y but not connected to any node of the set $[\epsilon, u] \setminus \{x, y\}$, where ϵ is the length-lexicographical least element. Such a node always exists in the Rado graph, bccause it has the extension property:

- For all disjoint, nonempty finite sets $U, W \subseteq V$ there exists a node $v \in V$, so that $\forall u \in U (v,u) \in E \wedge \forall w \in W (v,w) \notin E$

If $X = \{x_1 <_{llex} x_2 <_{llex} \dots <_{llex} x_n\} \subseteq V$ is any set with n Elements, then at least the elements $f(x_i, x_j, x_n)$ must be different for all $i < j$ so that $MIS_f(n) \geq \frac{n(n-1)}{2}$. Let

$$\phi(x,y,u,z) := \forall v (v \leq_{llex} u \wedge Ezv \leftrightarrow v = x \vee v = y)$$

$$\varphi_f(x,y,u,z) := \phi(x,y,u,z) \wedge \forall w (\phi(x,y,u,w) \to z \leq_{llex} w)$$

□

5.3 Sum Augmentations and the VD hierarchy

Another method for proving that a structure does not admit an automatic presentation was discovered by Delhommé [10] and has been successfully applied by himself and others [18] to prove the non-automaticity of certain classes of linear orders and trees. Consider an automatic structure $\mathfrak{A} = (A, R_1, \dots, R_n)$ with regular domain $A \subseteq \Sigma^*$ and regular r_i-ary relations $R_i \subseteq (\Sigma^*)^{r_i}$. Delhommé's criterion can be inferred from the observation that $\mathfrak{A}$ has up to isomorphism only finitely many substructures of the form $\langle wL \rangle_{\mathfrak{A}}$, for a given $L \subseteq \Sigma^*$. I.e. for any L the class $\mathcal{K}_L := \{\langle wL \rangle_{\mathfrak{A}} : w \in \Sigma^*, wL \subseteq A\}$ is finite, when isomorphic structures are identified in $\mathcal{K}_L$. In fact, if q_a is the number of states of a deterministic finite automaton that recognizes A and q_i the number of states of the deterministic automata that recognize R_i for $i = 1, \ldots, n$, then $C := q_a \cdot q_1 \cdot \ldots \cdot q_n$ is an upper bound for $|\mathcal{K}_L|$. This is so, because the isomorphism type $\langle wL \rangle_{\mathfrak{A}}$ is already determined by the tuple of states that the automata for A reaches after it has read the prefix w and the states that the automata for R_i reach after reading the r_i-ary convolution $\otimes \underbrace{(w, \ldots, w)}_{r_i}$ for $i = 1, \ldots, n$. Any word $v \in \Sigma^*$ that reaches the same tuple of states can then be substituted for w without changing the behaviour of the automata on words with prefix w, i.e. the mapping $\phi\colon wL \to vL, wx \mapsto vy$ that switches the prefixes is an isomorphism. A criterion for automatic presentability that derives from this combinatorial property is as follows:

Definition 5.3.1. Let $\mathfrak{A}$ be a τ-structure and $\mathcal{K}$ a class of τ-structures. $\mathfrak{A}$ is called *finite sum augmentation of* $\mathcal{K}$, if there are $\mathfrak{B}_1, \ldots, \mathfrak{B}_r \in \mathcal{K}$ (not necessarily distinct) and a partition $A = A_1 \cup \ldots \cup A_r$, so that the restrictions $\mathfrak{A}\big|_{A_i}$ are isomorphic to $\mathfrak{B}_i$ for $i = 1, \ldots, r$. In this case write

$$\mathfrak{A} = \mathfrak{B}_1 \sqcup \ldots \sqcup \mathfrak{B}_r$$

Lemma 5.3.1 (Delhommé [10]). *Let τ be a finite relational signature and $\mathfrak{A} \in$ AutStr a τ-structure. Then for every $FO(\tau)$-formula $\varphi(x, y_1, \ldots, y_p)$ there exists a finite class $\mathcal{K}_\varphi^{\mathfrak{A}}$ of substructures of $\mathfrak{A}$, so that for every $\bar{b} \in A^p$ the restriction $\mathfrak{A}\big|_{\varphi(A,\bar{b})} := \langle \{a \in A : \mathfrak{A} \models \varphi(a, \bar{b})\} \rangle_{\mathfrak{A}}$ is a finite sum augmentation of $\mathcal{K}_\varphi^{\mathfrak{A}}$. I.e. for every $\bar{b} \in A^p$ there are $\mathfrak{B}_1, \ldots, \mathfrak{B}_r \in \mathcal{K}_\varphi^{\mathfrak{A}}$ with*

$$\mathfrak{A}\big|_{\varphi(A,\bar{b})} = \mathfrak{B}_1 \sqcup \mathfrak{B}_2 \ldots \sqcup \mathfrak{B}_r$$

Proof. By taking an injective automatic presentation, we can assume wlog that $\mathfrak{A} = (A, R_1, \ldots, R_n)$ is itself an automatic structure, i.e. $A \subseteq \Sigma^*$, $R_i \subseteq A^{r_i}$ are regular relations. Then $\varphi(x, \bar{y})$ defines a regular relation over A and there exists a finite automaton $\mathcal{A}_\varphi = (Q, \Sigma_\square^{p+1}, \delta, q_0, F)$ that recognizes $\varphi^{\mathfrak{A}}$. For any state $q \in Q$ let $L_q^\varphi := \{x \in \Sigma^* : \delta^*(q, \otimes(x, \underbrace{\epsilon, \ldots, \epsilon}_{p})) \in F\}$ the set of words that $\mathcal{A}_\varphi$ recognizes when the other p tapes are empty beginning in state q and $\mathcal{K}_q := \{\langle wL_q^\varphi \rangle_{\mathfrak{A}} : w \in \Sigma^* \wedge wL_q^\varphi \subseteq A\}$. As outlined in the introduction the classes $\mathcal{K}_q$ are up to isomorphism only finite. Then $\mathcal{K}_\varphi^{\mathfrak{A}} := \bigcup_{q \in Q} \mathcal{K}_q \cup \{\langle w \rangle_{\mathfrak{A}} : w \in A\}$ is finite, since there are also only finitely many singleton substructures $\langle w \rangle_{\mathfrak{A}}$ of $\mathfrak{A}$ due to the finiteness of the signature τ. For every $\bar{b} \in A^p$ there exists then a finite sum augmentation of $\mathfrak{A}\big|_{\varphi(A,\bar{b})}$ with summands in $\mathcal{K}_\varphi^{\mathfrak{A}}$ given by

$$\mathfrak{A}\big|_{\varphi(A,\bar{b})} = \bigsqcup_{\substack{w \in \Sigma^{<|\bar{b}|}: \\ \delta^*(q, \otimes(w, \bar{b})) \in F}} \langle w \rangle_{\mathfrak{A}} \sqcup \bigsqcup_{\substack{w \in \Sigma^{|\bar{b}|}: \\ \delta^*(q_0, \otimes(w, \bar{b})) = q}} \langle wL_q^\varphi \rangle_{\mathfrak{A}}$$

$\square$

The proof doesn't work anymore when parameters are allowed in the presentation. For the sake of proving that certain classes of structures do not have an automatic presentation with parameters it suffices however to use a modification of the criterion.

