Université de Liège
Faculté des Sciences Appliquées

On the sets of real vectors recognized by finite automata in multiple bases

Thèse présentée par

Julien Brusten

en vue de l'obtention du grade de

Docteur en Sciences,
orientation Informatique

Année académique 2010-2011

Abstract

This thesis studies the properties of finite automata recognizing sets of real vectors encoded in positional notation using an integer base. We consider both general infinite-word automata, and the restricted class of *weak deterministic automata*, used, in particular, as symbolic data structures for representing the sets of vectors definable in the first order additive theory of real and integer numbers.

In previous work, it has been established that all sets definable in the additive theory of reals and integers can be handled by weak deterministic automata regardless of the chosen numeration base. In this thesis, we address the reciprocal property, proving that the sets of vectors that are simultaneously recognizable in all bases, by either weak deterministic or Muller automata, are those definable in the additive theory of reals and integers.

Precisely, for weak deterministic automata, we establish that the sets of real vectors simultaneously recognizable in two *multiplicatively independent bases* are necessarily definable in the additive theory of reals and integers. For general automata, we show that the multiplicative independence is not sufficient, and we prove that, in this context, the sets of real vectors that are recognizable in two bases that do not share the same set of prime factors are exactly those definable in the additive theory of reals and integers.

Those results lead to a precise characterization of the sets of real vectors that are recognizable in multiple bases, and provide a theoretical justification to the use of weak automata as symbolic representations of sets.

As additional contribution, we also obtain valuable insight into the internal structure of automata recognizing sets of vectors definable in the additive theory of reals and integers.

Acknowledgments

(In French)

Tout d'abord, je voudrais remercier mon promoteur, Bernard Boigelot. Il m'a non seulement proposé et permis de travailler sur un sujet passionnant, mais m'a aussi initié à la recherche scientifique. Sans sa grande disponibilité, ses nombreux conseils, ses bonnes idées et ses encouragements, cette thèse n'aurait sans doute jamais vu le jour.

Véronique Bruyère et Jérôme Leroux ont également fortement infuencé ce travail. Cela a été un plaisir de collaborer avec eux.

Je souhaiterais remercier les membres du jury, Véronique Bruyère, Fabien Durand, Pascal Gribomont, Jérôme Leroux, Michel Rigo et Pierre Wolper, pour le temps qu'ils prendront à évaluer cette thèse.

Je profite de ces quelques lignes pour remercier les collègues avec qui j'ai partagé ces cinq dernières années. Je ne tenterai pas de les citer tous de peur d'en oublier, mais je tiens particulièrement à remercier Jean-François Degbomont et Alexander Linden pour les nombreuses discussions, les temps de midi et les aides ponctuelles lors de la rédaction.

J'ai été financé par le Fonds de la Recherche Scientifique – FNRS en tant qu'aspirant, et ensuite par la Politique scientifique fédérale. Je suis reconnaissant envers ces deux institutions.

Finalement, je voudrais remercier mes amis et ma famille. En particulier, je remercie mes parents pour leur soutien et leur confiance tout au long de mes études.

Contents

List of figures

Chapter 1

Introduction

1.1 Motivations

Computers take a more and more important role in everyday life. A huge range of applications depends on computer programs, and their errors often lead to catastrophic effects. To avoid these consequences, it is fundamental for developers to detect software bugs. This has prompted the development of automated verification methods of computer systems.

A first step for verifying systems algorithmically consists in building an abstract model of the system that has to be analyzed [CGP00]. In this work, we consider models that are composed of a finite-state machine extended with variables. Mathematically, those models are composed of states, each of them characterizing a control location of the system, combined with a value for each variable. If the set of control locations is L and the system contains n variables whose respective domains are $D_1, D_2, \ldots, D_n$, then a state is an element of $L \times D_1 \times D_2 \times \cdots \times D_n$ and can thus be seen as a vector of values. In addition, a set of initial states is defined, as well as a reachability relation defining the ways to arrive at a state from another one. The modelization of concurrent systems can be reduced to the same formalism : The set of locations L becomes a set of tuples of locations, and the reachability relation has to cover all possible interleavings.

We want to verify that the system satisfies safety properties, i.e., that some correctness properties are satisfied for each of its reachable states. For this purpose, a usual technique is to compute the set of reachable states of the system, which can be done by performing a state-space exploration [Wes78, Hol88]. Classically, such an exploration is done by starting from the initial state of the program, and adding new states to the set of reachable states using the reachability relation, until reaching a fixpoint, which is eventually the case when there are finitely many states.

However, when the state space is too large, this technique is not applicable in practice since the number of states that have to be explored becomes huge, and so is the time needed to perform those verifications. The situation is the same when dealing with infinite-state systems, e.g., when unbounded integer variables are used, or when we introduce real variables in the model for representing the continuous flow of time of real-time systems. To solve this problem, a possible approach is to handle sets of states symbolically instead of explicitly [McM93]. We then have to test the safety properties on sets of states instead of individual ones, and, to detect that a fixpoint has been reached, a convenient solution consists in testing inclusion of the reached set of states in the previous one.

When we restrict to applications relying on integer variables (i.e., with respective domains equal to $\mathbb{Z}$) on which linear operations are performed, the sets that need to be represented share a common particular structure, combining essentially linear constraints and discrete periodicities [Boi98]. If we use real variables in addition to integer ones, i.e., when each domain D_i is either $\mathbb{Z}$ of $\mathbb{R}$, then the structure of those sets also contains linear continuous constraints.

Verification tools adapted to mixed integer and real programs whose operations are linear thus have to be based on data structures suited for the representation of such sets. Those data structures must have several properties. First, they need to be expressive enough, i.e., the sets of real and integer variables combining linear constraints and discrete periodicities have to be representable. Next, in order to be able to carry out state-space exploration, classical Boolean operations like union, intersection, complement, Cartesian product, ..., must be efficiently computable on the represented sets, as well as inclusion and emptiness tests. Finally, the representation should use only a reasonable amount of memory.

Several symbolic representation systems are currently known. A widely used solution is to represent sets as arithmetic formulas [CH78, Pug92, BHZ08]. An important drawback of this method is that simplifying a formula into a minimal form is not easy to carry out, and so are the operations consisting in testing if two formulas are equivalent, or in testing inclusion between sets represented by formulas. It follows that a simple set built from a long sequence of operations, which is usually the case in verification applications, is generally represented in a needlessly complicated form.

A completely different approach consists in representing sets by finite automata, which are finite-state machines able to recognize sets of words.

To represent sets of integer numbers with an automaton, one considers a numeration base $r \in \mathbb{N}_{>1}$, and defines an encoding relation that maps positionally numbers to words on the alphabet $\{0, 1, \ldots, r-1\}$. This encoding can easily be extended to represent integer vectors with a fixed dimension by words. If the language of the encodings of a set $S \subseteq \mathbb{Z}^n$ is regular, then it is accepted by a finite automaton which recognizes S. It is known for a long time that this representation system permits to represent all the sets definable in the first-order theory $\langle \mathbb{Z}, +, <, V_r \rangle$, i.e., Presburger arithmetic [Pre29] extended with a base-dependent predicate $V_r(x, y)$ that is true if and only if y is the highest integer power of r dividing x [Büc62, Bru85, BHMV94, WB95, Boi98, Lat05]. Intuitively, the sets definable in Presburger arithmetic are those that can be expressed as a finite combination of linear constraints and periodicities. This method is thus well adapted to those sets. One of its advantages is that performing set operations reduces to carry out the similar operations on languages, for which there exist easy-to-implement and efficient algorithms [HMU01]. In addition, since some forms of finite automata admit a canonical minimal form [Hop71], this method provides a simple canonical representation of sets.

The properties of this representation scheme have been intensively studied. In particular, the well-known *Cobham's and Semenov's* theorems state that if a set $S \subseteq \mathbb{Z}^n$ is simultaneously recognizable in two bases $r, s \in \mathbb{N}_{>1}$ that are *multiplicatively independent*, i.e., such that $r^p \neq s^q$ for all $p, q \in \mathbb{N}_{>0}$, then S is definable in the additive theory of integers $\langle \mathbb{Z}, +, < \rangle$ without use of any base-dependent predicate [Cob69, Sem77]. It follows that such a set is recognizable in every base. Cobham's and Semenov's theorems have been generalized in several ways. Among others, one can cite [Vil92, Fab94, Bès97, PB97, Han98, Bès00, Dur02, AB08, Dur08].

For representing sets definable in mixed integer and real arithmetic, and since a real number has generally an infinite expansion, one can move to infinite-word encodings and ω-regular languages [BBR97]. In this setting, the base-r encodings of numbers and vectors take the form of infinite words over the alphabet $\{0, 1, \ldots, r-1, \star\}$, where "$\star$" is a separator symbol used for distinguishing their integer and fractional parts. The sets recognizable in base r are those definable in $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$, i.e., the first-order additive theory of mixed integer and real variables extended with a predicate $X_r(x, u, k)$ that is true if and only if there exists a base-r encoding of x such that the digit at the position specified by u is equal to k [BRW98]. Similarly, the article [JS01] gives a geometrical characterization of the subsets of $[0, 1]^n$ recognizable in a given base.

Nevertheless, a practical limitation of this approach is the high computational cost of some operations involving infinite-word automata, in particular language complementation [Var07]. However, it has been shown [BJW05] that a restricted form of automata, *weak deterministic* ones, actually suffices for handling the sets definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. This theory is the additive first-order theory of mixed integer and real variables, and it is known that one of its syntactic extensions admits a quantifier elimination procedure [Wei99]. Weak automata can be manipulated with essentially the same cost as finite-word ones [Wil93], which alleviates the problem and leads to an effective representation of the sets of $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

Unfortunately, representing real vectors by automata still presents several drawbacks. For example, the sets defined by linear constraints with large coefficients are often represented in a rather inefficient way. It is thus important to deeply understand the properties of the representation of sets of real vectors by automata, in order to be able to improve its efficiency. In this thesis, we give, among others, a characterization of the sets $S \subseteq \mathbb{R}^n$ simultaneously recognizable in multiple bases, either by general infinite-word automata (i.e., not necessarily weak), or by weak deterministic ones.

1.2 Contributions

The first aim of this thesis consists in extending Cobham's and Semenov's theorems to sets of real vectors by precisely characterizing the sets that are recognizable by infinite-word automata in multiple bases.

We consider the case, relevant for practical applications, of weak deterministic automata. For this kind of automata, we prove that Cobham's and Semenov's theorems can be extended, i.e., that a set of real vectors is simultaneously recognizable by weak deterministic automata in two multiplicatively independent bases if and only if this set is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

This thesis also studies the sets recognized by general infinite-word automata. We establish that there exists a set of real numbers recognizable in two multiplicatively independent bases, but that is not definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. This shows that Cobham's and Semenov's theorems do not directly generalize to general infinite-word automata recognizing sets of real vectors. Hence, it was necessary to find an additional criterion on the bases in which a set has to be recognizable for being definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. This criterion turned out to be the following : Those bases have to admit different sets of prime factors.

We thus establish that a set $S \subseteq \mathbb{R}^n$ is simultaneously recognizable in two bases that do not share the same set of prime factors if and only if S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. As a corollary, such a set must then be recognizable by a weak deterministic automaton.

Since recognizability in two multiplicatively dependent bases is equivalent to recognizability in only one of them, those results provide a complete characterization of the sets that are recognizable in multiple bases by weak deterministic automata or by general infinite word automata. They also lead to a theoretical justification to the use of weak deterministic automata, by showing that the sets recognizable by general infinite-word automata in every encoding base are exactly the same as the ones recognizable by weak deterministic automata in every base.

As an additional contribution, we document precisely the internal structure of automata recognizing sets of vectors definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. This documentation has been exploited in order to improve the efficiency of the decision procedure for that arithmetic [BBD10].

Most of those contributions have first been published in [BB07, BBB08, BB09, BBL09, BBB10].

1.3 Outline

Chapter 2 introduces preliminary notions about numbers, vectors, words, automata and encodings.

In Chapter 3, we establish topological links between sets of vectors and sets of words, we present weak deterministic automata, and we prove some related expressiveness properties.

The problem of characterizing the sets of real vectors that are simultaneously recognized in multiple bases, either by weak deterministic automata or by general ones, is introduced in Chapter 4. This chapter also provides a reduction of this problem.

Chapters 5 and 6 respectively solve this problem for one-dimensional sets and for multi-dimensional ones.

In Chapter 7, we use the results of Chapters 5 and 6 to characterize the internal structure of sets of real vectors that are recognized in multiple bases.

Finally, Chapter 8 concludes this thesis.

Chapter 2

Basic notions

In this chapter, we recall preliminary notions about the representation of numbers and vectors by finite automata. After setting notation for sets of numbers and vectors, and recalling some background of logics, we introduce words, automata, encodings and their properties, and automata recognizing set of vectors. We assume that the reader is familiar with elementary set theory.

2.1 Sets of numbers

2.1.1 Numbers and vectors

The notations $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ respectively represent the sets of natural numbers $\{0, 1, 2, \ldots\}$, integer numbers $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$, rational numbers $\left\{\frac{p}{q} \mid p \in \mathbb{Z} \wedge q \in \mathbb{N}_{>0}\right\}$, and real numbers (containing $\sqrt{2}$, π, ...).

If $S \in \{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$, $\# \in \{>, \geq, <, \leq, \neq\}$, and $x \in \mathbb{R}$, then the notation $S_{\#x}$ is used to represent the set $\{y \in S \mid y\#x\}$.

Example 2.1 *The set of strictly positive integers is denoted by* $\mathbb{N}_{>0}$. ◇

For numbers $x, y \in \mathbb{R}$,

- the notation $[x, y]$ (resp. $[x, y[$, $]x, y]$, $]x, y[$) is used for representing the set $\{z \in \mathbb{R} \mid x \leq z \leq y\}$ (resp. $\{z \in \mathbb{R} \mid x \leq z < y\}$, $\{z \in \mathbb{R} \mid x < z \leq y\}$, $\{z \in \mathbb{R} \mid x < z < y\}$);

- $]-\infty, x]$ (resp. $]-\infty, x[$, $[x, \infty[$, $]x, \infty]$) denotes the set $\{z \in \mathbb{R} \mid z \leq x\}$ (resp. $\{z \in \mathbb{R} \mid z < x\}$, $\{z \in \mathbb{R} \mid z \geq x\}$, $\{z \in \mathbb{R} \mid z > x\}$);

- $|x|$ denotes the absolute value of x;

- $\lfloor x \rfloor$ denotes the greatest integer $z \in \mathbb{Z}$ such that $z \leq x$;

- $\lceil x \rceil$ denotes the smallest integer $z \in \mathbb{Z}$ such that $z \geq x$.

For $x, y \in \mathbb{Z}$ and $n \in \mathbb{N}_{>0}$, the notation $x \equiv_n y$ is used to specify that x and y are congruent modulo n, i.e., that n divides the number $x - y$. Similarly, $x \not\equiv_n y$ denotes the fact that x and y are not congruent modulo n.

If $n \in \mathbb{N}_{>0}$ is a dimension and S is a set, then $S^n = \underbrace{S \times S \ldots \times S}_{n}$ denotes the set of n-dimensional vectors with components in S.

When the dimension n is clear from context, we sometimes use the symbol $\vec{0}$ as a shorthand for $(\underbrace{0, 0, \ldots, 0}_{n})$.

If $\vec{v}, \vec{v'} \in S^n$ are vectors, and $i \in \{1, 2, \ldots, n\}$ is the index of a component, then $\vec{v}[i]$ denotes the ith component of $\vec{v}$, and $\vec{v}\vec{v'}$ represents the scalar product of $\vec{v}$ and $\vec{v'}$, i.e.,

$$\vec{v}\vec{v'} = \sum_{i=1}^{n} \vec{v}[i].\vec{v'}[i].$$

If $m, n \in \mathbb{N}_{>0}$ are dimensions and if $D \subseteq \mathbb{R}$ is a domain, then $D^{m \times n}$ is the set of matrices with m rows and n columns, with components in D. Let $M \in \mathbb{R}^{m \times n}$, and $S \subseteq \mathbb{R}^n$. The set MS is defined by $MS = \{M\vec{x} \mid \vec{x} \in S\}$.

If $\vec{v} = (v_1, v_2, \ldots, v_n) \in S^n$ is a vector, then $\mathrm{diag}(\vec{v})$ is the diagonal matrix

$$\mathrm{diag}(\vec{v}) = \begin{pmatrix} v_1 & 0 & \cdots & 0 \\ 0 & v_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_n \end{pmatrix}.$$

Let $S \subseteq \mathbb{R}^n$. The *affine hull* of S, denoted $\mathrm{aff}(S)$ is

$$\begin{aligned}\mathrm{aff}(S) = \{\vec{x} \in \mathbb{R}^n \quad | \quad & (\exists k \in \mathbb{N}_0) \\ & (\exists \vec{x_1}, \vec{x_2}, \ldots, \vec{x_k} \in S) \\ & (\exists \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}) \\ & \left(\vec{x} = \sum_{i=1}^{k} \lambda_i \vec{x_i} \wedge \sum_{i=1}^{k} \lambda_i = 1\right)\}.\end{aligned}$$

S is said to be an *affine space* if $S = \mathrm{aff}(S)$. Each affine space S is the translation of a vector space; the dimension of this vector space is the *dimension* of S, and is denoted $\dim(S)$.

2.2 Arithmetic theories

In this section, we introduce some logical arithmetic theories that will be used throughout this thesis. We describe those arithmetic theories by structures of first-order logic.

The structures we consider are characterized by *domains* D_i, *constants* in the domain $\bigcup_i D_i$, and *relations* R_i of given arities over $\bigcup_i D_i$. Formulas are built from variables (each of these variables is associated with a domain D_i), constants and relations, and are closed under finite Boolean combinations, existential and universal quantifications. If φ is a formula, and if $\{x_1, x_2, \ldots, x_n\}$ is the set of variables in φ that are not quantified, then one can attribute a validity value to $\varphi(x_1, x_2, \ldots, x_n)$ by fixing to the variables $x_1, x_2, \ldots, x_n$ values of $\bigcup_i D_i$, and by applying the usual semantic rules. If $S = \{(x_1, x_2, \ldots, x_n) \in (\bigcup_i D_i)^n \mid \varphi(x_1, x_2, \ldots, x_n)\}$, then S is said to be *definable* by φ.

2.2.1 Presburger arithmetic and additive theory of integers

Presburger arithmetic is defined by the structure $\langle \mathbb{N}, + \rangle$ [Pre29]. The symbol "+" denotes the ternary relation $+(x, y, z)$ containing the tuples $(x, y, z) \in \mathbb{N}^3$ such that $x + y = z$. For convenience, we write in the sequel $x + y = z$ instead of $+(x, y, z)$.