Lemma 5.3.2. *Let τ be a finite relational signature and $\mathfrak{A} \in \mathrm{AutStr}[\alpha]$ a τ-structure. Then for every $FO(\tau)$-formula $\varphi(x, y_1, \ldots, y_p)$ there exists family $\left(\mathcal{K}^{\mathfrak{A}}_{\varphi,n}\right)_{n\in\mathbb{N}}$ of substructure classes of $\mathfrak{A}$, so that*

1. *There exists a constant C with $|\mathcal{K}^{\mathfrak{A}}_{\varphi,n}| \leq C$ for all n.*
2. *For all $n < m$ every element of $\mathcal{K}^{\mathfrak{A}}_{\varphi,n}$ is a finite sum augmentation of $\mathcal{K}^{\mathfrak{A}}_{\varphi,m}$.*
3. *For all $\bar{b} \in A^p$ there is an n, so that $\mathfrak{A}\big|_{\varphi(A,\bar{b})}$ is a finite sum augmentation of $\mathcal{K}^{\mathfrak{A}}_{\varphi,n}$.*

Proof. Let $\mathfrak{A} = (A, R_1, \ldots, R_m)$ be an automatic structure with parameter α, so that $A \subseteq \Sigma^*$, $R \subseteq A^{r_i}$ are regular with advice $\alpha \in \Gamma^\omega$. Then $\varphi(x, \overline{y})$ defines a relation over A that is regular with some advice. Let $\mathcal{A}_\varphi = (Q, \Sigma_\square^{p+1}, \Gamma, \delta, q_0, F)$ be a finite advice automaton that recognizes $\varphi^{\mathfrak{A}}$ with advice. For any state $q \in Q$ and any $n \in \mathbb{N}$ let $L^\varphi_{q,n} := \{x \in \Sigma^* : \delta^*_n(q, \otimes(x, \underbrace{\epsilon, \ldots, \epsilon}_{p})) \in F\}$ be the set of words that $\mathcal{A}_\varphi$ recognizes with the other p input tapes being empty, beginning in state q and from position n on the advice tape. Let $\mathcal{K}_{q,n} := \{\langle wL^\varphi_{q,n}\rangle_{\mathfrak{A}} : w \in \Sigma^n \wedge wL^\varphi_{q,n} \subseteq A\}$ for every $q \in Q$ and $\mathcal{K}^{\mathfrak{A}}_{\varphi,n} := \bigcup_{q\in Q} \mathcal{K}_{q,n} \cup \{\langle w\rangle_{\mathfrak{A}} : w \in A\}$ for every $n \in \mathbb{N}$.

First we show that there is a constant C with $|\mathcal{K}^{\mathfrak{A}}_{\varphi,n}| \leq C$ for all $n \in \mathbb{N}$:

Let $\mathcal{A}_A := \left(Q_A, \Sigma, \Gamma, q^A_0, \delta_A, F_A\right)$, $\mathcal{A}_i := \left(Q_i, \Sigma^{r_i}_\square, \Gamma, q^i_0, \delta_i, F_i\right)$ be finite advice automata with $L(\mathcal{A}_A[\alpha]) = A$ and $L(\mathcal{A}_i[\alpha]) = R_i$ for $i = 1, \ldots, m$. For each n define an equivalence relation $\sim_n \subseteq \Sigma^n \times \Sigma^n$ by $w \sim_n v :\Leftrightarrow \delta^*_A(q_0, w) = \delta^*_A(q_0, v)$ and $\delta^*_i(q_0, \otimes\underbrace{(w, \ldots, w)}_{r_i}) = \delta^*_i(q_0, \otimes\underbrace{(v, \ldots, v)}_{r_i})$ for all $i \in \{1, \ldots, m\}$. Then it holds that $w \sim_n v \Rightarrow \langle wL\rangle_{\mathfrak{A}} \cong \langle vL\rangle_{\mathfrak{A}}$ for all $L \subseteq \Sigma^*$ with $wL \subseteq A$. Indeed $wx \mapsto vx$ is an isomorphism, because $wx \in A \Leftrightarrow \delta^*_A(q^A_0, wx) = \delta^*_{n,A}(\delta^*_A(q^A_0, w), x) \in F \Leftrightarrow \delta^*_{n,A}(\delta^*_A(q^A_0, v), x) \in F \Leftrightarrow vx \in A$. And $(wx, \ldots, wx) \in R_i \Leftrightarrow \delta^*_i(q^i_0, \otimes\underbrace{(wx, \ldots, wx)}_{r_i}) = \delta^*_{n,i}(\delta^*_i(q^i_0, \otimes\underbrace{(w, \ldots, w)}_{r_i}), \otimes\underbrace{(x, \ldots, , x)}_{r_i}) \in F \Leftrightarrow \delta^*_{n,i}(\delta^*_i(q^i_0, \otimes\underbrace{(v, \ldots, v)}_{r_i}), \otimes\underbrace{(x, \ldots, , x)}_{r_i}) = \delta^*_i(q^i_0, \otimes\underbrace{(vx, \ldots, vx)}_{r_i}) \in F \Leftrightarrow (vx, \ldots, vx) \in R_i$.

The number of equivalence classes of $\sim_n$ is bounded by $C_1 := |Q_A| \cdot |Q_1| \cdot \ldots \cdot |Q_m|$ for every $n \in \mathbb{N}$. Then $|\mathcal{K}^{\mathfrak{A}}_{\varphi,n}| \leq |\bigcup_{q\in Q} \mathcal{K}_{q,n}| + |\{\langle w\rangle_{\mathfrak{A}} : w \in A\}| \leq |Q| \cdot C_1 + C_2$, where C_2 is the number of isomorphism types of τ-structures with one element. Thus $C := |Q| \cdot C_1 + C_2$ is the constant we've been looking for.