Example 2.2 *The set $\{(x, y) \in \mathbb{N}^2 \mid x \leq y\}$ is definable in Presburger arithmetic by the formula $(\exists z \in \mathbb{N})(x + z = y)$. Remark that the relation $x = y$ is definable by $x + 0 = y$. The set $\{(x, y) \in \mathbb{N}^2 \mid x < y\}$ is also definable.*

For a given $n \in \mathbb{N}_{>2}$*, the relation* $x = x_1 + x_2 + \ldots + x_n$ *is definable by* $(\exists y_1 \in \mathbb{N})(\exists y_2 \in \mathbb{N}) \ldots (\exists y_{n-2} \in \mathbb{N})(x = x_1 + y_1 \wedge y_1 = x_2 + y_2 \wedge \ldots \wedge y_{n-2} = x_{n-1} + x_n)$*.*

Each constant $c \in \mathbb{N}$ *is definable in Presburger arithmetic. Indeed, the constant* 0 *is defined by* $\{x \in \mathbb{N} \mid x + x = x\}$*, the constant* 1 *is defined by* $\{x \in \mathbb{N} \mid x > 0 \wedge (\forall y \in \mathbb{N})(y > 0 \Rightarrow y \geq x)\}$*, and the constant* n *(with* $n \geq 1$*) is defined by* $\{x \in \mathbb{N} \mid x = \underbrace{1 + 1 + \ldots + 1}_{n}\}$*.*

The set of natural numbers that are multiples of a constant n *is definable by the formula* $\{x \in \mathbb{N} \mid (\exists y \in \mathbb{N})(x = ny)\}$*, where* ny *is an abbreviation for* $\underbrace{y + y + \ldots + y}_{n}$*.* ◇

It is proved in [GS66] that the sets definable in Presburger arithmetic coincide with the *semilinear sets*, defined as followed.

Definition 2.3 *Let* $n \in \mathbb{N}_{>0}$ *be a dimension. A set* $S \subseteq \mathbb{N}^n$ *is* linear *if there exists a finite set of vectors* $\vec{c}, \vec{x_1}, \vec{x_2}, \ldots, \vec{x_m} \in \mathbb{N}^n$ *such that*

$$S = \{\vec{c} + \sum_{i=1}^{m} c_i \vec{x_i} \mid c_1, c_2, \ldots, c_m \in \mathbb{N}\}.$$

A set $S \subseteq \mathbb{N}^n$ *is* semilinear *if it can be expressed as a finite union of linear sets.* □

Theorem 2.4 ([GS66]) *Let* $n \in \mathbb{N}_{>0}$ *be a dimension. A set* $S \subseteq \mathbb{N}^n$ *is semilinear if and only if it is definable in* $\langle \mathbb{N}, + \rangle$*.*

In particular, the one-dimensional sets $S \subseteq \mathbb{N}$ that are definable in Presburger arithmetic are exactly the ultimately periodic sets, i.e., the sets that differ from a periodic subset of $\mathbb{N}$ only by a finite set. More formally, we have the following corollary.

Corollary 2.5 *A set* $S \subseteq \mathbb{N}$ *is definable in* $\langle \mathbb{N}, + \rangle$ *if and only if there exists* $k \in \mathbb{N}_{>0}$ *such that*

$$(\forall x \in \mathbb{N}_{\geq k})(x \in S \Leftrightarrow x + k \in S).$$

Presburger arithmetic can be directly extended to the additive theory of integers $\langle \mathbb{Z}, +, < \rangle$. The sets definable in that theory can also be characterized with an adapted notion of semilinear sets :

Definition 2.6 *Let $n \in \mathbb{N}_{>0}$ be a dimension. A set $S \subseteq \mathbb{Z}^n$ is* linear *if there exists a finite set of vectors $\vec{c}, \vec{x_1}, \vec{x_2}, \ldots, \vec{x_m} \in \mathbb{Z}^n$ such that*

$$S = \{\vec{c} + \sum_{i=1}^{m} c_i \vec{x_i} \mid c_1, c_2, \ldots, c_m \in \mathbb{N}\}.$$

A set $S \subseteq \mathbb{Z}^n$ is semilinear *if it can be expressed as a finite union of linear sets.* □

2.2.2 Additive theory of reals

The additive theory of reals is defined by the structure $\langle \mathbb{R}, +, <, 1 \rangle$.

This theory admits the elimination of quantifiers [FR75]. Hence, the sets that are definable in $\langle \mathbb{R}, +, <, 1 \rangle$ are exactly the sets that can be expressed as finite Boolean combinations of linear constraints with rational coefficients. As a consequence, the one-dimensional sets $S \subseteq \mathbb{R}$ that are definable in $\langle \mathbb{R}, +, <, 1 \rangle$ are exactly the finite unions of intervals with rational boundaries, as noticed in [Wei99].

2.2.3 Additive theory of reals and integers

In this thesis, the additive theory of reals and integers will be used intensively : It is the theory of the structure $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. This theory thus contains two domains ($\mathbb{R}$ and $\mathbb{Z}$), which implies that each (possibly quantified) variable is either real or integer. The quantifiers that are available for building formulas are $(\forall x \in \mathbb{Z})$, $(\exists x \in \mathbb{Z})$, $(\forall x \in \mathbb{R})$ and $(\exists x \in \mathbb{R})$.

In the following example, extracted from [BJW05], we consider a set definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

Example 2.7 *The set depicted in Figure 2.1, representing a periodic tiling of triangles, is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. Indeed, it can be defined by the formula*

$$\begin{aligned}\{(x_1, x_2) \in \mathbb{R}^2 \mid & (\exists x_3, x_4 \in \mathbb{R})(\exists x_5, x_6 \in \mathbb{Z}) \\ & (x_1 = x_3 + 2x_5 \wedge x_2 = x_4 + 2x_6 \wedge \\ & x_3 \geq 0 \wedge x_4 \leq 1 \wedge x_4 \geq x_3)\}.\end{aligned}$$

◇

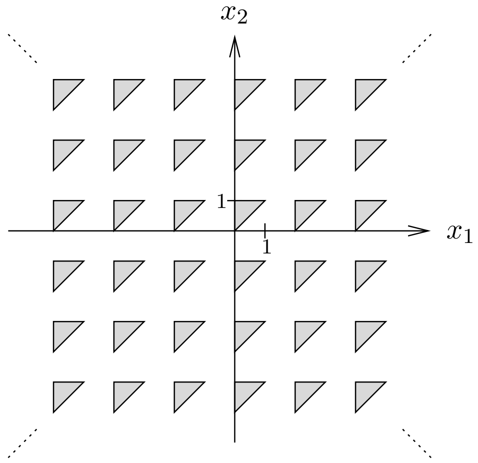

Figure 2.1: Set definable in the additive theory of reals and integers.

Remark that any set definable in $\langle \mathbb{R}, +, <, 1 \rangle$ is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$ as well.

2.3 Words and automata

In this section, we give basic definitions and facts about finite and infinite words, and about different kinds of automata recognizing sets of such words.

2.3.1 Words and languages

An *alphabet* Σ is a finite set of *symbols*.

Definition 2.8 *A* word *defined over an alphabet* Σ *is a (finite or infinite) sequence of symbols of* Σ.

The length *of a finite word* w*, denoted by* $|w|$*, is the number of symbols it contains. The unique finite word* w *of length* $|w| = 0$ *is the* empty word*, denoted* ε.

Given a finite word w_1 *and a (possibly infinite) word* w_2*, the word* $w_1 w_2$ *denotes the concatenation of* w_1 *and* w_2*, in that order.* □

A set of words defined over a given alphabet is called a *language*.

Definition 2.9 *Let Σ be a given alphabet. A* finite-word language *(resp.* infinite-word language*, or* ω-language*) is a (possibly infinite) set of finite words (resp. infinite words) over Σ.* □

Languages can be combined together :

Definition 2.10 *The* concatenation *of a finite-word language L_1 with a (possibly infinite-word) language L_2 is the language $L_1L_2 = \{w_1w_2 \mid w_1 \in L_1 \wedge w_2 \in L_2\}$.*

The iterative closure *(or* Kleene's closure*) of a finite-word language L is the language $L^* = \{w \mid (\exists k \in \mathbb{N})(\exists w_1, w_2, \ldots, w_k \in L)(w = w_1w_2 \ldots w_k)\}$. The abbreviation L^+ stands for the language LL^*.*

An operator similar to Kleene's closure exists for generating infinite languages : If L is a finite-word language, then $L^\omega = \{w_0w_1w_2 \ldots \mid (\forall i \in \mathbb{N})(w_i \in L \wedge w_i \neq \varepsilon)\}$. □

As formalized in the next definition, each infinite word w is either aperiodic, or ultimately periodic. Particular ultimately periodic words are purely periodic ones.

Definition 2.11 *Let Σ be an alphabet, and $w \in \Sigma^\omega$ be an infinite word over Σ. If there exist $w_1, w_2 \in \Sigma^*$ such that $w_2 \neq \varepsilon$ and $w = w_1w_2^\omega$, then w is* ultimately periodic*. If w_1 can be set to the empty word ε, then w is* purely periodic*. Otherwise, w is* aperiodic*.* □

2.3.2 Automata

Automata are finite-state machines that are able to recognize (possibly infinite) sets of words. There exist several kinds of automata, each of them recognizing words of a given nature, and having a particular expressive power. They have been studied for a long time [Kle56, RS59, Büc60, Büc62, McN66]. In this section, we introduce automata adapted to finite and infinite words.

Automata on finite words

Definition 2.12 *A* (possibly non-deterministic) finite automaton on finite words *is, syntactically, a quintuple*

$$(Q, \Sigma, \Delta, q_0, F),$$

where

- Q *is a finite set of* states*;*
- Σ *is an alphabet;*
- $\Delta \subseteq Q \times \Sigma^* \times Q$ *is a* transition relation*;*
- $q_0 \in Q$ *is an* initial state*;*
- $F \subseteq Q$ *is a set of* accepting states*.*

□

In the previous definition, we restrict ourselves, but without loss of generality, to consider a unique initial state q_0 instead of a set of initial states.

For states $q, q' \in Q$ and a word $w \in \Sigma^*$, we write $(q, w, q') \in \Delta^*$ if and only if there exist a number $k \in \mathbb{N}$, words $w_1, w_2, \ldots, w_k \in \Sigma^*$ and states $q_1, q_2, \ldots, q_{k-1}$ such that $w = w_1 w_2 \ldots w_k$ and

$$\{(q, w_1, q_1), (q_1, w_2, q_2), (q_2, w_3, q_3), \ldots, (q_{k-1}, w_k, q')\} \subseteq \Delta.$$

Intuitively, a finite automaton is *deterministic* if the reading of a word from a given state is unique, i.e., if, for every $q \in Q$ and $w \in \Sigma^*$, there exists at most one state $q' \in Q$ such that $(q, w, q') \in \Delta^*$. In the sequel, for technical reasons, we use a syntactic restriction of deterministic automata. Formally, deterministic automata are defined as follows.

Definition 2.13 *A finite automaton on finite words $\mathcal{A} = (Q, \Sigma, \Delta, q_0, F)$ is said to be* deterministic *if, for each $(q, w, q') \in \Delta$, the following two conditions hold :*

- $|w| = 1$*, and*
- $(\forall (q_1, w', q_2) \in \Delta)((q_1 = q \wedge w' = w) \Rightarrow q_2 = q')$.

□

In other words, the transition relation Δ of a deterministic finite automaton $\mathcal{A} = (Q, \Sigma, \Delta, q_0, F)$ can be expressed as a partial transition function

$\delta : Q \times \Sigma \to Q$ such that, for each $q, q' \in Q$ and $\sigma \in \Sigma$, $(q, \sigma, q') \in \Delta$ if and only if $\delta(q, \sigma) = q'$.

From the transition function of a deterministic finite automaton, one can extend its definition, moving its domain from single symbols to words.

Definition 2.14 *Let $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ be a deterministic finite automaton on finite words. The function $\delta^* : Q \times \Sigma^* \to Q$ is defined in the following way : For $q \in Q$ and $w = \sigma_0\sigma_1 \dots \sigma_{k-1}$, where $\sigma_i \in \Sigma$ for each $i \in \{0, 1, \dots, k-1\}$, we have*

$$\delta^*(q, w) = q'$$

if and only if

$$\delta(\dots \delta(\delta(q, \sigma_0), \sigma_1) \dots, \sigma_{k-1}) = q'.$$

□

A word w is accepted by a finite automaton on finite words $\mathcal{A}$ if there exists a path in $\mathcal{A}$ labeled by w, ending in an accepting state.

Definition 2.15 *A word $w \in \Sigma^*$ is* accepted *by a finite automaton on finite words $\mathcal{A} = (Q, \Sigma, \Delta, q_0, F)$ if there exist $n \in \mathbb{N}$, $w_1, w_2, \dots, w_n \in \Sigma^*$ and $q_1, q_2, \dots, q_n \in Q$ such that $w = w_1w_2 \dots w_n$, $q_n \in F$ and, for each $i \in \{1, 2, \dots, n\}$, $(q_{i-1}, w, q_i) \in \Delta$.* □

Definition 2.16 *The* language accepted *by a finite automaton on finite words $\mathcal{A}$ is the set of words accepted by $\mathcal{A}$, and is denoted by $L(\mathcal{A})$.* □

Example 2.17 *Let L be the language of words, built on the alphabet $\{a, b\}$, that contain an odd number of symbols a. The language L is accepted by the finite automaton $(\{1, 2\}, \{a, b\}, \Delta, 1, \{2\})$, where Δ is the set of transitions*

$$\{(1, a, 2), (1, b, 1), (2, a, 1), (2, b, 2)\}.$$

Alternatively, this automaton is represented by the graph depicted in Figure 2.2. Circles denote states, arrows denote transitions, the initial status of State 1 is expressed with the incoming arrow, and the accepting state is denoted by a double-circle. ◇

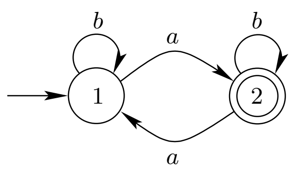

Figure 2.2: Automaton accepting the set of finite words on the alphabet $\{a, b\}$ that contain an odd number of symbols a.

Automata on infinite words

Finite automata on finite words can be adapted to recognize sets of infinite words. A classical set of such automata are *Büchi automata*, and share the syntax of finite automata on finite words.

Definition 2.18 *A* (possibly non-deterministic) Büchi automaton *is, syntactically, a quintuple*

$$(Q, \Sigma, \Delta, q_0, F),$$

where Q is a finite set of states, Σ is an alphabet, $\Delta \subseteq Q \times \Sigma^* \times Q$ is a transition relation, $q_0 \in Q$ is an initial state, and $F \subseteq Q$ is a set of accepting states. □

Similarly to automata on finite words, the transition relation Δ can be extended to Δ^*.

An infinite word w is accepted by a Büchi automaton $\mathcal{A}$ if there exists a path in $\mathcal{A}$ labeled by w, and that visits infinitely often accepting states.

Definition 2.19 *An infinite word $w \in \Sigma^\omega$ is* accepted *by a Büchi automaton $\mathcal{A} = (Q, \Sigma, \Delta, q_0, F)$ if there exist $w_1, w_2, \ldots \in \Sigma^*$ and $q_1, q_2, \ldots \in Q$ such that $w = w_1 w_2 \ldots$, for each $i \in \mathbb{N}_{>0}$, $(q_{i-1}, w_i, q_i) \in \Delta$, and, for an infinite number of values $i \in \mathbb{N}$, $q_i \in F$.* □

Definition 2.20 *The* language accepted *by a Büchi automaton $\mathcal{A}$ is the set of infinite words accepted by $\mathcal{A}$, and is denoted by $L(\mathcal{A})$.* □

Example 2.21 *Let L be the ω-language built on the alphabet $\{a, b\}$ of the words that end either with a^ω, or with b^ω. The language L is accepted*

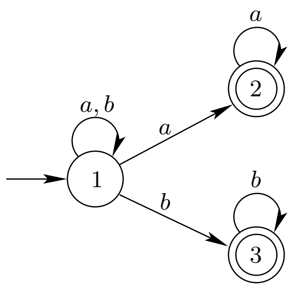

Figure 2.3: Automaton accepting the set of infinite words on the alphabet $\{a, b\}$ that end either with a^ω or with b^ω.

by the Büchi automaton $(\{1, 2, 3\}, \{a, b\}, \Delta, 1, \{2, 3\})$*, where* Δ *is the set of transitions*

$$\{(1, a, 1), (1, b, 1), (1, a, 2), (1, b, 3), (2, a, 2), (3, b, 3)\}.$$

Figure 2.3 contains the graph representation of this automaton. ◇

The following two theorems are proved in [PP04], and state that the languages accepted by Büchi automata are completely determined by their ultimately periodic words.

Theorem 2.22 ([PP04]) *Let L be a non-empty language. If L is accepted by a Büchi automaton, then L contains at least one ultimately periodic word.*

Theorem 2.23 ([PP04]) *Let L_1 and L_2 be languages accepted by Büchi automata. The languages L_1 and L_2 are equal if and only if they contain the same ultimately periodic words.*

It is known that the expressive power of (possibly non-deterministic) Büchi automata is strictly more important than the expressive power of deterministic ones. For example, it can be shown that the language of the words built on the alphabet $\{a, b\}$ and containing only finitely many occurrences of the symbol a is accepted by a non-deterministic Büchi automaton, but cannot be accepted by a deterministic one [Saf88].

However, there exists a class of deterministic infinite-word automata, known as *Muller automata*, and that shares the expressive power of non-deterministic Büchi automata. This class is obtained by using several sets of accepting states, instead of only one, and by adapting the notion of accepted word.

Definition 2.24 *A* (deterministic) Muller automaton *is, syntactically, a quintuple*

$$(Q, \Sigma, \delta, q_0, F),$$

where

- Q *is a finite set of* states*;*
- Σ *is an alphabet;*
- $\delta : Q \times \Sigma \to Q$ *is a* partial transition function*;*
- $q_0 \in Q$ *is an* initial state*;*
- $F \subseteq 2^Q$ *is a set of sets of* accepting states *(*2^Q *denotes the powerset of* Q*, i.e., its set of subsets).*

□

The function δ can be extended to a function δ^*, reading words instead of single symbols.

For a word w to be accepted by a Muller automaton, the set of states visited infinitely often during the reading of w has to be exactly a set in F.

Definition 2.25 *A word* $w \in \Sigma^\omega$ *is* accepted *by a Muller automaton* $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ *if there exist* $\sigma_1, \sigma_2, \ldots \in \Sigma$ *and* $q_1, q_2, \ldots \in Q$ *such that* $w = w_1 w_2 \ldots$*, for each* $i \in \mathbb{N}_{>0}$*,* $\delta(q_{i-1}, \sigma_i) = q_i$*, and the set* S *of states that occur infinitely many times in the sequence* $q_1, q_2, \ldots$*, is such that* $S \in F$*.* □

The following theorem is due to McNaughton.

Theorem 2.26 ([McN66, PP04]) *An* ω*-language* L *is accepted by a (possibly non-deterministic) Büchi automaton if and only if it is accepted by a Muller automaton.*

The notion of *co-Büchi* automata will also be useful. The syntax of such automata is exactly the one of Büchi automata, but their accepting condition specifies that accepting runs visit only finitely often accepting states, instead of infinitely often.

Definition 2.27 *A word* $w \in \Sigma^\omega$ *is* accepted *by a co-Büchi automaton* $\mathcal{A} = (Q, \Sigma, \Delta, q_0, F)$ *if there exist* $w_1, w_2, \ldots \in \Sigma^*$ *and* $q_1, q_2, \ldots \in Q$ *such that* $w = w_1 w_2 \ldots$*,* $(q_{i-1}, w, q_i) \in \Delta$ *for each* $i \in \mathbb{N}_{>0}$*, and* $q_i \in F$ *only for a finite number of values* $i \in \mathbb{N}$*.* □

2.3.3 Algorithms

A large set of operations can be applied to languages accepted by automata.

Theorem 2.28 ([HMU01]) *Let*

$$\mathcal{A}_1 = (Q, \Sigma, \Delta, q_0, F)$$

and

$$\mathcal{A}_2 = (Q', \Sigma, \Delta', q_0', F')$$

be two finite automata on finite words, defined on the same alphabet Σ*, and let* L_1 *and* L_2 *be the languages respectively accepted by* $\mathcal{A}_1$ *and* $\mathcal{A}_2$*. There exist algorithms for computing automata accepting the languages* $L_1 \cup L_2$*,* $L_1 \cap L_2$*,* $L_1 \times L_2$ *and* $L_1 \setminus L_2$*.*

Intuitively, the algorithms of Theorem 2.28 compute a *product automaton* of $\mathcal{A}_1$ and $\mathcal{A}_2$. In this automaton, the states are composed of pairs (q_1, q_2), with $q_1 \in Q$ and $q_2 \in Q'$. The transition relation and the accepting condition need to be adapted according to the operation that is achieved.

Definition 2.29 *Let* $\Sigma = \Sigma_1 \times \Sigma_2 \times \cdots \times \Sigma_n$*, with* $n \in \mathbb{N}_{>1}$*.*

Let $\vec{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n) \in \Sigma$*. The* projection of $\vec{\sigma}$ on each component except the ith *(*$i \in \{1, 2, \ldots, n\}$*) is* $\vec{\sigma}|_{\neq i} = (\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots, \sigma_n)$*.* □

Definition 2.30 *Let* $\Sigma = \Sigma_1 \times \Sigma_2 \times \cdots \times \Sigma_n$*, with* $n \in \mathbb{N}_{>1}$*.*

Let $w \in \Sigma^*$*. If* $w = \vec{\sigma_1}\vec{\sigma_2}\ldots\vec{\sigma_k}$*, then the* projection of w on each component except the ith *is* $w|_{\neq i} = \vec{\sigma_1}|_{\neq i}\vec{\sigma_2}|_{\neq i}\ldots\vec{\sigma_k}|_{\neq i}$*.*

Let $L \subseteq \Sigma^*$*. The* projection of L on each component except the ith *(*$i \in \{1, 2, \ldots, n\}$*) is defined as the set*

$$L|_{\neq i} = \{w \in (\Sigma_1 \times \cdots \times \Sigma_{i-1} \times \Sigma_{i+1} \times \cdots \times \Sigma_n)^* \mid (\exists w' \in L)(w = w'|_{\neq i})\}.$$

□

Given a finite automaton on finite words $\mathcal{A}$, a (possibly non-deterministic) finite automaton accepting the projection of the accepted language of $\mathcal{A}$ can be obtained by removing the ith component from the transition's labels of $\mathcal{A}$. The following theorem formalizes the existence of an algorithm performing that operation.

Theorem 2.31 ([HMU01]) *Let $\mathcal{A} = (Q, \Sigma_1 \times \Sigma_2 \times \cdots \times \Sigma_n, \Delta, q_0, F)$ be a finite automaton on finite words, and let $i \in \{1, 2, \ldots, n\}$. There exists an algorithm that computes an automaton accepting the projection $L(\mathcal{A})|_{\neq i}$ of the language accepted by $\mathcal{A}$ on each on its components, except the ith.*

As noticed in the following definition and theorem, deterministic finite automata on finite words admit canonical minimal forms.

Definition 2.32 *Let $\mathcal{A} = (Q, \Sigma, \Delta, q_0, F)$ be a deterministic finite automaton on finite words, and let L be the language accepted by $\mathcal{A}$. The automaton $\mathcal{A}$ is* minimal *if and only if for each deterministic finite automaton $\mathcal{A}' = (Q', \Sigma, \Delta', q'_0, F')$ accepting L, we have $|Q| \leq |Q'|$.* □

Theorem 2.33 ([Hop71]) *Let $\mathcal{A}$ be a deterministic finite automaton on finite words. A minimal deterministic finite automaton accepting the same language as $\mathcal{A}$ always exists, and is unique up to isomorphism. Moreover, there exists an algorithm that computes it.*

Given a non-deterministic finite automaton $\mathcal{A}$ on finite words, it is possible to compute a deterministic finite automaton $\mathcal{A}'$ accepting the same language : The basic idea is that each state of $\mathcal{A}'$ corresponds to a subset of the states of $\mathcal{A}$. The states of $\mathcal{A}'$ are then aimed at memorizing the possible sets of states in which $\mathcal{A}$ could be at a given instant of its execution.

Theorem 2.34 ([HMU01]) *A finite-word language is accepted by a finite automaton if and only if it is accepted by a deterministic finite automaton. Moreover, this deterministic automaton is computable in time $\mathcal{O}(2^n)$, where n is the number of states of the non-deterministic automaton.*

Some results also exist for Büchi automata.

Theorem 2.35 ([Büc62]) *Let*

$$\mathcal{A}_1 = (Q, \Sigma, \Delta, q_0, F)$$

and

$$\mathcal{A}_2 = (Q', \Sigma, \Delta', q'_0, F')$$

be two Büchi automata, defined on the same alphabet Σ, and let L_1 and L_2 be the languages accepted respectively by $\mathcal{A}_1$ and $\mathcal{A}_2$. There exist algorithms for computing Büchi automata accepting the languages $L_1 \cup L_2$, $L_1 \cap L_2$, $L_1 \times L_2$ and $L_1 \setminus L_2$.

Similarly to the case of automata on finite-words, the algorithms cited in Theorem 2.35 intuitively perform a product construction of the automata $\mathcal{A}_1$ and $\mathcal{A}_2$ and adapt the transition relation and the accepting condition of the resulting automaton. However, the case $L_1 \setminus L_2$ is practically highly non trivial.

Definition 2.36 *Let* $\Sigma = \Sigma_1 \times \Sigma_2 \times \cdots \times \Sigma_n$*, with* $n \in \mathbb{N}_{>1}$*.*

Let $w \in \Sigma^\omega$*. If* $w = \vec{\sigma_1}\vec{\sigma_2}\ldots$*, then the* projection of w on each component except the ith *is* $w|_{\neq i} = \vec{\sigma_1}|_{\neq i}\vec{\sigma_2}|_{\neq i}\ldots$*.*

Let $L \subseteq \Sigma^\omega$*. The* projection of L on each components except the ith *(*$i \in \{1, 2, \ldots, n\}$*) is defined as the set*

$$L|_{\neq i} = \{w \in (\Sigma_1 \times \cdots \times \Sigma_{i-1} \times \Sigma_{i+1} \times \cdots \times \Sigma_n)^\omega \mid (\exists w' \in L)(w = w'|_{\neq i})\}.$$

□

It is possible to compute a Büchi automaton accepting the projection of the language accepted by another Büchi automaton $\mathcal{A}$ by, intuitively, removing the ith components from the labels of the transitions of $\mathcal{A}$.

Theorem 2.37 ([Büc62]) *Let* $\mathcal{A} = (Q, \Sigma_1 \times \Sigma_2 \times \cdots \times \Sigma_n, \Delta, q_0, F)$ *be a Büchi automaton, and let* $i \in \{1, 2, \ldots, n\}$*. There exists an algorithm that computes a Büchi automaton accepting the projection* $L(\mathcal{A})|_{\neq i}$ *of the language accepted by* $\mathcal{A}$ *on each on its components, except the* i*th.*

Contrary to finite-word automata, Büchi automata do not generally have neither deterministic versions, nor minimized canonical ones [Saf88, Sta83].

Note that the problem of complementing Büchi automata is solvable, as a consequence of Theorem 2.35, but is known to be much more difficult than the corresponding problem for finite-word automata : There do not exist constructions that are usable in practice. The article [Var07] presents a survey of this problem.

Finally, let us mention that co-Büchi automata can always be determinized. The worst-case complexity of the determinization is very similar to the same problem for finite-word automata.

Theorem 2.38 ([KV97, MH84]) *An infinite-word language is accepted by a co-Büchi automaton if and only if it is accepted by a deterministic co-Büchi automaton. Moreover, this deterministic automaton is computable in time* $\mathcal{O}(2^n)$*, where* n *is the number of states of the non-deterministic automaton.*

2.4 Words representing real vectors

We have seen in the previous section that automata recognize sets of words. In order to make them represent sets of vectors, one thus needs to encode such vectors into words. This is the subject of the current section, in which we will also study the properties of the encodings of numbers by words.

2.4.1 Encodings of integers and vectors of integers

Let $r \in \mathbb{N}_{>1}$ be a *numeration base*. The alphabet Σ_r is defined as $\Sigma_r = \{0, 1, \ldots, r-1\}$, and each of its symbols is called a *digit*.

Using the classical positional encoding in base r, each natural number $n \in \mathbb{N}$ can be encoded by a word w of sufficiently large length $p \in \mathbb{N}_{>0}$, that is composed of digits in Σ_r. The word $w = \sigma_{p-1}\sigma_{p-2}\ldots\sigma_0$, where $\sigma_i \in \Sigma_r$ for each $i \in \{0, 1, \ldots, p-1\}$, *encodes* the number n if

$$n = \sum_{i=0}^{p-1} \sigma_i r^i.$$

The digits σ_{p-1} and σ_0 are respectively called the *most significant* and the *least significant digits* of w.

By using the r-complement scheme, this encoding can be generalized in order to represent signed integer numbers : The encodings of $z \in \mathbb{Z}_{<0}$ are formed by the last p digits of the encodings of $r^p + z$. The length p of z is not fixed but has to be large enough for $-r^{p-1} \leq z < r^{p-1}$ to hold. As a consequence, the most significant digit of an encoding is equal to 0 for positive integers and to $r-1$ for negative ones.

Formally, an encoding can be decoded using the following definition.

Definition 2.39 *Let $r \in \mathbb{N}_{>1}$ be a base. A word $w = \sigma_{p-1}\sigma_{p-2}\ldots\sigma_0$, with $p \in \mathbb{N}_{>0}$, $\sigma_{p-1} \in \{0, r-1\}$ and $\sigma_i \in \{0, 1, \ldots, r-1\}$ for each $i \in \{0, 1, \ldots, p-2\}$ is a* r-encoding *of an integer number $z \in \mathbb{Z}$ if and only if*

$$z = \sum_{i=0}^{p-2} \sigma_i r^i + \begin{cases} 0 & \text{if } \sigma_{p-1} = 0; \\ -r^{p-1} & \text{otherwise .} \end{cases}$$

If a word $w \in \{0, r-1\}\{0, 1, \ldots, r-1\}^*$ is a r-encoding of the integer $z \in \mathbb{Z}$, then we write $z = \langle w \rangle_r$, and we say that w encodes *z in base r.* □

Algorithm 1 `encodeInteger`$(z \in \mathbb{Z}, r \in \mathbb{N}_{>1})$

$negative \leftarrow (z < 0)$
if $negative$ **then**
  $p \leftarrow 1$
  while $z < -r^{p-1}$ **do**
    $p \leftarrow p + 1$
  end while
  $z \leftarrow z + r^p$
end if
$enc \leftarrow \varepsilon$
while $z > 0$ **do**
  $\sigma \leftarrow z \bmod r$
  $enc \leftarrow \sigma\ enc$
  $z \leftarrow \lfloor z/r \rfloor$
end while
if $\neg(negative)$ **then**
  $enc \leftarrow 0\ enc$
end if
return enc

It follows from Definition 2.39 that the first digit of an encoding, called its *sign digit*, can be repeated without altering the number that is encoded, as expressed in the following theorem.

Theorem 2.40 *Let $r \in \mathbb{N}_{>1}$ be a base, let $\sigma \in \{0, r-1\}$ and let $w \in \{0, 1, \ldots, r-1\}^*$. It holds that $\langle \sigma w \rangle_r = \langle \sigma \sigma w \rangle_r$.*

Proof If $\sigma = 0$, then the conclusion directly follows from Definition 2.39. On the other hand, if $\sigma = r - 1$, and if $w = \sigma_{|w|-1}\sigma_{|w|-2} \ldots \sigma_0$, then, by Definition 2.39,

$$\langle (r-1)(r-1)w \rangle_r = -r^{|w|+1} + (r-1)r^{|w|} + \sum_{i=0}^{|w|-1} \sigma_i r^i = \langle (r-1)w \rangle_r.$$

■

Algorithm 1 computes the encoding `encodeInteger`$(z \in \mathbb{Z}, r \in \mathbb{N}_{>1})$ of z in base r that has the smallest possible length.

This encoding scheme can be extended to vectors of $\mathbb{Z}^n$, where $n \in \mathbb{N}_{>0}$ is a given dimension. This extension can be achieved by extending the used alphabet : Each symbol will represent a tuple of digits, instead of a single one. The digits of the components of a vector $\vec{v} \in \mathbb{Z}^n$ are then read synchronously, provided that the encodings of the components of $\vec{v}$ share the same length. This is always possible, since the most significant digit of an encoding w of a number $z \in \mathbb{Z}$ can be repeated without altering the fact that w encodes z, as shown in Theorem 2.40.

The encoding of vectors of integers is formalized in the following definition.

Definition 2.41 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $r \in \mathbb{N}_{>1}$ be a base. A word $w = \vec{\sigma_{p-1}}\vec{\sigma_{p-2}} \dots \vec{\sigma_0}$, with $p \in \mathbb{N}_{>0}$, $\vec{\sigma_{p-1}} \in \{0, r-1\}^n$ and $\vec{\sigma_i} \in \{0, 1, \dots, r-1\}^n$ for each $i \in \{0, 1, \dots, p-2\}$ is a* r-encoding *of a vector $\vec{z} \in \mathbb{Z}^n$ if and only if, for each $j \in \{1, 2, \dots, n\}$, we have*

$$\langle \vec{\sigma_{p-1}}[j]\vec{\sigma_{p-2}}[j] \dots \vec{\sigma_0}[j] \rangle_r = \vec{z}[j].$$

The notation $\langle w \rangle_{r,n} = \vec{z}$ is used to mean that w is a r-encoding of $\vec{z}$. □

Example 2.42 *The word*

$$\begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

is a 3*-encoding of the vector* $(3, 2, -9)$*.* ◇

Remark that the alphabet used for encoding vectors has a size r^n, which becomes problematic when n is not restricted to have small values. A solution consists in *serializing* the encodings, by reading successively the digits of the components [Boi98, WB00].

Example 2.43 *The word* 002100020 *is a serialized* 3*-encoding of the vector* $(3, 2, -9)$*.* ◇

In this thesis, we only consider encodings that are not serialized (i.e., that are *synchronized*), but our results can be adapted to serialized ones.

2.4.2 Encodings of reals and vectors of reals

Since real numbers have generally an infinite expansion in base r, one cannot encode them positionally using only a finite sequence of digits. We will instead use infinite words.

Each real number x can (non necessarily uniquely) be decomposed into $x = x_I + x_F$, where $x_I \in \mathbb{Z}$ and $x_F \in [0, 1]$. Using this decomposition, the word $w = w_I \star w_F$ encodes the number x, where w_I is an encoding of x_I (i.e., $\langle w_I \rangle_r = x_I$), and w_F is an infinite word that encodes x_F. The special symbol $\star$ in w is just aimed at separating the encodings of x_I and x_F.

This leads to the following definition.

Definition 2.44 *Let $r \in \mathbb{N}_{>1}$ be a base. A word $w = \sigma_{p-1}\sigma_{p-2}\ldots\sigma_0 \star \sigma'_1\sigma'_2\ldots$, with $p \in \mathbb{N}_{>0}$, $\sigma_{p-1} \in \{0, r-1\}$, $\sigma_i \in \{0, 1, \ldots, r-1\}$ for each $i \in \{0, 1, \ldots, p-2\}$, and $\sigma'_i \in \{0, 1, \ldots, r-1\}$ for each $i \in \mathbb{N}_{>0}$ is a r-*encoding *of $x \in \mathbb{R}$ if and only if*

$$x = \sum_{i=0}^{p-2} \sigma_i r^i + \sum_{i \in \mathbb{N}_{>0}} \sigma'_i r^{-i} + \begin{cases} 0 & \text{if } \sigma_{p-1} = 0; \\ -r^{p-1} & \text{otherwise}. \end{cases}$$

If a word $w \in \{0, r-1\}\{0, 1, \ldots, r-1\}^* \star \{0, 1, \ldots, r-1\}^\omega$ is a r-encoding of $x \in \mathbb{R}$, then we write $x = \langle w \rangle_r$, and we say that w encodes *x in base r.* □

Example 2.45 *The word $200\star(10)^\omega$ is a 3-encoding of the number $-9 + \frac{3}{8} = -\frac{69}{8}$.* ◇

For a r-encoding $w = w_1 \star w_2$ of a real number x, the finite word w_1 is called the *integer part* of w, and w_2 its *fractional part*.

It is important to notice that there exist numbers that admit distinct encodings, but with the same integer part length.

Example 2.46 *The number $\frac{1}{2}$ admits the languages of 10-encodings $0^+\star 50^\omega$ and $0^+ \star 49^\omega$.* ◇

This fact is formalized in the following two theorems : The first one deals with numbers $x \in [0, 1]$; the second one with arbitrary numbers $x \in \mathbb{R}$.

Theorem 2.47 ([Eil74]) *Let $r \in \mathbb{N}_{>1}$ be a base, and $x \in [0, 1]$ be a fractional number. The number x admits a r-encoding $w = 0 \star w'$, where*

$w' \in \{0,1,\ldots,r-1\}^\omega$. If the word w' is not unique, and if the integer part of w is fixed to 0, then x admits exactly two r-encodings

$$w_H = 0 \star w'' d0^\omega$$

and

$$w_L = 0 \star w''(d-1)(r-1)^\omega,$$

with $w'' \in \{0,1\ldots,r-1\}^*$ and $d \in \{1,2,\ldots,r-1\}$.

When dealing with real numbers with a non-zero integer part, Theorem 2.47 is directly extended into the following one.