Next we show 2: Let $n < m = n + k$ and $\mathfrak{B} \in \mathcal{K}^{\mathfrak{A}}_{\varphi,n}$. Any finite structure is a finite sum augmentation of singletons, so let $\mathfrak{B} = \langle wL^\varphi_{q,n}\rangle_{\mathfrak{A}}$ with $wL^\varphi_{q,n} \subseteq A$ be infinite. Then $wL^\varphi_{q,n} = \bigcup\{wvL^\varphi_{p,n+k} : v \in \Sigma^k \wedge \exists x \in L^\varphi_{q,n}(v \leq_p x \wedge \delta^*_n(q, \otimes(v, \epsilon, \ldots, \epsilon)) = p)\} \cup \{wv : v \in L^\varphi_{q,n} \wedge |v| < k\}$ is a finite sum augmentation of $\mathcal{K}^{\mathfrak{A}}_{\varphi,n+k}$.

Finally we show 3: Let $\bar{b} \in A^p$ with $|\bar{b}| = n$. Then $\mathfrak{A}\big|_{\varphi(A,\bar{b})}$ is a finite sum augmentation of $\mathcal{K}^{\mathfrak{A}}_{\varphi,n}$:

$$\mathfrak{A}\big|_{\varphi(A,\bar{b})} = \bigsqcup_{\substack{w\in\Sigma^{<n}:\\ \delta^*(q,\otimes(w,\bar{b}))\in F}} \langle w\rangle_{\mathfrak{A}} \sqcup \bigsqcup_{\substack{w\in\Sigma^{n}:\\ \delta^*(q_0,\otimes(w,\bar{b}))=q}} \langle wL^{\varphi}_{q,n}\rangle_{\mathfrak{A}}$$

□

From this lemma we can conclude the same criterion that Delhommé used to prove that ω^ω is not automatic presentable. The criterion puts a bound on the number of *indecomposable* substructures that are definable in $\mathfrak{A} \in$ AutStr$[\alpha]$ by a parametrised formula..

Definition 5.3.2. Let $\mathcal{K}$ be a class of τ-structures. $\mathfrak{A} \in \mathcal{K}$ is called indecomposable in $\mathcal{K}$ iff for any finite sum augmentation $\mathfrak{A} = \mathfrak{B}_1 \sqcup \ldots \sqcup \mathfrak{B}_r$ with $\mathfrak{B}_i \in \mathcal{K}$ for $i = 1, \ldots, r$ there exists i with $\mathfrak{A} \cong \mathfrak{B}_i$.

Theorem 5.3.1. *Let τ be a finite relational signature and $\mathfrak{A}$ a τ-structure. If $\mathfrak{A} \in$ AutStr$[\alpha]$ for a parameter $\alpha \in \Gamma^\omega$ and $\varphi(x, \overline{y})$ is a FO(τ)-formula, then the class $\mathcal{K}^{\mathfrak{A}}_{\varphi} := \{\mathfrak{A}\big|_{\varphi^{\mathfrak{A}}(A,\bar{b})} : \bar{b} \in A^p\}$ contains only finitely many structures that are indecomposable in $\mathcal{K}^{\mathfrak{A}}_{\varphi}$.*

Proof. Towards a contradiction assume that $\mathcal{K}^{\mathfrak{A}}_{\varphi}$ contains infinitely many indecomposable structures. Let $\left(\mathcal{K}^{\mathfrak{A}}_{\varphi,n}\right)_{n\in\mathbb{N}}$ be the family of lemma 5.3.2. Let C be the constant with $|\mathcal{K}^{\mathfrak{A}}_{\varphi,n}| \leq C$ for all $n \in \mathbb{N}$. Take $C+1$ pairwise non-isomorphic, indecomposable structures $\mathfrak{B}_1, \ldots, \mathfrak{B}_{C+1} \in \mathcal{K}^{\mathfrak{A}}_{\varphi}$. According to Lemma 5.3.2 there are then $n_1, \ldots, n_{C+1} \in \mathbb{N}$ with $\mathfrak{B}_i \in \mathcal{K}_{\varphi,n_i}$ for $i = 1, \ldots, C+1$. Let $m := \max_{i=1}^{C+1} m_i$. Then every $\mathfrak{B}_i$ is a finite sum augmentation of $\mathcal{K}_{\varphi,m}$ and since $\mathfrak{B}_i$ is indecomposable it is therefore an element of $\mathcal{K}_{\varphi,m}$ for $i = 1, \ldots, C+1$ which is a contradiction to $|\mathcal{K}_{\varphi,m}| \leq C$. □

Theorem 5.3.1 can now be applied to generalize some of the results that have already been known for automatic structures without parameters. We begin by showing that ω^ω is the smallest well-ordering that is not automatic presentable with any parameter:

Theorem 5.3.2. *The well-ordering ω^ω is not automatic presentable with any parameter.*

Proof. Assume $\mathfrak{A} := (\omega^\omega, <) \in$ Autstr$[\alpha]$ for some parameter α. Let $\varphi(x, y) := x < y$, then $\mathcal{K}^{\mathfrak{A}}_{\varphi}$ consists of all ordinals below ω^ω. In particular $\omega^i \in \mathcal{K}^{\mathfrak{A}}_{\varphi}$ for all $i > 0$. Thanks to theorem 5.3.1 it suffices to show that the ordinals ω^i are indecomposable in $\mathcal{K}^{\mathfrak{A}}_{\varphi}$, so that $\mathcal{K}^{\mathfrak{A}}_{\varphi}$ contains infinitely many indecomposable elements. Since any subset of an ordinal is itself an ordinal, an ordinal is a finite sum augmentation of ordinals if and only if it is the disjoint union of finitely many ordinals. The disjoint union $\omega^i = \beta_1 \cup \ldots \cup \beta_r$ is equivalent to a colouring $c\colon \omega^i \to \{1, \ldots, r\}$ with finitely many colours. The proposition that ω^i is indecomposable in $\mathcal{K}^{\mathfrak{A}}_{\varphi}$ is thus equivalent to the claim that for any finite colouring $c\colon \omega^i \to D$ there exists a colour $d \in D$ so that $c^{-1}(d) \cong \omega^i$.