Theorem 2.48 *Let $r \in \mathbb{N}_{>1}$ be a base, and $x \in \mathbb{R}$ be a real number. The number x admits a r-encoding $w = \sigma w' \star w''$, such that $\sigma \in \{0, r-1\}$, $w' \in \{0,1,\ldots,r-1\}^*$, $w'' \in \{0,1,\ldots,r-1\}^\omega$, and σ is not a prefix of w'. The language $\sigma^+ w' \star w''$ is a subset of the language L_x of encodings of x. If $L_x \neq \sigma^+ w' \star w''$, then either*

- *$x = 0$ and*
$$L_x = 0^+ \star 0^\omega \cup (r-1)^+ \star (r-1)^\omega, \text{ or}$$
- *$x \in \mathbb{Z}_{\neq 0}$ and*
$$L_x = \sigma'^+ w''' \sigma'' 0^n \star 0^\omega \cup \sigma'^+ w''' (\sigma''-1)(r-1)^n \star (r-1)^\omega,$$
for some $n \in \mathbb{N}$, $\sigma' \in \{0, r-1\}$, $w''' \in \{0,1,\ldots,r-1\}^*$ and $\sigma'' \in \{1,2,\ldots,r-1\}$, or
- *$x \in \mathbb{R} \setminus \mathbb{Z}$ and*
$$L_x = \sigma'^+ w''' \star w'''' \sigma'' 0^\omega \cup \sigma'^+ w''' \star w'''' (\sigma''-1)(r-1)^\omega,$$
for some $\sigma' \in \{0, r-1\}$, $w''', w'''' \in \{0,1,\ldots,r-1\}^*$ and $\sigma'' \in \{1,2,\ldots,r-1\}$.

The r-encodings that end with 0^ω are said to be *high*, and the ones ending with $(r-1)^\omega$ are said to be *low*. If, for a given length of integer parts, a number x admits two r-encodings (ending respectively with 0^ω and $(r-1)^\omega$ as a consequence of Theorem 2.48), then such encodings are said to be *dual*.

Algorithm 2 computes an arbitrarily large prefix `encodeReal`($x \in \mathbb{R}, r \in \mathbb{N}_{>1}, p \in \mathbb{N}$) of an encoding of x in base r that has $\lfloor x \rfloor$ as integer part, encoded with the smallest possible amount of digits, and $x - \lfloor x \rfloor$ as fractional

Algorithm 2 `encodeReal`$(x \in \mathbb{R}, r \in \mathbb{N}_{>1}, p \in \mathbb{N})$

$enc \leftarrow$ `encodeInteger`$(x - \lfloor x \rfloor)$
$enc \leftarrow enc \star$
$x \leftarrow x - \lfloor x \rfloor$
while $p \geq 0$ **do**
  $\sigma \leftarrow \lfloor rx \rfloor$
  $enc \leftarrow enc\ \sigma$
  $x \leftarrow rx - \lfloor rx \rfloor$
  $p \leftarrow p - 1$
end while
return enc

part. If x admits dual encodings, then the high one will be chosen. The argument p specifies the required length of the prefix of the fractional part.

This encoding system generalizes to vectors $\vec{x} \in \mathbb{R}^n$ of dimension $n \in \mathbb{N}_{>0}$. As in the case of integers, the idea is to encode each component separately into a word, in such a way that these words share the same integer-part length. One thus obtains a vector of encodings in which the separator symbol $\star$ occurs at the same position in each component. The word encoding $\vec{x}$ reads those components one symbol at a time. Since the separators $\star$ are simultaneously read, the tuple $\star^n$ can be replaced by a single symbol $\star$. One thus obtains an encoding of $\vec{x}$ as a single word $w_I \star x_F$ over the alphabet $\{0, 1, \ldots, r-1\}^n \cup \{\star\}$, as expressed in the following definition.

Definition 2.49 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $r \in \mathbb{N}_{>1}$ be a base. A word $w = \vec{\sigma_{p-1}}\vec{\sigma_{p-2}} \ldots \vec{\sigma_0} \star \vec{\sigma'_1}\vec{\sigma'_2} \ldots$, with $p \in \mathbb{N}_{>0}$, $\vec{\sigma_{p-1}} \in \{0, r-1\}^n$, $\vec{\sigma_i} \in \{0, 1, \ldots, r-1\}^n$ for each $i \in \{0, 1, \ldots, p-2\}$, and $\vec{\sigma'_i} \in \{0, 1, \ldots, r-1\}^n$ for each $i \in \mathbb{N}_{>0}$, is a* r-encoding *of a vector $\vec{x} \in \mathbb{R}^n$ if and only if, for each $j \in \{1, 2, \ldots, n\}$, we have*

$$\langle \vec{\sigma_{p-1}}[j]\vec{\sigma_{p-2}}[j] \ldots \vec{\sigma_0}[j] \star \vec{\sigma'_1}[j]\vec{\sigma'_2}[j] \ldots \rangle_r = \vec{x}[j].$$

The notation $\langle w \rangle_{r,n} = \vec{x}$ is used to mean that w is a r-encoding of $\vec{x}$. □

Example 2.50 *The word*

$$\begin{pmatrix}0\\0\\2\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix}\begin{pmatrix}0\\2\\0\end{pmatrix} \star \begin{pmatrix}0\\0\\1\end{pmatrix}\begin{pmatrix}0\\0\\0\end{pmatrix}^{\omega}$$

is a 3-encoding of the vector $(3, 2, -\frac{26}{3})$. ◇

2.4.3 Properties of encodings

Properties of encodings of real numbers have been studied for a long time. In this section, adapted from [HW85], we recall results linking reals, rationals, and periodic encodings.

The first theorem is aimed at showing that each ultimately periodic encoding, in the sense of Definition 2.11, encodes a rational number.

Theorem 2.51 *Let $r \in \mathbb{N}_{>1}$ be a base, and $w \in \{0, r-1\}\{0, 1, \ldots, r-1\}^* \star \{0, 1, \ldots, r-1\}^\omega$ be a r-encoding. If w is ultimately periodic, then $\langle w \rangle_r \in \mathbb{Q}$.*

Proof Since w is ultimately periodic, there exist $w_1 \in \{0, r-1\}\{0, 1, \ldots, r-1\}^*$, $w_2 \in \{0, 1, \ldots, r-1\}^*$ and $w_3 \in \{0, 1, \ldots, r-1\}^+$ such that $w = w_1 \star w_2 w_3^\omega$.

We successively have

$$r^{|w_2|}\langle w \rangle_r = \langle w_1 w_2 \star w_2^\omega \rangle_r,$$

$$r^{|w_2|+|w_3|}\langle w \rangle_r = \langle w_1 w_2 w_3 \star w_3^\omega \rangle_r.$$

Hence,

$$\langle w \rangle_r = \frac{\langle w_1 w_2 w_3 \star w_3^\omega \rangle_r - \langle w_1 w_2 \star w_3^\omega \rangle_r}{r^{|w_2|}(r^{|w_3|} - 1)}. \tag{2.1}$$

By Definition 2.44, we have $\langle w_1 w_2 w_3 \star w_3^\omega \rangle_r = \langle w_1 w_2 w_3 \rangle_r + \langle 0 \star w_3^\omega \rangle_r$, as well as $\langle w_1 w_2 \star w_3^\omega \rangle_r = \langle w_1 w_2 \rangle_r + \langle 0 \star w_3^\omega \rangle_r$. Equation 2.1 can then be rewritten as

$$\langle w \rangle_r = \frac{\langle w_1 w_2 w_3 \rangle_r - \langle w_1 w_2 \rangle_r}{r^{|w_2|}(r^{|w_3|} - 1)}.$$

The conclusion follows since the numerator and the denominator of this fraction are both integers. ■

The converse of Theorem 2.51 will be tackled by decomposing the problem into three cases, depending on the value of the base $r \in \mathbb{N}_{>1}$. This decomposition will give us strong properties about the respective lengths of the period of the encoding, and of the prefix leading to the first occurrence of the period.

The first case handles the situation where the prime factors of the denominator of a rational $x \in \mathbb{Q}$ are also prime factors of the base r.

Theorem 2.52 *Let $r \in \mathbb{N}_{>1}$ be a base, and let $\{q_1, q_2, \ldots, q_n\}$ be the set of prime factors of r. Suppose that $r = q_1^{i_1} q_2^{i_2} \ldots q_n^{i_n}$. Let $x \in \mathbb{Q}$ be such that*

$x = \frac{p}{q}$, $p \in \mathbb{Z}_{\neq 0}$, $q \in \mathbb{N}_{>0}$, *and* $\gcd(|p|, q) = 1$. *If there exist* $j_1, j_2, \ldots, j_n \in \mathbb{N}$ *such that* $q = q_1^{j_1} q_2^{j_2} \ldots q_n^{j_n}$, *then there exists a* r*-encoding* $w \in \{0, r-1\}\{0, 1, \ldots, r-1\}^* \star \{0, 1, \ldots, r-1\}^\omega$ *such that* w *ends with* 0^ω. *Furthermore, the smallest finite word* w_2 *such that* $w = w_1 \star w_2 0^\omega$ *has a length* $|w_2| = \max_{k \in \{1,2,\ldots,n\}} \left\lceil \frac{j_k}{i_k} \right\rceil$.

Proof This is immediate, since the number

$$m = \max_{k \in \{1,2,\ldots,n\}} \left\lceil \frac{j_k}{i_k} \right\rceil$$

is the smallest one such that $r^m x \in \mathbb{Z}$. ■

Before considering other situations, let us show that, reciprocally, each real number that admits an encoding ending with 0^ω can be expressed as a fraction $x = \frac{p}{q}$ where q is a product of factors of the base r.

Theorem 2.53 *Let* $r \in \mathbb{N}_{>1}$ *be a base, let* $S = \{q_1, q_2, \ldots, q_n\}$ *be the set of prime factors of* r, *and let* $x \in \mathbb{R}$. *If* x *admits a* r*-encoding ending with* 0^ω, *then there exist* $p \in \mathbb{Z}_{\neq 0}$, $q \in \mathbb{N}_{>0}$ *such that* $x = \frac{p}{q}$, *gcd*$(|p|, q) = 1$ *and the set the prime factors of* q *is included in* S.

Proof Since x admits an encoding ending with 0^ω, there exist finite words $w_1 \in \{0, r-1\}\{0, 1, \ldots, r-1\}^*$ and $w_2 \in \{0, 1, \ldots, r-1\}^*$ such that $x = \langle w_1 \star w_2 0^\omega \rangle_r$. It follows that $r^{|w_2|} x = \langle w_1 w_2 \rangle_r$, and thus

$$x = \frac{\langle w_1 w_2 \rangle_r}{r^{|w_2|}}.$$

The conclusion follows by reducing this fraction. ■

By combining Theorems 2.52 and 2.53, we obtain the following result.

Corollary 2.54 *Let* $r \in \mathbb{N}_{>1}$ *be a base, let* $\{q_1, q_2, \ldots, q_n\}$ *be the set of prime factors of* r, *and let* S *be the set*

$$S = \left\{ \frac{p}{q_1^{i_1} q_2^{i_2} \cdots q_n^{i_n}} \mid p \in \mathbb{Z}, i_1, i_2, \ldots, i_n \in \mathbb{N} \right\}.$$

A number $x \in \mathbb{R}$ *belongs to* S *if and only if* x *admits a* r*-encoding ending with* 0^ω.

Next, we consider rationals with denominators that are relatively prime with the base r.

The following lemma, which is a classical result of number theory, will be useful.

Lemma 2.55 ([HW85]) *Let $m, n \in \mathbb{N}$ be two relatively prime numbers. There exists $k \in \mathbb{N}_{>0}$ such that $n^k \equiv_m 1$.*

Theorem 2.56 *Let $r \in \mathbb{N}_{>1}$ be a base, and let $x \in \mathbb{Q}$ be such that $x = \dfrac{p}{q}$, $p \in \mathbb{Z}_{\neq 0}$, $q \in \mathbb{N}_{>0}$, and $\gcd(|p|, q) = 1$. If $\gcd(q, r) = 1$, then there exist an encoding $w \in \{0, r-1\}\{0, 1, \ldots, r-1\}^* \star \{0, 1, \ldots, r-1\}^\omega$ and two finite words $w_1, w_2 \neq \varepsilon$ such that $w = w_1 \star w_2^\omega$ and $\langle w \rangle_r = x$. Moreover, the smallest such w_2 cannot be equal to 0, and has a length equal to the smallest $k \in \mathbb{N}_{>0}$ such that $r^k \equiv_q 1$.*

Proof By Lemma 2.55, and since r and q are relatively prime numbers, there exists $k \in \mathbb{N}_{>0}$ such that $r^k \equiv_q 1$. Suppose that k has its smallest possible value, i.e., that for all $k' \in \mathbb{N}_{>0}$ such that $k' < k$, we have $r^{k'} \not\equiv_q 1$. It follows that there exists $m \in \mathbb{N}$ such that

$$r^k x = r^k \frac{p}{q} = (mq + 1)\frac{p}{q} = mp + \frac{p}{q}.$$

Since $mp \in \mathbb{Z}$, the fractional parts of x and $r^k x$ are identical. By the encoding scheme, we then have that the encoding of x given by Algorithm 2, called with a sufficiently large expected precision, ends with a purely periodic fractional part, with period of length k. Moreover, the period cannot be equal to 0^k, since this encoding would then end with 0^ω, which would contradict Corollary 2.54.

It remains to show that this length is the smallest possible one. By contradiction, suppose that there exist $w_1 \in \{0, r-1\}\{0, 1, \ldots, r-1\}^*$ and $w_2 \in \{0, 1, \ldots, r-1\}^+$ such that $|w_2| < k$ and

$$x = \langle w_1 \star w_2^\omega \rangle_r.$$

We also have

$$r^{|w_2|} x = \langle w_1 w_2 \star w_2^\omega \rangle_r,$$

hence

$$x = \frac{\langle w_1 w_2 \rangle_r - \langle w_1 \rangle_r}{r^{|w_2|} - 1}.$$

Since x is equal to the reduced fraction $\frac{p}{q}$, the denominator q divides $r^{|w_2|}-1$. That leads to a contradiction since k is the smallest number such that $r^k \equiv_q 1$. ■

Finally, we will consider rationals such that the sets of prime factors of the denominators of their reduced fractions contain at the same time factors of the base r and numbers that are relatively prime with r.

Theorem 2.57 *Let $r \in \mathbb{N}_{>1}$ be a base, and let $\{q_1, q_2, \ldots, q_n\}$ be the set of prime factors of r. Suppose that $r = q_1^{i_1} q_2^{i_2} \ldots q_n^{i_n}$. Let $x \in \mathbb{Q}$ be such that $x = \frac{p}{q}$, $p \in \mathbb{Z}_{\neq 0}$, $q \in \mathbb{N}_{>0}$, and $\gcd(|p|, q) = 1$. If there exist $j_1, j_2, \ldots, j_n \in \mathbb{N}$ such that*

$$\sum_{i \in \{1,2,\ldots,n\}} j_i \neq 0$$

and

$$q = q_1^{j_1} q_2^{j_2} \ldots q_n^{j_n} Q$$

with $\gcd(Q, r) = 1$, then the smallest words w_1, w_2 such that

- *$w_1 \in \{0, 1, \ldots, r-1\}^*$,*
- *$w_2 \in \{0, 1, \ldots, r-1\}^+$,*
- *there exists $w \in \{0, r-1\}\{0, 1, \ldots, r-1\}^*$, and*
- *$x = \langle w \star w_1 w_2^\omega \rangle_r$*

are such that

$$|w_1| = \max_{k \in \{1,2,\ldots,n\}} \left\lceil \frac{j_k}{i_k} \right\rceil$$

and

$$|w_2| \textit{ is the smallest } k \in \mathbb{N}_{>0} \textit{ such that } r^k \equiv_Q 1.$$

Proof Let m be defined by

$$m = \max_{k \in \{1,2,\ldots,n\}} \left\lceil \frac{j_k}{i_k} \right\rceil.$$

We have

$$r^m x = \frac{p'}{Q},$$

where $p' \in \mathbb{Z}$. By the division algorithm, the number p' can be expressed as

$$p' = p_1 Q + p_2,$$

where $p_1 \in \mathbb{Z}$, $p_2 \in \mathbb{N}$, $0 < p_2 < Q$ and $\gcd(p_2, Q) = 1$. We have

$$r^m x = p_1 + \frac{p_2}{Q}.$$

Since $p_1 \in \mathbb{Z}$ and $\gcd(p_2, Q) = 1$, then, by Theorem 2.56, there exist $w \in \{0, r-1\}\{0, 1, \ldots, r-1\}^*$ and $w' \in \{0, 1, \ldots, r-1\}^+$ such that

$$r^m x = \langle w \star (w')^\omega \rangle_r$$

and $|w'|$ is the smallest $k \in \mathbb{N}_{>0}$ such that $r^k \equiv_Q 1$. Moreover, w' has the smallest possible length.

The conclusion follows by remarking that an encoding of x can be obtained from $w \star (w')^\omega$ by shifting the separator $\star$ leftwise by m positions. ■

Theorems 2.52, 2.56 and 2.57 cover all possible rationals. A corollary is then the following.

Corollary 2.58 *Let $r \in \mathbb{N}_{>1}$ be a base, and $x \in \mathbb{Q}$ be a rational number. There exists a ultimately periodic r-encoding $w \in \{0, r-1\}\{0, 1, \ldots, r-1\}^* \star \{0, 1, \ldots, r-1\}^\omega$ such that $\langle w \rangle_r = x$.*

Finally, Theorem 2.51 and Corollary 2.58 lead to the following corollary.

Corollary 2.59 *Let $r \in \mathbb{N}_{>1}$ be a base, and let $x \in \mathbb{R}$. The number x admits a ultimately periodic r-encoding if and only if $x \in \mathbb{Q}$.*

Taking into account Definition 2.49, this result can directly be extended to the following one.

Corollary 2.60 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and let $\vec{x} \in \mathbb{R}^n$. The vector $\vec{x}$ admits a ultimately periodic r-encoding if and only if $\vec{x} \in \mathbb{Q}^n$.*

2.5 Automata recognizing sets of vectors

We have seen that vectors of integers or of reals can be encoded as words. Automata are thus able to represent languages of encodings, i.e., sets of vectors of numbers.

2.5.1 Definitions

Numbers admit infinitely many encodings. In particular, by Theorem 2.40, their sign digits can be repeated at will. Moreover, Theorem 2.48 expresses that, for some real numbers, there exist distinct encodings with the same integer part length. To represent sets of vectors by languages of words, we thus have to choose the encodings that will form those languages. A natural choice consists in keeping all encodings. With this technique, some operations on sets, like conjunctions and disjunctions, directly correspond to operations on languages accepted by automata, for which simple and efficient algorithms generally exist. Another advantage of using automata to represent sets of vectors is that some classes of automata can be minimized into a canonical form. Hence, the equality test of sets represented by automata can be efficiently implemented.