We prove this claim by induction over i. For $i = 0$, i.e. $\omega^0 = 1$ this is clear. Consider now a finite colouring $c\colon \omega^{i+1} \to D$. ω^{i+1} can be identified with the well-ordering $(\omega^i \times \omega, <)$ with $(x, y) < (x', y') :\Leftrightarrow y < y' \vee (y = y' \wedge x < x')$. So that $\omega^{i+1} = \bigcup_{j<\omega} \omega^i \times \{j\}$ and $c\big|_{\omega^i \times \{j\}}$ is a finite colouring of ω^i for every j. According to induction hypothesis there is therefore for every j a colour $d(j) \in D$ with $\omega^i \times \{j\} \cap c^{-1}(d(j)) \cong \omega^i$. Since D is finite there must then be one colour $d \in D$ for which there are infinitly many $j_1 < j_2 < j_3 < \ldots$ with $d = d(j_1) = d(j_2) = \ldots$. It follows that $c^{-1}(d) \cong \bigcup_{k<\omega} \omega \times \{i_k\} \cong \omega^{i+1}$. □

The ordinals below ω^ω all have automatic presentations even without parameters. An automatic presentation of ω^n over alphabet $\Sigma = \{0, \ldots, (n-1)\}$ for example is given by $(A, <_{lex})$ with $A := (n-1)^* \ldots 0^*$ with the lexicographic ordering. Another consequence is that there is no parametrised FO-interpretation of ω^ω in any of the structures $\bar{l}$-$\mathfrak{W}(\Sigma^*)$ or $\bar{l}$-$\mathfrak{W}(\mathbb{N})$. There is however a FO-interpretation of ω^ω in $\mathfrak{W}(\Sigma^*)$:

Corollary 5.3.1. $\mathfrak{W}(\Sigma^*)$ *is not automatic presentable with any parameter.*

Proof. We show that $\omega^\omega \leq_{FO} \mathfrak{W}(\Sigma^*)$. Since ω^ω is not automatic presentable with any parameter and FO-interpretations preserve automatic presentability with parameters then $\mathfrak{W}(\Sigma^*)$ can't be automatic presentable with any parameter either. Let $a \in \Sigma$.

$$\Delta(x) := \forall i(|i| \leq |x| \to x(i) \in a^*) \wedge (|x| = 0 \vee \neg|x(x)| = 0)$$

Ordinals $\alpha \in \omega^\omega$ are represented by the coefficient list of their Cantor normal form $\alpha = \omega^n k_n + \ldots + \omega k_1 + k_0$ with $k_i \in \omega$ for all $0 \leq i \leq n$ and $k_n \neq 0 \mapsto (a^{k_0}, a^{k_1}, \ldots, a^{k_n})$. The empty ordinal $0 = \emptyset$ is represented by the empty sequence $()$.

$$\begin{aligned}\psi_<(x,y) := \quad & |x| < |y| \vee \\ & (|x| = |y| \wedge \exists |j| \leq |x|(|x(j)| < |y(j)| \wedge \forall i(|j| < |i| \leq |x| \to |x(i)| = |y(i)|)))\end{aligned}$$

The coefficient lists are ordered length-lexicographically from right to left. □

In his Phd thesis [18] about automatic structures Rubin could generalize the criterion by Delhommé and use it to prove that more generally any linear ordering that is automatic presentable can only have a finite VD-rank. Before we give a definition of the VD-rank of a linear ordering, we introduce the notion of a rank function on structures and see how it relates to theorem 5.3.1.

Definition 5.3.3. Let τ be a signature and $\mathcal{K}$ a class of τ-structures. A rank function $r\colon \mathcal{K} \to On$ on $\mathcal{K}$ assigns to each structure $\mathfrak{A} \in \mathcal{K}$ an ordinal number $r(K) \in On$. Since $\mathcal{K}$ is closed under isomorphism, isomorphic structures get the same rank, i.e.

$$r(\mathfrak{A}) \neq r(\mathfrak{B}) \Rightarrow \mathfrak{A} \not\cong \mathfrak{B}$$

With rank functions one can generalize the concept of indecomposability under finite sum augmentations:

Definition 5.3.4. Let $\mathcal{K}$ be a class of τ-structures and r a rank function on $\mathcal{K}$. r is called *indecomposable under sum augmentations of* $\mathcal{K}$ iff for any finite sum augmentation $\mathfrak{A} = \mathfrak{B}_1 \sqcup \ldots \sqcup \mathfrak{B}_r$ of $\mathcal{K}$ there exists i with $r(\mathfrak{A}) = r(\mathfrak{B}_i)$.

The trivial rank function that assigns to every structure the same rank 0 would be an example for an indecomposable rank function but is for the following not very useful.

Theorem 5.3.3. *Let τ be a finite relational signature and $\mathfrak{A}$ a τ-structure. If $\mathfrak{A} \in$ $\textsc{AutStr}[\alpha]$ for a parameter $\alpha \in \Gamma^\omega$ and $\varphi(x, \overline{y})$ is a FO(τ)-formula, then $r(\mathcal{K}^{\mathfrak{A}}_\varphi) :=$ $\{r(\mathfrak{B}) : \mathfrak{B} \in \mathcal{K}^{\mathfrak{A}}_\varphi\}$ is a finite set for every rank function that is indecomposable under sum augmentations of the substructures of $\mathfrak{A}$.*

Proof. As in the proof of theorem 5.3.1 assume that $r(\mathcal{K}^{\mathfrak{A}}_\varphi)$ is infinite and take $C + 1$ structures $\mathfrak{B}_i \in \mathcal{K}^{\mathfrak{A}}_\varphi$ with different ranks. Again, the $\mathfrak{B}_i$ are finite sum augmentations of $\mathcal{K}^{\mathfrak{A}}_{\varphi,n_i}$. Since r is indecomposable there are structures $\mathfrak{B}'_i \in \mathcal{K}^{\mathfrak{A}}_{\varphi,n_i}$ with $r(\mathfrak{B}_i) =$ $r(\mathfrak{B}'_i)$. Choose an $m \geq n_i$, so that all $\mathfrak{B}'_i$ are sum augmentations of $\mathcal{K}^{\mathfrak{A}}_{\varphi,m}$. Using indecomposability of r again $\mathcal{K}_{\varphi,m}$ contains $C + 1$ structures with different ranks which are therefore non-isomorphic which contradicts $|\mathcal{K}^{\mathfrak{A}}_{\varphi,m}| \leq C$. □

From this one can distill a general approach for proving that a structure $\mathfrak{A}$ doesn't have an automatic presentation with parameters:

- Find a rank function on the substructures of $\mathfrak{A}$ that is indecomposable under finite sum augmentations but seperates enough non-isomorphic structures.
- Find a formula $\varphi(x, y_1, \ldots, y_p)$ or more generally a relation on $\mathfrak{A}$ that is *intrinsic regular*, i.e. regular under any automatic presentation with parameters so that the class $\mathcal{K}^{\mathfrak{A}}_\varphi$ contains infinitely many structures with pairwise different ranks.

We apply the approach to countable linear orders based on Rubin's idea [18]. Here are first some basic notions on linear orders. A comprehensive reference where basic definitions and many results about linear orders are presented is the book by Rosenstein [17].

Let ω denote the order type of the naturals, ω^* the order type of the negative integers, η the order type of the rationals,ζ the order type of the integers and $\mathbf{n}$ the order type of the linear order on n elements. A linear order is *scattered* if it doesn't embed η. A theorem by Hausdorff says that every countable linear order is the dense sum of countably many scattered linear orders, so that it is enough for the moment to focus on countable scattered linear orders. A rank function on countable scattered linear orders is the VD("very discrete")-rank which is defined by transfinite recursion over the countable ordinals ω_1:

- $VD_0 := \{\mathbf{0}, \mathbf{1}\}$
- $VD_\alpha := \left\{ \sum_{i \in \mathcal{I}} \mathcal{L}_i : \mathcal{I} \in \{\omega, \omega^*, \zeta\}, \mathcal{L}_i \in \bigcup_{\beta < \alpha} VD_\beta \right\}$

It can be shown that the class of countable scattered linear orders is already exhausted by the union of the VD_α, i.e. $VD := \bigcup_{\alpha\in\omega_1} VD_\alpha$. Thus we can assign to every countable scattered linear order $\mathcal{L}$ the smallest ordinal α with $\mathcal{L} \in VD_\alpha$ as a rank. Denote this rank function also by $VD(\mathcal{L}) := \inf\{\alpha \in \omega_1 : \mathcal{L} \in VD_\alpha\}$.

Example 5.3.1.
- $VD(\mathbf{0}) = VD(\mathbf{1}) = 0$.
- $VD(\mathbf{n}) = VD(\underbrace{1+1+\ldots+1}_{n}+0+\ldots) = 1$ for $\mathbf{n} > 1$, $VD(\omega) = VD(1+1+\ldots) = 1$, $VD(\omega^*) = VD(\ldots+1+1) = 1$, $VD(\zeta) = VD(\ldots+1+1+\ldots) = 1$.
- $VD(\omega\mathbf{n}) = VD(\underbrace{\omega+\omega+\ldots+\omega}_{n}+0+0+\ldots) = 2$, $VD(\omega^*\mathbf{n}) = 2$, $VD(\zeta\mathbf{n}) = 2$, $VD(\zeta+\mathbf{n}_1+\zeta+\mathbf{n}_2+\ldots) = 2$
- $VD((\zeta\omega)\omega^*) = 3$
- etc.

The VD-rank is however not indecomposable under finite sum augmentations, since $VD(\omega+2) = VD((1+1+\ldots)+1+1+0+\ldots) = 2$, but $VD(\omega) = VD(2) = 1$. One defines therefore another rank function on the scattered linear orders, which differs not too much from VD, but has the advantage of being indecomposable under finite sum augmentations. The VD_*-rank $VD_*(\mathcal{L})$ of a countable scattered linear order $\mathcal{L}$ is defined as the smallest ordinal α so that $\mathcal{L}$ is a finite sum of scattered linear orders with VD-rank less or equal α, i.e. so that $\mathcal{L} = \sum_{i=1}^{k} \mathcal{L}_i$ with $VD(\mathcal{L}_i) \leq \alpha$. Thus $VD_*(\omega+1+1) = \max\{VD(\omega), VD(1)\} = 1$. Furthermore it can be shown that $VD_*(\mathcal{L}) \leq VD(\mathcal{L}) \leq VD_*(\mathcal{L}) + 1$. The following facts are direct consequences of propositions proved by Rubin and are here reformulated within our general framework.

Proposition 5.3.1.
1. *VD_* is indecomposable under finite sum augmentations. ([18] Proposition E.2.4)*
2. *Let $\varphi(x, y_1, y_2) := y_1 \leq x \wedge x \leq y_2$ then $VD_*(\mathcal{K}_\varphi^{\mathcal{L}})$ is infinite for every countable scattered linear order $\mathcal{L}$ of infinite VD_* rank. ([18] Proposition E.2.2 3. + Proposition E.2.5)*

From proposition 5.3.1 and theorem 5.3.3 we get that scattered linear orders with infinite VD_* (equivalently infinite VD) rank have no automatic presentation with parameters. As Rubin shows for the case of automatic structures without parameters the result can be further strengthened to the class of all linear orders. The stronger result can also be generalized to automatic structures with parameters. One defines another rank function which extends the VD rank function from the class of countable scattered linear orders to the class of all linear orders. For a linear order $\mathcal{L}$ consider the equivalence relation $x \sim_{FC} y :\Leftrightarrow [\min(x,y), \max(x,y)]$ is a finite interval of $\mathcal{L}$. Then the equivalence classes are (not necessarily finite) intervals of $\mathcal{L}$. Let $c_{FC}\colon \mathcal{L} \to c_{FC}[\mathcal{L}]$ be the canonical projection map, that maps each x to its equivalence class $c_{FC}(x) = [x]_{FC}$ which is an interval. The set $c_{FC}[\mathcal{L}]$ of equivalence classes can again be equipped with a linear order via $[x]_{FC} < [y]_{FC} :\Leftrightarrow x' < y'$ for all $x' \in [x]_{FC}$ and $y' \in [y]_{FC}$. c_{FC} is also called "finite condensation" map, because it condenses elements that have finite distance to each other into one interval. This condensation procedure can now be repeated on the condensed linear order $c_{FC}[\mathcal{L}]$, so that all intervals within a finite distance from each other get again condensed into one interval and so on. By transfinite recursion over the ordinal numbers one thereby defines the α-th condensation $c^{\alpha}_{FC}[\mathcal{L}]$ of $\mathcal{L}$ for all $\alpha \in On$ in the usual way. The first ordinal α with $c^{\alpha+1}_{FC}[\mathcal{L}] = c^{\alpha}_{FC}[\mathcal{L}]$, i.e. the closure ordinal in the fixed point iteration of the operator $\mathcal{L} \mapsto c_{FC}[\mathcal{L}]$ is the *FC-rank of* $\mathcal{L}$ denoted by $FC(\mathcal{L})$. It can be shown that FC rank and VD rank coincide on countable scattered linear orders.