Definition 2.61 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq \mathbb{Z}^n$ be a set of vectors of integers. A* Number Decision Diagram (NDD) recognizing *S in base r is a finite automaton on finite words $\mathcal{A}$ such that*

$$L(\mathcal{A}) = \{w \in \{0, r-1\}^n(\{0,1,\ldots,r-1\}^n)^* \mid \langle w\rangle_{r,n} \in S\}.$$

The set S is said to be r-recognizable*, and $\mathcal{A}$* recognizes *S in base r.* □

Definition 2.62 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq \mathbb{R}^n$ be a set of vectors of reals. A* Real Vector Automaton (RVA) recognizing *S in base r is a Büchi automaton $\mathcal{A}$ such that*

$$L(\mathcal{A}) = \{w \in \{0, r-1\}^n(\{0,1,\ldots,r-1\}^n)^* \star (\{0,1,\ldots,r-1\}^n)^\omega \mid \langle w\rangle_{r,n} \in S\}.$$

The set S is said to be r-recognizable*, and $\mathcal{A}$* recognizes *S in base r.*

When using Muller automata instead of Büchi ones, which can be done without loss of generality by Theorem 2.26, then we obtain Muller RVA. □

2.5.2 Expressiveness

In this section, we study logical characterizations of the r-recognizable sets of vectors of integers, and of the r-recognizable sets of vectors of reals. The main results are that those sets are exactly the sets definable in base-dependent extensions of Presburger arithmetic, and of the additive theory of reals and integer, introduced in Sections 2.2.1 and 2.2.3.

The following two theorems characterize the r-recognizable sets of integer vectors. Actually, those theorems were originally proved for sets of natural numbers, but can trivially be generalized to sets of integer vectors.

Theorem 2.63 ([Büc62, Bru85, BHMV94]) *Let $n \in \mathbb{N}_{>0}$ be a dimension and $r \in \mathbb{N}_{>1}$ be a base. A set $S \subseteq \mathbb{Z}^n$ is r-recognizable if and only if it is definable in the structure*

$$\langle \mathbb{Z}, +, <, V_r \rangle,$$

where $V_r(x, u)$ is defined as the predicate that holds if and only if u is the greatest power of r that divides x if it exists, and 0 otherwise.

Theorem 2.64 ([BHMV94]) *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, $\ell \in \mathbb{N}_{>0}$, and $S \subseteq \mathbb{Z}^n$. The set S is r-recognizable if and only if it is r^ℓ-recognizable.*

The intuition behind this result is the following : A r-encoding can be obtained from a r^ℓ-encoding by simply replacing its digits by sequences of ℓ digits. The reading of a transition in a NDD recognizing a set S in base r^ℓ thus corresponds to the reading of ℓ transitions in a NDD recognizing S in base r, and conversely.

Similar results hold for sets of real vectors.

Theorem 2.65 ([BRW98]) *Let $n \in \mathbb{N}_{>0}$ be a dimension and $r \in \mathbb{N}_{>1}$ be a base. A set $S \subseteq \mathbb{R}^n$ is r-recognizable if and only if it is definable in the structure*

$$\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle,$$

where $X_r(x, u, k)$ is the predicate that holds if and only if u is an integer power of the base r, and there exists an encoding of x such that the digit at the position specified by u, i.e. at the position $\log_r(u)$ with respect to the symbol $\star$, is equal to k, where $k \in \{0, 1, \ldots, r-1\}$.

The idea behind the proof of Theorem 2.65, presented in [BRW98], is that, given a RVA $\mathcal{A}$, each of its states can be encoded by a symbol in a finite alphabet. Hence, each execution of $\mathcal{A}$ can be represented by an infinite word, which can be seen as an encoding of a vector of reals. One then shows that it is possible to build a formula of $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$ that specifies that a given execution accepts an encoding of a real number.

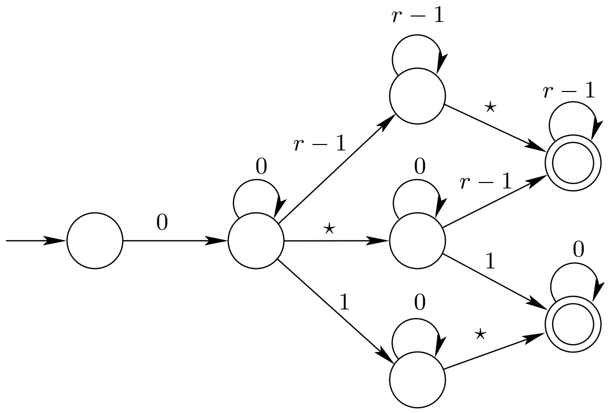

Figure 2.4: RVA representing the set of integer powers of a base $r \in \mathbb{N}_{>1}$.

Let $V_r(x, u)$ be the generalization to $\mathbb{R}^2$ of the predicate defined in Theorem 2.63. It has been shown [Bru06] that the structure $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$ is *equivalent* to the structure $\langle \mathbb{R}, \mathbb{Z}, +, <, V_r \rangle$, in the sense that each formula definable in $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$ is definable in $\langle \mathbb{R}, \mathbb{Z}, +, <, V_r \rangle$, and reciprocally.

The theories $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$ and $\langle \mathbb{R}, \mathbb{Z}, +, <, V_r \rangle$ are strictly more expressive than the additive theory of reals and integers $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. Indeed, in particular, they make it possible to define sets including non-linear periodicities. For instance, the set P_r of integer powers of the base r is defined in $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$ by the formula

$$\{x \in \mathbb{R} \mid X_r(0, x, 0)\}.$$

Alternatively, P_r is recognized in base r by the automaton depicted in Figure 2.4.

The following theorem states that the expressiveness of RVA is not affected when the base r is replaced by one of its powers r^ℓ, with $\ell \in \mathbb{N}_{>0}$. Its proof establishes that the predicate $X_r(x, u, k)$ can be expressed in terms of $X_{r^\ell}(x, u, k)$, and reciprocally. Indeed, intuitively, testing the value of the digit at a given position in an encoding in base r^ℓ can be reduced to the test of ℓ digits in base r, and conversely.

Theorem 2.66 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, $\ell \in \mathbb{N}_{>0}$, and $S \subseteq \mathbb{R}^n$. The set S is r-recognizable if and only if it is r^ℓ-recognizable.*

Proof By Theorem 2.65, it suffices to establish that the structures

$$\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$$

and

$$\langle \mathbb{R}, \mathbb{Z}, +, <, X_{r^\ell} \rangle$$

are equivalent. This can be achieved by showing that the sets $\{(x,u,k) \in \mathbb{R}^3 \mid X_{r^\ell}(x,u,k)\}$ and $\{(x,u,k) \in \mathbb{R}^3 \mid X_r(x,u,k)\}$ are respectively definable in the structures $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$ and $\langle \mathbb{R}, \mathbb{Z}, +, <, X_{r^\ell} \rangle$.

First, in order to express the predicate $X_r(x,u,k)$ in terms of $X_{r^\ell}(x,u,k)$, one needs to consider the possible positions of u with respect to the greatest power of r^ℓ lower than or equal to u, since u is not necessarily a power of r^ℓ. Formally,

$$\{(x,u,k) \in \mathbb{R}^3 \mid X_r(x,u,k)\} =$$
$$\Bigg\{ (x,u,k) \in \mathbb{R}^3 \mid k \in \{0,1,\ldots,r-1\} \wedge$$
$$\bigvee_{k_0,k_1,\ldots,k_{\ell-1} \in \{0,1,\ldots,r-1\}} \ \bigvee_{i \in \{0,1,\ldots,\ell-1\}} X_{r^\ell}\left(x, \frac{u}{r^i}, kr^i + \sum_{j \in \{0,1,\ldots,\ell-1\} \setminus \{i\}} k_j r^j \right) \Bigg\}.$$

Expressing $X_{r^\ell}(x,u,k)$ in terms of $X_r(x,u,k)$ is more tricky, due to dual encodings, and requires to define three additional auxiliary predicates.

- $G_r(x,u)$ is true if and only if $x \neq 0$, x admits dual encodings, and u is the greatest integer power of r such that there exist two encodings of w in which the digits at the position specified by u are different. It can be defined by

$$\{(x,u) \in \mathbb{R}^2 \mid G_r(x,u)\} =$$
$$\{(x,u) \in \mathbb{R}^2 \mid$$
$$\bigvee_{k \in \{0,1,\ldots,r-1\}} \ \bigvee_{l \in \{0,1,\ldots,r-1\}} (k \neq l \wedge X_r(x,u,k) \wedge X_r(x,u,l) \wedge (\forall u' \in \mathbb{R})$$
$$\bigwedge_{k' \in \{0,1,\ldots,r-1\}} \ \bigwedge_{l' \in \{0,1,\ldots,r-1\}} ((u' > u \wedge X_r(x,u',k') \wedge X_r(x,u',l')) \Rightarrow$$
$$k' = l'))\}.$$

- $X_r^H(x,u,k)$ is true if and only if $X_r(x,u,k)$ is true and either if x does not admit dual encodings, or if k is the digit at the position specified

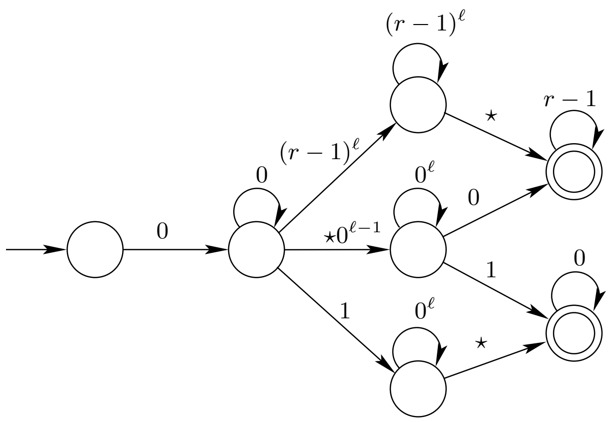

Figure 2.5: RVA representing in a base $r \in \mathbb{N}_{>1}$ the set of integer powers of r^ℓ, with $\ell \in \mathbb{N}_{>0}$.

by u of a high encoding of x. It can be defined by

$$\begin{gathered}\{(x,u,k) \in \mathbb{R}^3 \mid X_r^H(x,u,k)\} = \\ \{(x,u,k) \in \mathbb{R}^3 \mid \\ X_r(x,u,k) \wedge ((x = 0 \wedge k = 0)\vee \\ (x \neq 0 \wedge (\forall u \in \mathbb{R})(\neg G_r(x,u)))\vee \\ (\exists u' \in \mathbb{R})(G_r(x,u')\wedge \\ ((u > u') \vee (u = u' \wedge X_r(x,u,k-1)) \vee (u < u' \wedge k = 0))))\}.\end{gathered}$$

- $X_r^L(x,u,k)$ is true if and only if $X_r(x,u,k)$ is true and either if x does not admit dual encodings, or if k is the digit at the position specified by u of a low encoding of x. It can be defined by

$$\begin{gathered}\{(x,u,k) \in \mathbb{R}^3 \mid X_r^L(x,u,k)\} = \\ \{(x,u,k) \in \mathbb{R}^3 \mid \\ X_r(x,u,k) \wedge ((x = 0 \wedge k = r-1)\vee \\ (x \neq 0 \wedge (\forall u \in \mathbb{R})(\neg G_r(x,u)))\vee \\ (\exists u' \in \mathbb{R})(G_r(x,u')\wedge \\ ((u > u') \vee (u = u' \wedge X_r(x,u,k+1)) \vee (u < u' \wedge k = r-1))))\}.\end{gathered}$$

- $P_{r^\ell}(u)$ is true if and only if u is an integer power of r^ℓ. By Theorem 2.65, it is definable in $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r\rangle$ since it is recognized by the RVA given in Figure 2.5.

In the definitions of $X_r^H(x,u,k)$ and $X_r^L(x,u,k)$, the number 0 is handled as a particular case, since it is the only number x that admits dual encoding but is such that $G_r(x,u)$ is false for each possible value of u.

For $k \in \{0,1,\ldots,r^\ell - 1\}$, let $k_{\ell-1}k_{\ell-2}\ldots k_0$ denote its r-encoding of size ℓ (with $k_{\ell-1}, k_{\ell-2}, \ldots, k_0 \in \{0,1,\ldots,r-1\}$).

The predicate $X_{r^\ell}(x,u,k)$ can now be expressed in terms of X_r :

$$\Big\{(x,u,k) \in \mathbb{R}^3 \mid X_r^\ell(x,u,k)\} = \{(x,u,k) \in \mathbb{R}^3 \mid \bigvee_{k' \in \{0,1\ldots,r^\ell-1\}} \Big(k = k' \wedge P_{r^\ell}(u) \wedge \Big(\bigwedge_{i\in\{0,1,\ldots,\ell-1\}} X_r^H(x, r^i u, k_i) \vee \bigwedge_{i\in\{0,1,\ldots,\ell-1\}} X_r^L(x, r^i u, k_i)\Big)\Big)\Big\}.$$

■

Using the following definition, Theorem 2.66 admits an important corollary.

Definition 2.67 *Two numbers $r, s \in \mathbb{N}_{>1}$ are* multiplicatively independent *if and only if $r^p \neq s^q$ for all $p, q \in \mathbb{N}_{>0}$.* □

Corollary 2.68 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be multiplicatively dependent bases, and $S \subseteq \mathbb{R}^n$. The set S is r-recognizable if and only if it is s-recognizable.*

Proof By Definition 2.67, there exist $p, q \in \mathbb{N}_{>0}$ such that $r^p = s^q$.

By Theorem 2.66, S is r-recognizable if and only if it is r^p-recognizable. Using a second time this theorem, and since $r^p = s^q$, this condition holds if and only if S is s-recognizable. ■

To conclude this section, the following theorem states that if two r-recognizable subsets of $\mathbb{R}^n$ coincide over $\mathbb{Q}^n$, then they are equal.

Theorem 2.69 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $S_1, S_2 \subseteq \mathbb{R}^n$ be two r-recognizable sets. If $S_1 \cap \mathbb{Q}^n = S_2 \cap \mathbb{Q}^n$, then $S_1 = S_2$.*

Proof This is an immediate consequence of Theorem 2.23 and Corollary 2.60. ■

2.5.3 Algorithms

The algorithms cited in Section 2.3.3 can be adapted to perform set operations on RVA. The only differences rely in the fact that the automata obtained from those operations have to accept all the encodings of the vectors they represent.

Let $n_1, n_2 \in \mathbb{N}_{>0}$ be dimensions, $r \in \mathbb{N}_{>1}$ be a base, and $\mathcal{A}_1, \mathcal{A}_2$ be RVA respectively representing sets $S_1 \subseteq \mathbb{R}^{n_1}$ and $S_2 \subseteq \mathbb{R}^{n_2}$.

If $n_1 = n_2$, then RVA representing the union $S_1 \cup S_2$, the intersection $S_1 \cap S_2$ and the difference $S_1 \setminus S_2$ can be obtained by computing respectively the union, the intersection and the difference of the Büchi automata $\mathcal{A}_1$ and $\mathcal{A}_2$, without any adaptation of the algorithms.

To build an automaton representing the Cartesian product $S_1 \times S_2$, the algorithm given in Section 2.3.3 has to map the transitions labeled by $(\star, \star)$ into one transition labeled by $\star$. Transitions labeled by $(\sigma, \star)$, or by $(\star, \sigma')$, where $\sigma \in \{0, 1, \ldots, r-1\}^{n_1}$ and $\sigma' \in \{0, 1, \ldots, r-1\}^{n_2}$ have to be deleted.

The projection of a set of vectors $S \subseteq \mathbb{R}^n$ on each of their components except the ith ($i \in \{1, 2, \ldots, n\}$) is defined by

$$S|_{\neq i} = \{(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \mathbb{R}^{n-1} \mid (\exists x_i \in \mathbb{R})((x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) \in S)\}.$$

By using the algorithm on languages, and by defining that the projection of the symbol $\star$ is itself, one obtains a language that may not accept all the encodings of the vectors it represents.

Example 2.70 *The language of* 2*-encodings of the set* $S = \{(0, 2)\} \subset \mathbb{R}^2$ *is*

$$(0,0)^+(0,1)(0,0) \star (0,0)^\omega \cup (0,0)^+(0,1) \star (0,1)^\omega \cup (1,0)^+(1,1)(1,0) \star (1,0)^\omega \cup (1,0)^+(1,1) \star (1,1)^\omega,$$

and the projection of this language that eliminates the second component is equal to

$$0^+00 \star 0^\omega \cup 0^+0 \star 0^\omega \cup 1^+11 \star 1^\omega \cup 1^+1 \star 1^\omega = 0^+0 \star 0^\omega \cup 1^+1 \star 1^\omega,$$

which does not contain the encodings $0 \star 0^\omega$ *and* $1 \star 1^\omega$*.* ◇

A solution exists for tackling this problem, and consists in considering all vectors of digits $\vec{a} \in \{0, r-1\}^n$ and in detecting, with an exploration that begins at the initial state of the automaton, the states that are reachable by the reading of words $\vec{a}^i$, with $i \in \mathbb{N}_{>0}$. Then, for each reached state s, one adds the transition $(q_0, \vec{a}, s)$, where q_0 is the initial state of the automaton [Boi98]. When using serialized encodings, there exists a more efficient algorithm [BL04].

To conclude this section, let us mention that there exists an efficient way, given a dimension $n \in \mathbb{N}_{>0}$, a vector $\vec{a} \in \mathbb{Z}^n$ and a number $b \in \mathbb{Z}$, for building a RVA representing the set $\{\vec{x} \in \mathbb{R}^n \mid \vec{a}\vec{x} \leq b\}$. This algorithm is presented in [BRW98].

2.5.4 Cobham's and Semenov's theorems

This section gives the characterization, due to Cobham for the one-dimensional case and to Semenov for the multi-dimensional case, of the sets of integer vectors that are simultaneously recognizable in several bases. This characterization was originally developed for sets of non-negative integer vectors, but can be directly generalized to integer vectors.

Theorem 2.71 ([Cob69]) *Let $r, s \in \mathbb{N}_{>1}$ be multiplicatively independent bases, and let $S \subseteq \mathbb{N}$ (resp. $S \subseteq \mathbb{Z}$). The set S is simultaneously r- and s-recognizable if and only if S is definable in $\langle \mathbb{N}, + \rangle$ (resp. in $\langle \mathbb{Z}, +, < \rangle$).*

Theorem 2.72 ([Sem77]) *Let $r, s \in \mathbb{N}_{>1}$ be multiplicatively independent bases, and let $S \subseteq \mathbb{N}^n$ (resp. $S \subseteq \mathbb{Z}^n$). The set S is simultaneously r- and s-recognizable if and only if S is definable in $\langle \mathbb{N}, + \rangle$ (resp. in $\langle \mathbb{Z}, +, < \rangle$).*

The possible subsets $S \subseteq \mathbb{Z}^n$ can thus be classified into three distinct groups. A set S is either

- not r-recognizable for any $r \in \mathbb{N}_{>1}$,
- r- and s-recognizable for multiplicatively independent bases $r, s \in \mathbb{Z}_{>1}$. By Theorems 2.71, 2.72 and 2.63, such sets are recognizable in every base $r \in \mathbb{N}_{>1}$,
- or recognizable only in multiplicatively dependent bases. Indeed, a corollary of Theorem 2.64 states that, given two bases $r, s \in \mathbb{N}_{>1}$ that are multiplicatively dependent, a set $S \subseteq \mathbb{Z}^n$ is r-recognizable if and only if it is s-recognizable [BHMV94].