Example 5.3.2.

- $FC(\eta) = 0$, because in a dense linear order all elements $x \neq y$ have infinite distance from each other and so are not equivalent.
- $FC(\mathcal{L}_i) \leq FC(\sum_{i\in\eta} \mathcal{L}_i)$ for all $i \in \eta$ and therefore $FC(\sum_{i\in\eta} \mathcal{L}_i) \geq \sup\{FC(\mathcal{L}_i) : i \in \eta\}$.

The following is an adaptation of the corresponding proof for automatic structures [18].

Theorem 5.3.4. *Countable linear orders with infinite FC-rank have no automatic presentation with parameters.*

Proof. Towards a contradiction assume there is a countable linear order $\mathcal{L} \in \textsc{AutStr}[\alpha]$ that has infinite FC rank. $\mathcal{L}$ is according to Hausdorff's theorem a dense sum of scattered linear orders $\mathcal{L} = \sum_{i\in\eta} \mathcal{L}_i$. Since $FC(\mathcal{L}) \geq \sup\{FC(\mathcal{L}_i) : i \in \eta\} = \sup\{VD(\mathcal{L}_i) : i \in \eta\} \geq \omega$ for any $C \in \omega$ there is a scattered linear order with $VD_*(\mathcal{L}_i) + 1 \geq VD(\mathcal{L}_i) > C$. Furthermore it can be shown that $\sup\{VD([b_1, b_2]) : [b_1, b_2] \subseteq \mathcal{L}_i\} = VD(\mathcal{L}_i)$. So for every C there is an interval $[b'_1, b'_2] \subseteq \mathcal{L}$ with $VD_*([b'_1, b'_2]) > C$. For $\varphi(x, y_1, y_2) := y_1 \leq x \wedge x \leq y_2$ it holds therefore that $VD_*(\mathcal{K}^{\mathcal{L}}_{\varphi})$ is infinite. Contradiction to theorem 5.3.3. □

In the same publication Rubin could also prove that the Cantor Bendixson rank of any automatic presentable tree must be finite. Since the proof for this theorem doesn't involve any new automaton techniques anymore but associates the Cantor Bendixson rank of a tree with the VD rank of its associated Kleene Brouwer ordering and then essentially reduces the proof to the finiteness condition for the VD rank of automatic linear orders, we can state here without further elaboration that it still holds when parameters are allowed in the presentation.

Theorem 5.3.5. *Countable trees with infinite Cantor Bendixson rank have no automatic presentation with parameters.*

Having narrowed down the domain of linear orders that are potential candidates for automatic presentations to those with finite FC rank, we now give some positive examples. Let us introduce the notations $\textsc{AutLo}$,$\textsc{AutLo}[\alpha]$ and $\textsc{AutLo}[\text{all}]$ for the classes of linear orders that are automatic, automatic with parameter α and automatic with some parameter respectively. Here are first a few basic closure properties of these classes.

Lemma 5.3.3. *1.* $\mathcal{L}_1 \in \textsc{AutLo}[\alpha]$, $\mathcal{L}_2 \in \textsc{AutLo}[\beta] \Rightarrow \mathcal{L}_1 + \mathcal{L}_2 \in \textsc{AutLo}[\alpha \otimes \beta]$

2. $\mathcal{L}_1 \in \textsc{AutLo}[\alpha] \Rightarrow \mathcal{L}_1^{\mathrm{R}} \in \textsc{AutLo}[\alpha]$

3. $\mathcal{L}_1 \in \textsc{AutLo}[\alpha], \mathcal{L}_2 \in \textsc{AutLo}[\beta] \Rightarrow \mathcal{L}_1\mathcal{L}_2 \in \textsc{AutLo}[\alpha \otimes \beta]$

4. $\omega, \omega^*, \eta, \zeta, \boldsymbol{n} \in \textsc{AutLo}$

Proof. 1. Let $(\Delta_1, L_{<_1})$ and $(\Delta_2, L_{<_2})$ be injective automatic presentations, so that $\Delta_1 \subseteq \Sigma_1^*$, $L_{<_1} \subseteq \Sigma_1^* \times \Sigma_1^*$ are regular with advice $\alpha \in \Gamma_1^\omega$ and $\Delta_2 \subseteq \Sigma_2^*$,$L_{<_2} \subseteq \Sigma_2^* \times \Sigma_2^*$ regular with advice $\beta \in \Gamma_2^\omega$. We can assume wlog that Σ_1 and Σ_2 are disjoint. Let $\Sigma := \Sigma_1 \cup \Sigma_2$. Then all those relations are also regular with advice $\alpha \otimes \beta$ over Σ. Let $\Delta := \Delta_1 \cup \Delta_2$ and $L_< := L_{<_1} \cup L_{<_2} \cup \Delta_1 \times \Delta_2$ which are also regular with advice $\alpha \otimes \beta$ due to boolean closure properties. $(\Delta, L_<)$ is a presentation of $L_1 + L_2$.

2. This is another consequence of the closure properties: If $L \subseteq \Sigma^* \times \Sigma^*$ is regular with advice α, then so is L^R with $(x, y) \in L^R :\Leftrightarrow (y, x) \in L$.

3. Given $(\Delta_1, L_{<_1})$ and $(\Delta_2, L_{<_2})$ as in 1. Let $\Delta := \Delta_1 \times \Delta_2$ and $((v_1, v_2), (w_1, w_2)) \in L_< :\Leftrightarrow v_2 <_2 w_2 \vee (v_2 = w_2 \wedge v_1 < w_1)$ which is regular with advice $\alpha \otimes \beta$.