Chapter 3

Weak deterministic automata

The first objective of this chapter is to present weak deterministic automata, which is a particular subclass of Büchi automata, for which there exist efficient manipulation algorithms. The languages accepted by those automata admit a topological characterization that can be connected to the topology of the sets that are represented. Some links between weak deterministic automata and topology of sets of vectors were established in [BJW05]. As original contributions, we prove additional links, needed in the next chapters, and we demonstrate closure properties of the sets recognizable by weak deterministic automata.

Before introducing weak deterministic automata, we begin by giving some notions of topology.

3.1 Topologies of ω-words and real vectors

In this section, we recall some notions about topology, which is a useful tool for reasoning about the properties of sets of words and numbers [PP04].

3.1.1 General concepts

Given a set S, either of words or of numbers, a distance $d(x, y)$ defined on this set induces a metric topology on the subsets of S.

Definition 3.1 *Let S be a set. A function $d : S^2 \to \mathbb{R}_{\geq 0}$ is a* distance *on S if the three following conditions hold :*

- $(\forall x, y \in S)(d(x, y) = 0 \Leftrightarrow x = y)$;
- $(\forall x, y \in S)(d(x, y) = d(y, x))$;
- $(\forall x, y, z \in S)(d(x, z) \leq d(x, y) + d(y, z))$.

□

Definition 3.2 *Let S be a set, and d be a distance on S. A* neighborhood *of a point $x \in S$ with respect to $\varepsilon \in \mathbb{R}_{>0}$ is the set*

$$N_\varepsilon(x) = \{y \in S \mid d(x, y) < \varepsilon\}.$$

□

Definition 3.3 *Let S be a set, and d be a distance on S. A set $S' \subseteq S$ is said to be* open *with respect to d if for all $x \in S'$, there exists $\varepsilon \in \mathbb{R}_{>0}$ such that $N_\varepsilon(x) \subseteq S'$.* □

Definition 3.4 *Let S be a set, and d be a distance on S. A set $S' \subseteq S$ is said to be* closed *with respect to d if its complement with respect to S, i.e., $S \setminus S'$, is open.* □

Theorem 3.5 *Let S be a set, and d be a distance on S. A set $S' \subseteq S$ is closed with respect to d if and only if S' contains the limits of all its converging sequences of elements.*

Given a domain and a distance defined on this domain, the following notations will be used :

- F is the class of closed sets;
- G is the class of open sets;
- F_σ is the class of countable unions of closed sets;
- G_δ is the class of countable intersections of open sets.

Other classes can be defined from these notations : The class $\mathcal{B}(F) = \mathcal{B}(G)$ contains the finite Boolean combinations of open and closed sets, whereas $F_\sigma \cap G_\delta$ is the class of sets that can be expressed as countable unions of closed sets as well as countable intersections of open sets.

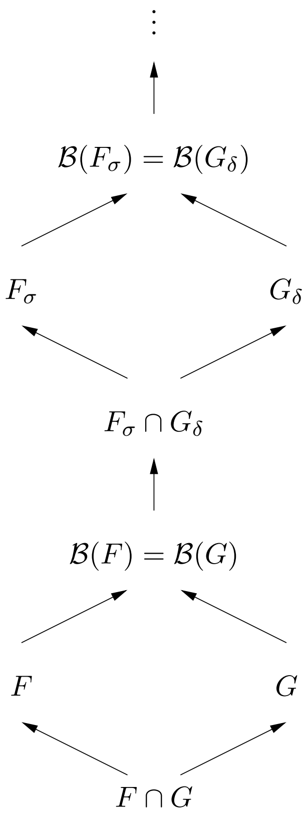

Figure 3.1: First levels of the Borel hierarchy in a metric topology.

Those classes of sets form the first levels of the Borel hierarchy. In a metric topology, this hierarchy states that F and G are subclasses of $\mathcal{B}(F) = \mathcal{B}(G)$, which is itself a subclass of $F_\sigma \cap G_\delta$. The first levels of the Borel hierarchy are schematized in Figure 3.1, taken from [BJW05], where an arrow between classes indicates proper inclusion.

The following two theorems will be useful.

Theorem 3.6 *Let S_1, S_2 be two sets. If S_1, S_2 both belong to G_δ, then $S_1 \cup S_2$ belongs to G_δ.*

Proof The sets S_1 and S_2 are both in the topological class G_δ. That means that there exist open sets $G_0, G_1, G_2, \ldots$ and $G'_0, G'_1, G'_2, \ldots$ such that

$$S_1 = \bigcap_{i \in \mathbb{N}} G_i$$

and

$$S_2 = \bigcap_{i\in\mathbb{N}} G'_i.$$

Without loss of generality, those two countable intersections can be supposed decreasing, i.e., for each $i \in \mathbb{N}$, $G_{i+1} \subseteq G_i$. Indeed, one can, if necessary, take intersections, because the finite intersections of open sets are open.

Under this assumption, let us prove that

$$S_1 \cup S_2 = \bigcap_{i\in\mathbb{N}} (G_i \cup G'_i),$$

which will be sufficient since the unions of open sets are open.

Let $x \in S_1 \cup S_2$. One either has $x \in G_i$ for each $i \in \mathbb{N}$, or $x \in G'_i$ for each $i \in \mathbb{N}$. Hence, $x \in \bigcap_{i\in\mathbb{N}}(G_i \cup G'_i)$.

On the other hand, let $x \in \bigcap_{i\in\mathbb{N}}(G_i \cup G'_i)$. For an infinite subset $N \subseteq \mathbb{N}$, we have $x \in G_n$ for all $n \in N$, or $x \in G'_n$ for all $n \in N$. Since the intersections $\bigcap_{i\in\mathbb{N}} G_i$ and $\bigcap_{i\in\mathbb{N}} G'_i$ are decreasing, belonging to an infinite number of G_i (or G'_i) is equivalent to belonging to all of them. It follows that $x \in S_1$ or $x \in S_2$. ■

Theorem 3.7 *Let S be a set, and S_1, S_2 be subsets of S. If S_1 and S_2 both belong to F_σ, then $S_1 \cap S_2$ belongs to F_σ.*

Proof The sets S_1 and S_2 both belong to F_σ. By duality, their complements $S \setminus S_1$ and $S \setminus S_2$ are in G_δ. By Theorem 3.6, the union $(S \setminus S_1) \cup (S \setminus S_2)$ is also in G_δ. Hence, by duality, the set $S_1 \cap S_2$ belongs to F_σ. ■

3.1.2 Topology of infinite words

Definition 3.8 *Let Σ be an alphabet. The* distance between infinite words *$w, w' \in \Sigma^\omega$ is defined by*

$$d(w, w') = \begin{cases} \dfrac{1}{|\,\mathrm{common}(w, w')| + 1} & \text{if } w \neq w' \\ 0 & \text{if } w = w', \end{cases}$$

where $|\,\mathrm{common}(w, w')|$ denotes the length of the longest common prefix of w and w'. □

This distance induces a topology on Σ^ω.

Example 3.9 *On the alphabet* $\{0, 1, \ldots, 9, \star\}$*, we have*

- $d(0438 \star 394(23)^\omega, 93 \star 290^\omega) = 1$;
- $d(0 \star (01)^\omega, 0 \star (10)^\omega) = \dfrac{1}{3}$.

◇

The notion of *dense oscillating sequence property* will be useful in the sequel.

Definition 3.10 ([BJW05]) *Let* Σ *be an alphabet, and* $L \subseteq \Sigma^\omega$ *be an* ω*-language. The language* L *satisfies the* dense oscillating sequence property *if,*

$$w_1, w_2, w_3, \ldots \in \Sigma^\omega$$

being infinite words, and

$$\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots \in \mathbb{R}_{>0}$$

being positive values, one has that

$$\exists w_1 \forall \varepsilon_1 \exists w_2 \forall \varepsilon_2 \exists w_3 \forall \varepsilon_3 \ldots$$

such that $d(w_i, w_{i+1}) \leq \varepsilon_i$ *for all* $i \in \mathbb{N}_{>0}$*,* $w_i \in L$ *for all odd* i*, and* $w_i \notin L$ *for all even* i*.* □

3.1.3 Topology of real vectors

The classical Euclidean distance will be used as distance between real vectors.

Definition 3.11 *Let* $n \in \mathbb{N}_{>0}$ *be a dimension. The* (Euclidean) distance between vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ *is defined by*

$$d(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^{n} (\vec{x}[i] - \vec{y}[i])^2}.$$

□

In the topology of the sets of real vectors induced by the Euclidean distance, a notion of dense oscillating property can be defined in the same way as for infinite words.

Definition 3.12 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $S \subseteq \mathbb{R}^n$ be a set of real vectors. The set S satisfies the* dense oscillating sequence property *if,*

$$\vec{x_1}, \vec{x_2}, \vec{x_3}, \ldots \in \mathbb{R}^n$$

being real vectors and

$$\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots \in \mathbb{R}_{>0}$$

being positive values, one has that

$$\exists \vec{x_1} \forall \varepsilon_1 \exists \vec{x_2} \forall \varepsilon_2 \exists \vec{x_3} \forall \varepsilon_3 \ldots$$

such that $d(\vec{x_i}, \vec{x_{i+1}}) \leq \varepsilon_i$ for all $i \in \mathbb{N}_{>0}$, $\vec{x_i} \in S$ for all odd i, and $\vec{x_i} \notin S$ for all even i. □

The following theorems give closure properties of the class $F_\sigma \cap G_\delta$ of the topology of real vectors.

Theorem 3.13 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $S_1, S_2 \subseteq \mathbb{R}^n$. If the sets S_1 and S_2 both belong to $F_\sigma \cap G_\delta$, then $S_1 \cup S_2$ belongs to $F_\sigma \cap G_\delta$ as well.*

Proof First, the sets S_1 and S_2 both belong to F_σ. Their union is thus a union of two countable unions of closed sets, which is itself a countable union of closed sets, i.e., a F_σ set. Next, the sets S_1 and S_2 both belong to G_δ. By Theorem 3.6, their union is a G_δ set. ■

Theorem 3.14 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $S_1, S_2 \subseteq \mathbb{R}^n$. If the sets S_1 and S_2 both belong to $F_\sigma \cap G_\delta$, then $S_1 \cap S_2$ belongs to $F_\sigma \cap G_\delta$ as well.*

Proof First, the sets S_1 and S_2 both belong to F_σ. By Theorem 3.7, their intersection is a F_σ set. Next, the sets S_1 and S_2 both belong to G_δ. Their intersection is thus an intersection of two countable intersections of open sets, which is itself a countable intersection of open sets, i.e., a G_δ set. ■

Theorem 3.15 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $S \in \mathbb{R}^n$. If the set S belongs to $F_\sigma \cap G_\delta$, then its complement $\mathbb{R}^n \setminus S$ belongs to $F_\sigma \cap G_\delta$ as well.*

Proof The set S belongs to F_σ (resp. G_δ). It follows by duality that its complement $\mathbb{R}^n \setminus S$ belongs to G_δ (resp. F_σ). ■

Theorem 3.16 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $S_1, S_2 \subseteq \mathbb{R}^n$. If the sets S_1 and S_2 both belong to $F_\sigma \cap G_\delta$, then $S_1 \setminus S_2$ belongs to $F_\sigma \cap G_\delta$ as well.*

Proof Since $S_1 \setminus S_2 = S_1 \cap (\mathbb{R}^n \setminus S_2)$, this theorem directly follows from Theorems 3.14 and 3.15. ■

Theorem 3.17 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $S \subseteq \mathbb{R}^n$, $C \in \mathbb{Q}^{n \times n}$ such that $det(C) \neq 0$, and $\vec{a} \in \mathbb{Q}^n$. If S belongs to $F_\sigma \cap G_\delta$, then the set $CS + \vec{a}$ belongs to $F_\sigma \cap G_\delta$ as well.*

Proof Since C is non-singular, the operation $\vec{x} \mapsto C\vec{x} + \vec{a}$ is bijective. In addition, it is continuous, as well as its inverse. Hence, it preserves open and closed sets [Bou71].

On the one hand, $S \in F_\sigma$, i.e, $S = \bigcup_{i \in \mathbb{N}} F_i$, where F_i is closed for each $i \in \mathbb{N}$. It follows that $CS + \vec{a} = \bigcup_{i \in \mathbb{N}} (CF_i + \vec{a})$ is also a F_σ set.

On the other hand, $S \in G_\delta$. By duality, its complement $\mathbb{R}^n \setminus S$ is in F_σ. Thus, $C(\mathbb{R}^n \setminus S) + \vec{a} \in F_\sigma$, and $CS + \vec{a} \in G_\delta$. ■

Theorem 3.18 *Let $n, m \in \mathbb{N}_{>0}$ be dimensions, and $S \subseteq \mathbb{R}^n$. If the set S is open (resp. closed), then the set $S \times \mathbb{R}^m$ is open (resp. closed) as well.*

Proof Suppose that S is open. Let $\vec{x} \in S \times \mathbb{R}^m$. By definition, there exist $\vec{y} \in S$ and $\vec{z} \in \mathbb{R}^m$ such that $\vec{x} = (\vec{y}, \vec{z})$. Since S is open, there exists $\varepsilon \in \mathbb{R}_{>0}$ such that $N_\varepsilon(\vec{y}) \subseteq S$. The neighborhood $N_\varepsilon(\vec{x})$ of $\vec{x}$ is such that $N_\varepsilon(\vec{x}) \subseteq S \times \mathbb{R}^m$. Hence, the set $S \times \mathbb{R}^m$ is open.

Now, suppose that S is closed, i.e., that its complement $\mathbb{R}^n \setminus S$ is open. We have $S \times \mathbb{R}^m = \mathbb{R}^{n+m} \setminus ((\mathbb{R}^n \setminus S) \times \mathbb{R}^m)$. Since $\mathbb{R}^n \setminus S$ is open, and by the first part of the proof, the set $S \times \mathbb{R}^m$ is the complement of an open set, i.e., a closed set. ■

Theorem 3.19 *Let $n, m \in \mathbb{N}_{>0}$ be dimensions, $S_1 \subseteq \mathbb{R}^n$, and $S_2 \subseteq \mathbb{R}^m$. If the sets S_1 and S_2 both belong to $F_\sigma \cap G_\delta$, then the Cartesian product $S_1 \times S_2$ belongs to $F_\sigma \cap G_\delta$ as well.*

Proof Since the sets S_1 and S_2 both belong to F_σ and to G_δ, there exist closed sets $F_0, F'_0, F_1, F'_1, F_2, F'_2, \ldots$ and open sets $G_0, G'_0, G_1, G'_1, G_2, G'_2, \ldots$ such that

$$S_1 = \bigcup_{i \in \mathbb{N}} F_i = \bigcap_{i \in \mathbb{N}} G_i$$

and

$$S_2 = \bigcup_{i \in \mathbb{N}} F'_i = \bigcap_{i \in \mathbb{N}} G'_i.$$

By Theorem 3.18, the sets $F_i \times \mathbb{R}^m$ and $F'_i \times \mathbb{R}^m$ are closed for each $i \in \mathbb{N}$. Similarly, $G_i \times \mathbb{R}^m$ and $G'_i \times \mathbb{R}^n$ are open for each $i \in \mathbb{N}$. It follows that

$$S_1 \times \mathbb{R}^m = \bigcup_{i \in \mathbb{N}} (F_i \times \mathbb{R}^m) = \bigcap_{i \in \mathbb{N}} (G_i \times \mathbb{R}^m)$$

and

$$\mathbb{R}^n \times S_2 = \bigcup_{i \in \mathbb{N}} (\mathbb{R}^n \times F'_i) = \bigcap_{i \in \mathbb{N}} (\mathbb{R}^n \times G'_i)$$

are two sets that belong to the class $F_\sigma \cap G_\delta$.

Since $S_1 \times S_2 = (S_1 \times \mathbb{R}^m) \cap (\mathbb{R}^n \times S_2)$, the conclusion follows by Theorem 3.14. ■

Although projection does not generally preserve $F_\sigma \cap G_\delta$ sets, it is sometimes possible to extract a $F_\sigma \cap G_\delta$ set from a set of larger dimension. This operation is described by the following theorem.

Theorem 3.20 *Let $n, m \in \mathbb{N}_{>0}$ be dimensions, $\vec{a} \in \mathbb{R}^m$ and $S \subseteq \mathbb{R}^n$. If the set $S \times \{\vec{a}\}$ belongs to $F_\sigma \cap G_\delta$, then S belongs to $F_\sigma \cap G_\delta$ as well.*

Proof Let f be the projection function $f : S \times \{\vec{a}\} \to S : (\vec{x}, \vec{a}) \mapsto \vec{x}$.

On the one hand, $S \times \{\vec{a}\} \in F_\sigma$, i.e., $S \times \{\vec{a}\} = \bigcup_{i \in \mathbb{N}} F_i$, where F_i is closed for each $i \in \mathbb{N}$. We have $S = \bigcup_{i \in \mathbb{N}} f(F_i)$. It is proved in [Mun75] that if Y is a compact space, then the projection $X \times Y \mapsto Y$ preserves closed sets. Since $\{\vec{a}\}$ is compact, the sets $f(F_i)$ are closed. Hence, S is F_σ.

On the other hand, $S \times \{\vec{a}\} \in G_\delta$, i.e., $S \times \{\vec{a}\} = \bigcap_{i \in \mathbb{N}} G_i$, where G_i is open for each $i \in \mathbb{N}$. For $\vec{x} \in \mathbb{R}^n$, we successively have

$$\begin{aligned} & x \in f(S \times \{\vec{a}\}) \\ \Leftrightarrow \quad & (\vec{x}, \vec{a}) \in S \times \{\vec{a}\} \\ \Leftrightarrow \quad & (\vec{x}, \vec{a}) \in \bigcap_{i \in \mathbb{N}} G_i \\ \Leftrightarrow \quad & \vec{x} \in \bigcap_{i \in \mathbb{N}} f(G_i). \end{aligned}$$

Hence, $S = f(S \times \{\vec{a}\}) = \bigcap_{i \in \mathbb{N}} f(G_i)$. Since the projections of open sets are always open [Bou71], the set S is G_δ. ■

3.1.4 Links between topologies

In this section, links between the topologies of real vectors and infinite words will be emphasized. The notations F_σ, G_δ and $F_\sigma \cap G_\delta$ (respectively F_σ, G_δ and $\mathsf{F}_\sigma \cap \mathsf{G}_\delta$) will be used when dealing with the topology of infinite words (respectively with the topology of real vectors).

Theorem 3.21 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $L \subseteq (\{0, 1, \ldots, r-1\}^n \cup \{\star\})^\omega$ be a language. If L belongs to F_σ, then the set of real vectors that have at least one r-encoding in L belongs to F_σ.*

Proof For $i_1 \in \mathbb{N}$, let W_{i_1} be the language of r-encodings with integer parts composed of a tuple of sign digits followed by i_1 tuples of digits, i.e.,

$$W_{i_1} = \{0, r-1\}^n (\{0, 1, \ldots, r-1\}^n)^{i_1} \star (\{0, 1, \ldots, r-1\}^n \cup \{\star\})^\omega.$$

By Definitions 3.3 and 3.8, the sets W_{i_1} are open for all $i_1 \in \mathbb{N}$.