4. $\omega \in \textsc{AutLo}$ was already shown earlier in this chapter. $\omega^* = \omega^R \in \textsc{AutLo}$ according to 2. $\zeta = \omega^* + \omega \in \textsc{AutLo}$ according to 1. $\eta \cong (\{0,1\}^*1, <_{lex}) \in \textsc{AutLo}$ and $\mathbf{n}$ is finite and therefore trivially automatic presentable.

□

A more interesting closure property concerns coloured linear orders. A *coloured linear order* $\mathcal{L} := (L, <, (P_a)_{a \in \Sigma})$ over Σ is a linear order expanded by finitely many monadic predicates P_a which are pairwise disjoint (possibly empty) so that they partition the domain of the linear order into differently coloured elements. A partition corresponds to a colouring $\mathcal{L}\colon L \to \Sigma$ with $\mathcal{L}(w) = a$ if and only if $w \in P_a^{\mathcal{L}}$. For a linear order type $\mathcal{L}$ denote by $\Sigma^{\mathcal{L}}$ the set of all Σ-coloured $\mathcal{L}$-orders.

Definition 5.3.5. Let $\mathcal{L}$ be a coloured linear order over Ω and $\mathcal{F} = (\mathcal{L}_a)_{a \in \Omega}$ a family of linear orders. The $\mathcal{L}$-sum over $\mathcal{F}$ is the $\bigcup_{a \in \Omega} \Omega_a$-coloured linear order

$$\sum_{\mathcal{L}} \mathcal{F} := \sum_{w \in \mathcal{L}} \mathcal{L}_{\mathcal{L}(w)}$$

Example 5.3.3. 1. Every ω-word $\alpha \in \Gamma^\omega$ corresponds to a Σ-coloured linear order of order type ω and vice versa. One can thereby also generalize the concept of ω-words to $\mathcal{L}$-words for any linear order $\mathcal{L}$.

2. Let $\Sigma_2 = \{0, 1\}$ and $\mathcal{F} := \{\mathcal{L}_0 := \mathbf{0}, \mathcal{L}_1 := \mathbf{1}\}$. For any ω-word $\alpha \in \Sigma_2^\omega$ interpreted as a coloured linear order, the α-sum over $\mathcal{F}$ is

$$\sum_{\alpha} \mathcal{F} = \sum_{i \in \omega} \alpha(i) = \begin{cases} \mathbf{n} & \text{if } \exists^{=n} i : \alpha(i) = 1 \\ \omega & \text{if } \exists^{\infty}\ i : \alpha(i) = 1 \end{cases}$$

Coloured $\mathcal{L}$-sums preserve automaticity:

Proposition 5.3.2. *Let $\mathcal{L}$ be a Ω-coloured linear order and $\mathcal{F} = \{\mathcal{L}_a : a \in \Omega\}$ a family of coloured linear orders where $\mathcal{L}_a$ is Ω_a-coloured for $a \in \Omega$.*

$$\mathcal{L} \in \textsc{AutLo}[\alpha], \forall a \in \Omega : \mathcal{L}_a \in \textsc{AutLo}[\beta_a] \Rightarrow \sum_{\mathcal{L}} \mathcal{F} \in \textsc{AutLo}[\alpha \otimes \otimes_{a\in\Omega}\beta_a]$$

Proof. Let $\mathcal{L} = (\Delta, <, (P_a)_{a\in\Omega})$, $\mathcal{L}_a = (\Delta_a, <_a, (P_{b,a})_{b\in\Omega_a})$ be injective automatic presentations with parameters α and β_a for all $a \in \Omega$ respectively. We can choose a common alphabet over which all relations in the presentation are regular with advice $\alpha \otimes_{a\in\Omega} \beta_a$. Then an automatic presentation of $\sum_{\mathcal{L}} \mathcal{F}$ with parameter $\alpha \otimes_{a\in\Omega} \beta_a$ is given by $\mathfrak{d} = (\Delta_\Sigma, <_\Sigma, (P_b^\Sigma)_{b\in\bigcup\Omega_a})$ with

- $\Delta_\Sigma := \bigcup\limits_{a\in\Omega} P_a \times \Delta_a$
- $(v_1, w_1) <_\Sigma (v_2, w_2) :\Leftrightarrow v_1 < v_2 \vee \bigvee\limits_{a\in\Omega} (v_1 = v_2 \wedge v_1 \in P_a \wedge w_1 <_a w_2)$
- $P_b^\Sigma := \bigcup\limits_{a\in\Omega} P_a \times P_{b,a}$

□

As a consequence we can determine how the automatic scattered linear orders with and without parameters are situated within the VD hierarchy.

Proposition 5.3.3. .

1. $VD_2 \subseteq \textsc{AutLo}[\text{all}]$
2. $VD_2 \not\subseteq \textsc{AutLo}[\alpha]$ *for all* α
3. $VD_1 \subseteq \textsc{AutLo}$
4. $VD_2 \not\subseteq \textsc{AutLo}$

Proof. 1. Recall from example 5.3.1 that $VD_1 = \{\mathbf{0}, \mathbf{1}, \mathbf{n}, \omega, \omega^*, \zeta : n \in \omega\}$. According to the definition of the VD hierarchy therefore

$$VD_2 := \left\{ \sum_{i\in\mathcal{I}} \mathcal{L}_i : \mathcal{I} \in \{\omega, \omega^*, \zeta\}, \mathcal{L}_i \in \{\mathbf{0}, \mathbf{1}, \mathbf{n}, \omega, \omega^*, \zeta : n \in \omega\} \right\}$$

This can be simplified. First of all ω- and ω^*-sums can both also be written as ζ-sums by extending the sum at the ends with infinitely many $\mathbf{0}$-summands, i.e.: $\sum\limits_{\omega} \mathcal{L}_i = \sum\limits_{\omega^*} \mathbf{0} + \sum\limits_{\omega} \mathcal{L}_i = \sum\limits_{\zeta} \mathcal{L}_i'$ with $\mathcal{L}_i' = \begin{cases} \mathcal{L}_i & \text{if } i \in \omega \\ \mathbf{0} & \text{if } i \in \omega^* \end{cases}$ and analogous for ω^*

sums. Furthermore any $\mathbf{n}$-summand can be written as $\mathbf{n} = \underbrace{\mathbf{1} + \mathbf{1} + \ldots + \mathbf{1}}_{n}$ and $\zeta = \omega^* + \omega$, so that

$$VD_2 := \left\{ \sum_{i \in \zeta} \mathcal{L}_i : \mathcal{L}_i \in \{\mathbf{0}, \mathbf{1}, \omega, \omega^*\} \right\}$$