Since L belongs to F_σ, it is a countable union of closed sets and can be expressed as

$$L = \bigcup_{i_2 \in \mathbb{N}} F_{i_2},$$

where each F_{i_2} is closed.

The language L may contain non-valid encodings. However, the language

$$\bigcup_{i_1 \in \mathbb{N}} \bigcup_{i_2 \in \mathbb{N}} (W_{i_1} \cap F_{i_2})$$

is a sublanguage of L such that the language of valid encodings it contains is exactly the language of valid encodings that belong to L. When i_1 and i_2

are fixed, the set $W_{i_1} \cap F_{i_2}$ is the intersection of a closed and an open set; hence, the Borel hierarchy (see Figure 3.1, page 43) states that it belongs to F_σ and is thus a countable union of closed sets:

$$W_{i_1} \cap F_{i_2} = \bigcup_{i_3 \in \mathbb{N}} L_{i_1,i_2,i_3}.$$

For each of these closed sets L_{i_1,i_2,i_3}, let us define $S_{i_1,i_2,i_3} \subseteq \mathbb{R}^n$ as the set of vectors that have at least one r-encoding in L_{i_1,i_2,i_3}.

The set S_{i_1,i_2,i_3} is closed for every $i_1, i_2, i_3 \in \mathbb{N}$. Indeed, suppose by contradiction that S_{i_1,i_2,i_3} is not closed for a given triplet $(i_1, i_2, i_3) \in \mathbb{N}^3$. By Theorem 3.5, there exists a converging sequence

$$(\vec{v_k})_{k \in \mathbb{N}}$$

of points $\vec{v_k} \in S_{i_1,i_2,i_3}$ whose limit $\vec{v}$ does not belong to S_{i_1,i_2,i_3}. If at least one component $\vec{v}[i]$ (with $i \in \{1, 2, \ldots, n\}$) admits dual r-encodings, then the corresponding sequence $(\vec{v_k}[i])_{k \in \mathbb{N}}$ may approach $\vec{v}[i]$ by alternating between smaller and greater points than $\vec{v}[i]$, which would imply that the sequence of words encoding it would not converge. We will bypass this problem : Since the sequence $(\vec{v_k})_{k \in \mathbb{N}}$ is infinite, and since it does not contain its limit $\vec{v}$, there exists at least a vector $\vec{a} \in \{-1, 1\}^n$ such that the set

$$(\vec{a}\{\vec{v_k} - \vec{v} \mid k \in \mathbb{N}\}) \cap (\mathbb{R}_{>0})^n$$

is infinite. Let us consider such a vector $\vec{a}$, and let

$$J = \{j \in \mathbb{N} \mid \vec{a}(\vec{v_j} - \vec{x}) \in (\mathbb{R}_{>0})^n\}.$$

The sequence $(\vec{v_j})_{j \in J}$ is a converging infinite subsequence of $(\vec{v_k})_{k \in \mathbb{N}}$ that approaches $\vec{v}$ following the combination of sign digits defined by $\vec{a}$. Each of the vectors of S_{i_1,i_2,i_3} has at least one r-encoding in $W_{i_1} \cap F_{i_2}$. Since the valid encodings in $W_{i_1} \cap F_{i_2}$ have the same integer part length, the sequence $(\vec{v_j})_{j \in J}$ is mapped to a converging sequence of words encoding those points. Since L_{i_1,i_2,i_3} is closed, it contains, by Theorem 3.5, the limit of its converging sequences; hence, the limit $\vec{v}$ of the sequence $(\vec{v_j})_{j \in J}$ has a r-encoding in L_{i_1,i_2,i_3}, which leads to a contradiction since this limit would be an element of S_{i_1,i_2,i_3}.

It follows that the set of real vectors that have an encoding in L is a countable union

$$\bigcup_{(i_1,i_2,i_3) \in \mathbb{N}^3} S_{i_1,i_2,i_3}$$

of closed sets in $\mathbb{R}^n$, and thus belongs to F_σ. ∎

Theorem 3.22 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and let $S \subseteq \mathbb{R}^n$. The set S belongs to $\mathsf{F}_\sigma \cap \mathsf{G}_\delta$ if and only if the language of r-encodings*

$$\{w \in \{0, r-1\}^n(\{0, 1, \dots, r-1\}^n)^* \star (\{0, 1, \dots, r-1\}^n)^\omega \mid \langle w \rangle_{r,n} \in S\}$$

belongs to $F_\sigma \cap G_\delta$.

Proof Let L be the language of r-encodings of S, i.e.,

$$\{w \in \{0, r-1\}^n(\{0, 1, \dots, r-1\}^n)^* \star (\{0, 1, \dots, r-1\}^n)^\omega \mid \langle w \rangle_{r,n} \in S\}.$$

It is known [BJW05] that if the set $S \subseteq \mathbb{R}^n$ belongs to $\mathsf{F}_\sigma \cap \mathsf{G}_\delta$, then L belongs to $F_\sigma \cap G_\delta$.

If L belongs to $F_\sigma \cap G_\delta$, then it belongs in particular to F_σ. By Theorem 3.21, S then belongs to F_σ. On the other hand, L belongs to G_δ, hence L is a countable intersection of open sets. It follows that the complement of L with respect to $(\{0, 1, \dots, r-1\}^n \cup \{\star\})^\omega$, denoted by L', is a countable union of complements of open sets, i.e., a countable union of closed sets. The language L' thus belongs to F_σ. By Theorem 3.21, the set of real vectors that have a r-encoding in L' belongs to F_σ, which implies that S belongs to G_δ. ■

3.2 Weak automata

As discussed in Section 2.5.3, applying basic set-theory operators to RVA reduces to carrying out the same operations on their accepted languages. Practically, this is somehow problematic, since operations like set complementation are typically costly and tricky to implement on infinite-word automata, as noticed in Section 2.3.3.

In order to bypass this problem, it has been shown that the full expressive power of Büchi automata is not needed for representing the subsets of $\mathbb{R}^n$, with $n \in \mathbb{N}_{>0}$, that are definable in the theory $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. Indeed, we will see in this section that such sets can be represented by a particular subclass of Büchi automata, *weak deterministic automata*. Those automata can be manipulated essentially in the same way as finite-word ones [Wil93].

3.2.1 Definition and link with topology

The following definition introduces the notion of *weak automata*.

Definition 3.23 *Let $\mathcal{A} = (Q, \Sigma, \Delta, q_0, F)$ be a Büchi automaton. The automaton $\mathcal{A}$ is said to be* weak *if its set of states Q can be partitioned into disjoint subsets $Q_1, Q_2, \ldots, Q_m$, where*

- *each Q_i contains either accepting or non-accepting states, i.e., either $Q_i \subseteq F$, or $Q_i \cap F = \emptyset$, and*
- *there exists a partial order $\leq$ on the sets $Q_1, Q_2, \ldots, Q_m$ such that for every transition $(q, w, q') \in \Delta$, with $q \in Q_i$, $q' \in Q_j$, and $w \in \Sigma^*$, we have $Q_j \leq Q_i$.*

□

Remark that, in Definition 3.23, the partition $\{Q_1, Q_2, \ldots, Q_m\}$ can always be chosen as being the decomposition of the transition graph of the automaton into its strongly connected components. The states contained in each strongly connected component of a weak automaton are thus always either accepting or non-accepting.

The languages accepted by weak automata have a particular characteristic. They can be represented both by Büchi and co-Büchi automata (such automata are defined at page 18), which can be justified as follows : The language accepted by weak automaton $(Q, \Sigma, \Delta, q_0, F)$ is exactly the language accepted by the co-Büchi automaton $(Q, \Sigma, \Delta, q_0, Q \setminus F)$. Indeed, each execution eventually stays within a single strongly connected component Q_i in which all states are either accepting or non-accepting.

As a consequence, weak automata can be determinized into deterministic co-Büchi automata, since they inherit this property from Theorem 2.38, which holds for co-Büchi automata.

The following two results will permit to show that weak deterministic automata accept exactly the ω-languages accepted by Büchi automata that belong to the topological class $F_\sigma \cap G_\delta$.

Theorem 3.24 ([Lan69]) *The languages accepted by deterministic Büchi automata are exactly the ω-languages accepted by Büchi automata that belong to the topological class G_δ.*

By duality, Theorem 3.24 admits the following corollary.

Corollary 3.25 *The languages accepted by deterministic co-Büchi automata are exactly the ω-languages accepted by Büchi automata that belong to the topological class F_σ.*

The next theorem is a consequence of Theorem 3.24 and of Corollary 3.25.

Theorem 3.26 ([SW74, Sta83, MS97]) *The languages accepted by weak deterministic automata are exactly the ω-languages accepted by Büchi automata that belong to the topological class $F_\sigma \cap G_\delta$.*

3.2.2 Algorithms

Like finite-word automata, an advantage of weak deterministic automata is that they admit a canonical form.

Theorem 3.27 ([Sta83, MS97]) *Each weak deterministic automaton admits a minimal form that is unique up to isomorphism.*

Furthermore, this normal form can efficiently be obtained.

Theorem 3.28 ([Löd01]) *The minimal form associated to a weak deterministic automaton is computable in time $\mathcal{O}(n \log n)$, where n is the number of states of the automaton.*

3.2.3 Expressiveness properties

In this section, we study the properties of the subsets of $\mathbb{R}^n$ that are recognizable by weak deterministic Büchi automata, and we give a sufficient condition for establishing that a set cannot be recognized by such an automaton.

Recall that the notation $F_\sigma \cap G_\delta$ (resp. $\mathsf{F}_\sigma \cap \mathsf{G}_\delta$) is used for the topology of infinite words (resp. the topology of real vectors).

Definition 3.29 *Let $n \in \mathbb{N}_{>0}$ be a dimension and $r \in \mathbb{N}_{>1}$ be a base. A set $S \subseteq \mathbb{R}^n$ is said to be* weakly r-recognizable *if and only if the set of r-encodings of the elements of S is recognizable by a weak deterministic Büchi automaton. Such an automaton is said to be a* weak deterministic RVA *recognizing S.* □

The following theorem links Theorems 3.22 and 3.26.

Theorem 3.30 *Let $n \in \mathbb{N}_{>0}$ be a dimension and $r \in \mathbb{N}_{>1}$ be a base. A set $S \subseteq \mathbb{R}^n$ is weakly r-recognizable if and only if it is r-recognizable, and it belongs to the topological class $\mathsf{F}_\sigma \cap \mathsf{G}_\delta$.*

Proof By Theorem 3.26, the language of r-encodings of S is weakly r-recognizable if and only if it is r-recognizable, and it belongs to $F_\sigma \cap G_\delta$. By Theorem 3.22, the last condition is equivalent to the membership of S to $\mathsf{F}_\sigma \cap \mathsf{G}_\delta$. ■

Theorem 3.30 admits the following corollary, since it has been shown that all sets definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$ belong to the class $\mathsf{F}_\sigma \cap \mathsf{G}_\delta$ [BJW05].

Theorem 3.31 ([BJW05]) *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $S \subseteq \mathbb{R}^n$ be a set definable in the first-order theory $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. The set S is weakly r-recognizable, for every base $r \in \mathbb{N}_{>1}$.*

The following theorems introduce some operations and transformations that preserve the weak recognizable nature of sets.

Theorem 3.32 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, $\ell \in \mathbb{N}_{>0}$, and $S \subseteq \mathbb{R}^n$. The set S is weakly r-recognizable if and only if it is weakly r^ℓ-recognizable.*

Proof If S is weakly r-recognizable (resp. weakly r^ℓ-recognizable), then, by Theorem 3.30, the set S belongs to the class $\mathsf{F}_\sigma \cap \mathsf{G}_\delta$. By Theorem 3.22, the language of r^ℓ-encodings (resp. r-encodings) of S belongs the the class $F_\sigma \cap G_\delta$. Moreover, this language is accepted by a Büchi automaton as a consequence of Theorem 2.66. The conclusion then follows from Theorem 3.30. ■

This theorem has the following corollary.

Corollary 3.33 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be multiplicatively dependent bases, and $S \subseteq \mathbb{R}^n$. The set S is weakly r-recognizable if and only if it is weakly s-recognizable.*

Proof By Definition 2.67, there exist $p, q \in \mathbb{N}_{>0}$ such that $r^p = s^q$.

By Theorem 3.32, S is weakly r-recognizable if and only if it is weakly r^p-recognizable. Using a second time this theorem, and since $r^p = s^q$, this condition holds if and only if S is weakly s-recognizable. ■

Theorem 3.34 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $S_1, S_2 \subseteq \mathbb{R}^n$. If S_1 and S_2 are both weakly r-recognizable, then the sets $S_1 \cup S_2$, $S_1 \cap S_2$, $S_1 \setminus S_2$ and $S_1 \times S_2$ are weakly r-recognizable as well.*

Proof The sets $S_1 \cup S_2$, $S_1 \cap S_2$, $S_1 \setminus S_2$ and $S_1 \times S_2$ are r-recognizable by Theorem 2.65.

By Theorem 3.30, the sets S_1 and S_2 both belong to the class $\mathsf{F}_\sigma \cap \mathsf{G}_\delta$, and so are the sets $S_1 \cup S_2$, $S_1 \cap S_2$, $S_1 \setminus S_2$ and $S_1 \times S_2$ by Theorems 3.13, 3.14, 3.16 and 3.19.

By Theorem 3.30, they are weakly r-recognizable. ■

Theorem 3.35 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, $C \in \mathbb{Q}^{n\times n}$ be such that $det(C) \neq 0$, $\vec{a} \in \mathbb{Q}^n$, and $S \subseteq \mathbb{R}^n$. If the set S is weakly r-recognizable, then the set $CS + \vec{a}$ is weakly r-recognizable as well.*

Proof The set $CS + \vec{a}$ is r-recognizable by Theorem 2.65.

By Theorems 3.30 and 3.17, it also belongs to the class $\mathsf{F}_\sigma \cap \mathsf{G}_\delta$.

By Theorem 3.30, it is weakly r-recognizable. ■

Theorem 3.36 *Let $n, m \in \mathbb{N}_{>0}$ be dimensions, $r \in \mathbb{N}_{>1}$ be a base, $\vec{a} \in \mathbb{R}^m$ and $S \subseteq \mathbb{R}^n$. If the set $S \times \{\vec{a}\}$ is weakly r-recognizable, then S is weakly r-recognizable as well.*

Proof The set S is r-recognizable by Theorem 2.65.

By Theorems 3.30 and 3.20, it also belongs to the class $\mathsf{F}_\sigma \cap \mathsf{G}_\delta$.

By Theorem 3.30, is is weakly r-recognizable. ■

In the remaining of this section, we exploit the notion of dense oscillating sequence property to prove that the subsets of $\mathbb{R}$ satisfying this property are not weakly recognizable. We will use the following theorem, that deals with languages instead of sets.

Theorem 3.37 ([BJW05]) *Let L be an ω-language accepted by a Büchi automaton. If L satisfies the dense oscillating sequence property, then L cannot be accepted by a weak deterministic automaton.*

Now, we move to one-dimensional sets $S \subseteq \mathbb{R}$.

Theorem 3.38 *Let $r \in \mathbb{N}_{>1}$ be a base. The r-recognizable sets $S \subseteq \mathbb{R}$ that satisfy the dense oscillating sequence property are not weakly r-recognizable.*

Proof Consider a r-recognizable set $S \subseteq \mathbb{R}$ satisfying the dense oscillating sequence property (in the sense of Definition 3.12). By Theorem 3.37, it is sufficient to establish that the language

$$L = \{w \in \{0, r-1\}\{0, 1, \ldots, r-1\}^* \star \{0, 1, \ldots, r-1\}^\omega \mid \langle w \rangle_r \in S\}$$

satisfies the dense oscillating sequence property as well (in the sense of Definition 3.10).

Recall that, by Theorem 2.48, each real number admits multiple encodings. First, the sign digit of an encoding can be repeated at will. Second, for a given length of integer part (assumed to be sufficiently large), a number admits either one or two (dual) encodings.

Let $S_1, S_2 \subseteq \mathbb{R}$ be sets of numbers such that $S_1 \cap S_2 = \emptyset$. Consider any number $x_1 \in S_1$ for which there exist arbitrarily close numbers in S_2.

For an encoding w, let $|w|_I$ be the length of the integer part of w.

There exists an encoding w_1 of x_1 for which there exist arbitrarily close encodings w_2 of numbers $x_2 \in S_2$, including the dual encodings with the same integer part length as w_1, if any. Indeed, if w_1 does not admit dual encodings, then the encodings of numbers that are arbitrarily close to x_1, with an integer part length equal to $|w_1|_I$, are all arbitrarily close to w_1. Otherwise, we choose w_1 as being a high (resp. low) encoding if there exist arbitrarily close numbers $x_2 \in S_2$ greater (resp. smaller) than x_1. One can then choose, among those numbers x_2, numbers that are arbitrarily close to x_1 and such that all their encodings w_2 such that $|w_2|_I = |w_1|_I$ are arbitrarily close to w_1.

Formally, we have

$$\begin{array}{rc}
& (x_1 \in S_1 \wedge (\forall \varepsilon \in \mathbb{R}_{>0})(\exists x_2 \in S_2)(d(x_1, x_2) < \varepsilon)) \\
\Rightarrow & (\exists w_1 \in \{0, r-1\}\{0, 1, \ldots, r-1\}^* \star \{0, 1, \ldots, r-1\}^\omega) \\
& ((\langle w_1 \rangle_r = x_1 \wedge (\forall \varepsilon' \in \mathbb{R}_{>0})(\exists x_2 \in S_2) \\
& ((\exists w_2 \in \{0, r-1\}\{0, 1, \ldots, r-1\}^* \star \{0, 1, \ldots, r-1\}^\omega) \\
& ((\langle w_2 \rangle_r = x_2 \wedge |w_1|_I = |w_2|_I) \\
& \wedge \\
& (\forall w_2 \in \{0, r-1\}\{0, 1, \ldots, r-1\}^* \star \{0, 1, \ldots, r-1\}^\omega) \\
& ((\langle w_2 \rangle_r = x_2 \wedge |w_1|_I = |w_2|_I \Rightarrow d(w_1, w_2) < \varepsilon'))).
\end{array}$$

By hypothesis, there exists $x_1 \in S$ such that

$$\forall \varepsilon_1 \exists x_2 \forall \varepsilon_2 \exists x_3 \forall \varepsilon_3 \ldots,$$

$d(x_i, x_{i+1}) \leq \varepsilon_i$ for all $i \in \mathbb{N}_{>0}$, $x_i \in S$ for all odd i, and $x_i \notin S$ for all even i. We choose $S_1 = S$, and define S_2 as the subset of $\mathbb{R} \setminus S$ whose elements x_2 satisfy

$$\forall \varepsilon_2 \exists x_3 \forall \varepsilon_3 \exists x_4 \forall \varepsilon_4 \ldots,$$

$d(x_i, x_{i+1}) \leq \varepsilon_i$ for all $i \in \mathbb{N}_{>1}$, $x_i \in S$ for all odd i, and $x_i \notin S$ for all even i.