With the alphabet $\Sigma := \{0, 1, \omega, \omega^*\}$ and $\mathcal{F} := \{\mathcal{L}_0 := \mathbf{0}, \mathcal{L}_1 := \mathbf{1}, \mathcal{L}_\omega := \omega, \mathcal{L}_{\omega^*} = \omega^*\}$ the VD_2 orders are therefore precisely the Σ-coloured ζ-sums:

$$VD_2 := \left\{ \sum_{\mu} \mathcal{F} : \mu \in \Sigma^\zeta \right\}$$

According to Lemma 5.3.3 $\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_\omega, \mathcal{L}_{\omega^*} \in \textsc{AutLo}$. Every ω-word $\alpha \in \Sigma^\omega$, i.e. every Σ-coloured linear order of order type ω, is in $\textsc{AutLo}[\alpha]$. Any $\mu \in \Sigma^\zeta$ can be written as $\mu = \alpha^* + \beta$ with $\alpha, \beta \in \Sigma^\omega$, so that using Lemma 5.3.3 again, $\mu \in \textsc{AutLo}[\alpha \otimes \beta]$. By 5.3.2 thus $\sum_{\mu} \mathcal{F} \in \textsc{AutLo}[\text{all}]$ for every $\mu \in \Sigma^\zeta$, i.e. $VD_2 \subseteq \textsc{AutLo}[\text{all}]$

2. This follows from a simple cardinality argument. The class $\textsc{AutLo}[\alpha]$ is only countably infinite for any α, because there are only countably many finite automata and thus only countably many automatic presentations with advice α. The class VD_2 however is uncountable infinite, since for instance the linear orders $\zeta + n_1 + \zeta + n_2 + \ldots$ for any number sequence with $n_i > 0$ are pairwise non-isomorphic.

3. see Lemma 5.3.3

4. special case of 2.

□

The author furthermore believes that $VD_3 \not\subseteq \textsc{AutLo}[\text{all}]$ holds, because while all ωn-words are automatic presentable with n parameters, it seems natural that ω many parameters would be required to represent all ω^2-words and it should be possible to prove it by a diagonalization argument of some sort.

In analogy to the full VD hierarchy one can now also define an automatic sub-hierarchy of VD which starts with VD_0, or any finite set of automatic linear orders. If on level $n+1$ only the ζ-sums over all finite subsets of level n are used, then every linear order of the hierarchy has an automatic presentation with parameter due to proposition 5.3.2.

- $\textsc{Aut}VD_0 := VD_0 = \{0, 1\}$
- $\textsc{Aut}VD_{n+1} := \left\{ \sum_{i \in \zeta} \mathcal{L}_i : \{\mathcal{L}_i : i \in \zeta\} \in \mathcal{P}_{fin}(\textsc{Aut}VD_n) \right\}$

Bibliography

[1] Faried Abu Zaid. Definability in ω-Automatic Structures. Diploma thesis, RWTH-Aachen, 2011.

[2] Faried Abu Zaid, Erich Grädel, and Łukasz Kaiser. The Field of Reals is not omega-Automatic. In Christoph Dürr and Thomas Wilke, editors, *Proceedings of the 29th International Symposium on Theoretical Aspects of Computer Science, STACS 2012*, 2012.

[3] Jean-Paul Allouche and Jeffrey O. Shallit. *Automatic Sequences - Theory, Applications, Generalizations.* Cambridge University Press, 2003.

[4] Reinhold Baer. Abelian groups without elements of finite order. *Duke Mathematical Journal*, 3:68–122, 1937.

[5] Vince Bárány. *Automatic Presentations of Infinite Structures.* Dissertation, RWTH Aachen, 2007.

[6] Achim Blumensath. Automatic Structures. Diploma thesis, RWTH-Aachen, 1999.

[7] Véronique Bruyère, Georges Hansel, Christian Michaux, and Roger Villemaire. Logic and p-recognizable sets of integers. *Bull. Belg. Math. Soc*, 1:191–238, 1994.

[8] Julius R. Büchi. On a decision method in restricted second order arithmetic. In Ernest Nagel, Patrick Suppes, and Alfred Tarski, editors, *Proceedings of the 1960 International Congress on Logic, Methodology and Philosophy of Science (LMPS'60)*, pages 1–11. Stanford University Press, June 1962.

[9] Thomas Colcombet, Christof Löding, Issn Aachener, and Informatik Berichte Aib. Transforming structures by set interpretations, 2006.

[10] Christian Delhommé. Automaticité des ordinaux et des graphes homogènes. *C. R. Math. Acad. Sci. Paris*, 339:5–10, 2004.

[11] László Fuchs. *Infinite Abelian Groups*, volume Bd. 1. Academic Press, New York/London, 1970.

[12] John E. Hopcroft, Rajeev Motwani, and Jeffrey D. Ullman. *Introduction to Automata Theory, Languages and Computation.* Pearson Addison-Wesley, Upper Saddle River, NJ, 3. edition, 2007.

[13] Lukasz Kaiser. *Logic and Games on Automatic Structures - Playing with Quantifiers and Decompositions*, volume 6810 of *Lecture Notes in Computer Science*. Springer, 2011.

[14] Bakhadyr Khoussainov and Anil Nerode. Automatic presentations of structures. In *LCC*, pages 367–392, 1994.

[15] Alex Kruckman, Sasha Rubin, John Sheridan, and Ben Zax. A myhill-nerode theorem for automata with advice. In *GandALF'12*, pages 238–246, 2012.

[16] Dominique Perrin and Jean-Éric Pin. *Infinite words : automata, semigroups, logic and games*. Pure and applied mathematics. London San Diego, Calif. Academic, 2004.

[17] Joseph G. Rosenstein. *Linear orderings*. Academic Press, New York, 1982.

[18] Sascha Rubin. *Automatic Structures*. Phd thesis, University of Auckland, 2004.

[19] Terence Tao and Van H. Vu. *Additive combinatorics*. Cambridge studies in advanced mathematics. Cambridge university press, Cambridge, New York, Melbourne, 2006.

[20] Simon Thomas. The classification problem for torsion-free abelian groups of finite rank. *J. Amer. Math. Soc*, 16:233–258, 2001.

[21] Todor Tsankov. The additive group of the rationals does not have an automatic presentation. *J. Symb. Log.*, 76(4):1341–1351, 2011.