By the previous property, there exists an encoding w_1 of x_1 such that for arbitrarily small $\varepsilon' \in \mathbb{R}_{>0}$, there exists an element $x_2 \in S_2$ whose all encodings w_2 satisfy $d(w_1, w_2) < \varepsilon'$, provided that $|w_2|_I = |w_1|_I$. Moreover, there exists at least one such encoding w_2. By applying a similar reasoning to $x_2, x_3, x_4, \ldots$, one obtains

$$\exists w_1 \forall \varepsilon'_1 \exists w_2 \forall \varepsilon'_2 \exists w_3 \forall \varepsilon'_3 \ldots,$$

$d(w_i, w_{i+1}) \leq \varepsilon'_i$ for all $i \in \mathbb{N}_{>0}$, $w_i \in L$ for all odd i, and $w_i \notin L$ for all even i. It follows that the language L satisfies the dense oscillating sequence property. ■

Chapter 4

Recognizability in multiple bases : From $\mathbb{R}^n$ to $[0, 1]^n$

In Chapters 2 and 3, we introduced general notions about the representation of vectors of reals by automata, and we gave related general results.

From now on, we will tackle the problem of extending the Cobham's and Semenov's theorems, introduced in Section 2.5.4, to automata recognizing real numbers and vectors.

4.1 Objectives

Our main objective will thus consist in characterizing precisely the conditions under which a set $S \subseteq \mathbb{R}^n$ is recognizable, or weakly recognizable, in multiple bases.

Precisely, we will in particular prove the following two statements :

1. Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be multiplicatively independent bases, and $S \subseteq \mathbb{R}^n$ be a set of real vectors. The set S is simultaneously weakly r- and weakly s-recognizable if and only if S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

2. Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases that do not share the same set of prime factors, and $S \subseteq \mathbb{R}^n$ be a set of real vectors. The set S is simultaneously r- and s-recognizable if and only if S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

For the second case, we will show in addition that considering bases with different sets of prime factors is essential. Indeed, we will provide an example of a set $S \subseteq \mathbb{R}$ that is recognizable in two multiplicatively independent bases, but not definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. This will lead to a complete characterization of the sets of real vectors recognizable in multiple bases.

Since, by Theorem 3.31 [BJW05], the sets definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$ are weakly recognizable (and thus also recognizable) in all bases $r \in \mathbb{N}_{>1}$, it remains to prove the following results :

1. Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be multiplicatively independent bases, and $S \subseteq \mathbb{R}^n$ be a set of real vectors. If S is simultaneously weakly r- and weakly s-recognizable, then S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.
2. Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases that do not share the same set of prime factors, and $S \subseteq \mathbb{R}^n$ be a set of real vectors. If S is simultaneously r- and s-recognizable, then S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

In the next section, we show that these problems can be reduced to simpler ones.

4.2 Reduction to the case of fractional parts

We consider sets $S \subseteq \mathbb{R}^n$ that are simultaneously recognizable, or weakly recognizable, in two bases r and s that are multiplicatively independent. This section is aimed at reducing this problem, by restricting the domain $\mathbb{R}^n$ to the n-cube $[0,1]^n$.

4.2.1 Decomposition

Before manipulating automata, we begin by general considerations about sets of vectors of reals.

For a set $S \subseteq \mathbb{R}^n$ and a vector $\vec{\imath} \in \mathbb{Z}^n$, let us define the set $F(S, \vec{\imath})$ of vectors of fractional parts that can be added to the vector $\vec{\imath}$ to obtain an element of S.

Definition 4.1 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $S \subseteq \mathbb{R}^n$, and $\vec{\imath} \in \mathbb{Z}^n$. The set $F(S, \vec{\imath})$ is defined by $F(S, \vec{\imath}) = \{\vec{x} \in [0,1]^n \mid \vec{\imath} + \vec{x} \in S\}$.* □

Equivalently, the set $F(S,\vec{\imath})$ is equal to

$$F(S,\vec{\imath}) = (S \cap \{\vec{\imath} + [0,1]^n\}) - \vec{\imath} = (S - \vec{\imath}) \cap [0,1]^n.$$

Remark that, for a set $S \subseteq \mathbb{R}^n$ and a vector $\vec{x} \in S$, there could exist several vectors $\vec{\imath} \in \mathbb{Z}^n$ such that $\vec{x} \in \vec{\imath} + F(S,\vec{\imath})$.

Example 4.2 *For the set $S = \{5\}$, we have $5 \in 4 + F(\{5\},4) = 4 + \{1\}$ and $5 \in 5 + F(\{5\},5) = 5 + \{0\}$.* ◇

Each set $S \subseteq \mathbb{R}^n$ can be built from its associated sets $F(S,\vec{\imath})$, as expressed in the following lemma.

Lemma 4.3 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $S \subseteq \mathbb{R}^n$. We have*

$$S = \bigcup_{i \in \mathbb{Z}^n} (\vec{\imath} + F(S,\vec{\imath})).$$

Proof This lemma directly follows from Definition 4.1. ■

Given a set $S \subseteq \mathbb{R}^n$, the set $\mathbb{Z}^n$ can be decomposed into equivalent classes such that two vectors $\vec{\imath_1}, \vec{\imath_2} \in \mathbb{Z}^n$ are in the same equivalence class if and only if their associated sets $F(S,\vec{\imath_1})$ and $F(S,\vec{\imath_2})$ are identical. The following definition formalizes this decomposition.

Definition 4.4 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $S \subseteq \mathbb{R}^n$, and $\vec{\imath} \in \mathbb{Z}^n$. The set $I(S,\vec{\imath})$ is defined by*

$$I(S,\vec{\imath}) = \{\vec{\imath'} \in \mathbb{Z}^n \mid F(S,\vec{\imath}) = F(S,\vec{\imath'})\}.$$

□

The equivalence class that contains a vector $\vec{\imath} \in \mathbb{Z}^n$ is the class $I(S,\vec{\imath})$. The union expressed in Lemma 4.3 can be adapted into a (finite or infinite) union, in terms of the sets $I(S,\vec{\imath})$, according to the following lemma.

Lemma 4.5 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $S \subseteq \mathbb{R}^n$. We have*

$$S = \bigcup_{(S^I,S^F) \in \mathcal{D}(S)} (S^I + S^F),$$

where

$$\mathcal{D}(S) = \{(I(\vec{\imath}), F(S,\vec{\imath})) \subseteq \mathbb{Z}^n \times [0,1]^n \mid \vec{\imath} \in \mathbb{Z}^n \wedge F(S,\vec{\imath}) \neq \emptyset\}.$$

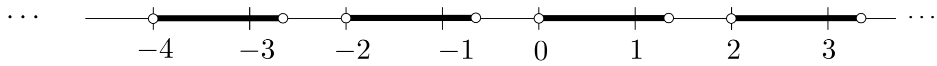

Figure 4.1: Representation of the set $2\mathbb{Z} + \left]0, \frac{4}{3}\right[$.

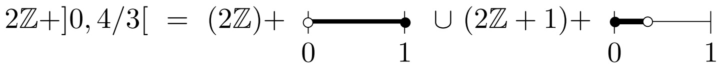

Figure 4.2: Decomposition of the set $2\mathbb{Z} + \left]0, \frac{4}{3}\right[$.

Proof This lemma directly follows from Lemma 4.3 and Definition 4.4. ■

Note that all the sets S^I and S^F belonging to the union expressed in Lemma 4.5 are non-empty, and that each set $S \subseteq \mathbb{R}^n$ is decomposed uniquely into that union.

The following two examples are aimed at illustrating this decomposition.

Example 4.6 *The set* $S = 2\mathbb{Z} + \left]0, \frac{4}{3}\right[$*, depicted in Figure 4.1, is the set of numbers* $x \in \mathbb{R}$ *that are the sum of an even integer (resp. an odd integer) and an element of* $]0, 1]$ *(resp.* $\left[0, \frac{1}{3}\right[$*). It thus admits the following decomposition :*

$$\mathcal{D}(S) = \{(2\mathbb{Z},]0, 1]), (2\mathbb{Z} + 1, \left[0, \frac{1}{3}\right[)\}.$$

This decomposition is illustrated in Figure 4.2. ◇

Example 4.7 *The set* $S \subseteq \mathbb{R}^2$ *defined by the periodic tiling of Example 2.7 admits the following decomposition :*

$$\begin{aligned}\mathcal{D}(S) = \{&((2\mathbb{Z} \times 2\mathbb{Z}), \{(x_1, x_2) \in [0,1]^n \mid x_2 \geq x_1\}),\\ &(((2\mathbb{Z}+1) \times 2\mathbb{Z}), (\{1\} \times [0,1]) \cup \{(0,1)\}),\\ &((2\mathbb{Z} \times (2\mathbb{Z}+1)), ([0,1] \times \{0\}) \cup \{(0,1)\}),\\ &(((2\mathbb{Z}+1) \times (2\mathbb{Z}+1)), \{(0,0), (1,0), (1,1)\})\}.\end{aligned}$$

This decomposition is illustrated in Figure 4.3. ◇

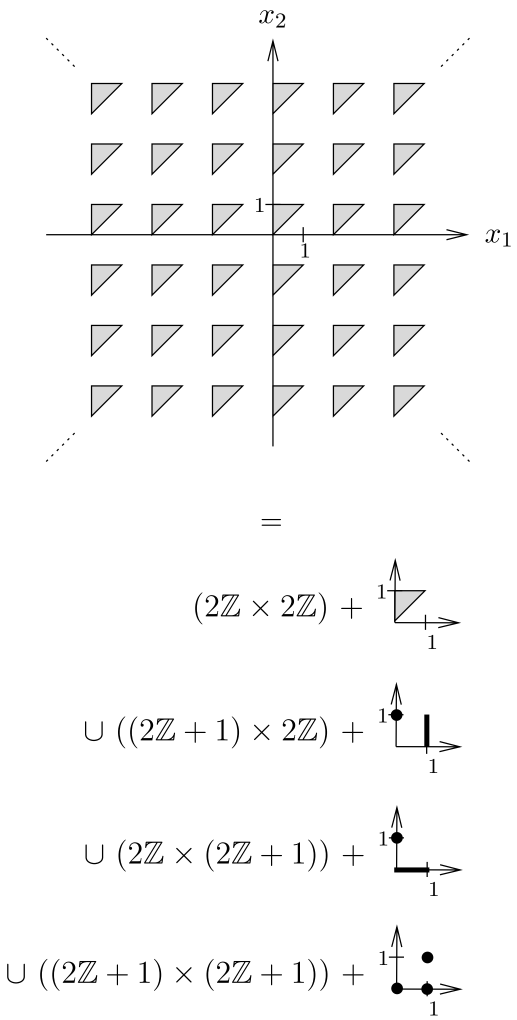

Figure 4.3: Decomposition of a periodic tiling in $\mathbb{R}^2$.

Assume that a set $S \subseteq \mathbb{R}^n$ is either r-recognizable, or weakly r-recognizable, and let $\mathcal{A}$ be a (possibly weak deterministic) RVA recognizing it. We now show that the decomposition, expressed in Lemma 4.5, of S into integer and fractional parts, is strongly related to the languages accepted in $\mathcal{A}$ from the states that are destinations of transitions labeled by the separator $\star$.

To simplify the technical developments, we assume that the manipulated automata have a deterministic transition relation : Instead of dealing with Büchi automata and RVA, we use Muller automata, introduced in Definition 2.24, and Muller RVA. We can make this assumption without loss of generality since, by Theorem 2.26, Büchi and Muller automata share the same expressive power. In the case of weak recognizability, we continue to use weak deterministic automata.

We first show that it is possible to map the Muller (resp. weak deterministic) automaton $\mathcal{A}$ into an Muller (resp. weak deterministic) automaton $\mathcal{A}'$ having the following properties : $\mathcal{A}'$ accepts the same language as $\mathcal{A}$, each state q of $\mathcal{A}'$ is useful, in the sense that there exists an accepting path visiting q, and the states that are destinations of transitions labeled by $\star$ accept languages that are pairwise different.

Definition 4.8 *Let $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ be either a weak deterministic automaton, or a Muller automaton, and let $q \in Q$. The language $L(q)$ of the suffixes, accepted from q, of words of $L(\mathcal{A})$ is defined by*

$$L(q) = \{w \mid (\exists w' \in \Sigma^*)(\delta^*(q_0, w') = q \wedge w'w \in L(\mathcal{A}))\}.$$

□

Lemma 4.9 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $r \in \mathbb{N}_{>1}$ be a base. Let $\mathcal{A}$ be a Muller (resp. weak deterministic) RVA. If $L(\mathcal{A}) \neq \emptyset$, then there exists a Muller (resp. weak deterministic) RVA*

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

such that three following conditions hold :

- *$L(\mathcal{A}) = L(\mathcal{A}')$;*
- *for each $q \in Q$, we have $L(q) \neq \emptyset$;*
- *let $Q_\star = \{q \in Q \mid (\exists q' \in Q)(\delta(q', \star) = q)\}$. For each pair of distinct states $q, q' \in Q_\star$, we have $L(q) \neq L(q')$.*

Proof First, let $\mathcal{A}'$ be a copy of $\mathcal{A}$ where the states $q \in Q$ such that $L(q) = \emptyset$ have been removed from Q.

Second, since $\mathcal{A}'$ is a RVA, it only accepts words of the form $w \in \{0, r-1\}^n(\{0,1,\ldots,r-1\}^n)^* \star (\{0,1,\ldots,r-1\}^n)^\omega$. Hence, there are no cycles following a transition labeled by the symbol $\star$. If $Q' \subseteq Q_\star$ is a (maximal) subset of states such that for each pair of states $q, q' \in Q'$ we have $L(q) = L(q')$, then let $q_1 \in Q'$ be an arbitrary state of this subset. For each $q_2 \in Q'$ such that $q_1 \neq q_2$, it suffices to modify the transition function δ in order to have $\delta(q_2', \star) = q_1$, where q_2' is the unique state such that $\delta(q_2', \star) = q_2$. Since $L(q) = L(q')$ for all $q, q' \in Q'$ and since $\mathcal{A}'$ has a deterministic transition relation, those operations do not alter the language accepted by $\mathcal{A}$. Moreover, they can be repeated for each such subset $Q' \subseteq Q_\star$. By doing so, we get a RVA $\mathcal{A}''$ satisfying the expected conditions. ■

The following lemma gives links between the sets $F(S, \vec{\imath})$ and the states that are destinations of transitions labeled by $\star$.

Lemma 4.10 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ be either a weak deterministic RVA, or a Muller RVA, recognizing a non-empty set $S \subseteq \mathbb{R}^n$ in base r. Suppose that*

- *for each $q \in Q$, we have $L(q) \neq \emptyset$;*
- *let $Q_\star = \{q \in Q \mid (\exists q' \in Q)(\delta(q', \star) = q)\}$. For each pair of distinct states $q, q' \in Q_\star$, we have $L(q) \neq L(q')$.*

The following two properties hold :

1. *Let $q \in Q_\star$, and let $w \in \{0, r-1\}^n(\{0,1,\ldots,r-1\}^n)^*$ be a r-encoding such that $\delta^*(q_0, w\star) = q$. We have $F(S, \langle w \rangle_{r,n}) = \{\langle \vec{0} \star w' \rangle_{r,n} \mid w' \in L(q)\}$.*
2. *Let $w_1, w_2 \in \{0, r-1\}^n(\{0,1,\ldots,r-1\}^n)^*$ be two r-encodings such that $F(S, \langle w_1 \rangle_{r,n})$ and $F(S, \langle w_2 \rangle_{r,n})$ are both non-empty. We have $F(S, \langle w_1 \rangle_{r,n}) \neq F(S, \langle w_2 \rangle_{r,n})$ if and only if $\delta^*(q_0, w_1\star) \neq \delta^*(q_0, w_2\star)$.*

Proof We prove the two parts separately.

1. One the one hand, let $\vec{x}$ in $F(S, \langle w \rangle_{r,n})$. We have $\vec{x} \in [0,1]^n$. Since, by definition, each RVA accepts all the encodings of the vectors it represents, $\mathcal{A}$ accepts at least a word $w' = w \star w_2$ such that $\langle \vec{0} \star w_2 \rangle_{r,n} =$

$\vec{x}$. Since $\mathcal{A}$ has a deterministic transition relation, the reading of w_2 in $\mathcal{A}$ starts at the state q. Hence, $\vec{x}$ belongs to $\{\langle\vec{0} \star w'\rangle_{r,n} \mid w' \in L(q)\}$.

On the other hand, let $\vec{x}$ be an element of $\{\langle\vec{0} \star w'\rangle_{r,n} \mid w' \in L(q)\}$. We fix such a word w'. Since $q = \delta^*(q_0, w\star)$, we know that the word $w \star w'$ is accepted by $\mathcal{A}$, that is, $\langle w \star w'\rangle_{r,n} = \langle w\rangle_{r,n} + \vec{x}$ belongs to S. It follows that $\vec{x}$ belongs to $F(S, \langle w\rangle_{r,n})$.

2. Let $q_1 = \delta^*(q_0, w_1\star)$ and $q_2 = \delta^*(q_0, w_2\star)$.

   If $F(S, \langle w_1\rangle_{r,n}) \neq F(S, \langle w_1\rangle_{r,n})$, then, by the first part, we have

$$\{\langle\vec{0} \star w'\rangle_{r,n} \mid w' \in L(q_1)\} \neq \{\langle\vec{0} \star w'\rangle_{r,n} \mid w' \in L(q_2)\}.$$

   Hence, $q_1 \neq q_2$.

   Now, we suppose $q_1 \neq q_2$. By hypothesis, we have $L(q_1) \neq L(q_2)$. Without loss of generality, let $w \in L(q_1) \setminus L(q_2)$. By the first part, we have to prove that

$$\{\langle\vec{0} \star w'\rangle_{r,n} \mid w' \in L(q_1)\} \neq \{\langle\vec{0} \star w'\rangle_{r,n} \mid w' \in L(q_2)\}.$$

   For this purpose, we will prove that

$$\langle\vec{0} \star w\rangle_{r,n} \notin \{\langle\vec{0} \star w'\rangle_{r,n} \mid w' \in L(q_2)\}.$$

   By contradiction, suppose that $(\exists w'' \in L(q_2))(\langle\vec{0} \star w''\rangle_{r,n} = \langle\vec{0} \star w\rangle_{r,n})$. Since $\mathcal{A}$ has a deterministic transition relation, we have $w_2 \star w'' \in L(\mathcal{A})$ and $w_2 \star w' \notin L(\mathcal{A})$. This is not possible, since $\mathcal{A}$ accepts all the encodings of the vectors it represents.

■

Lemma 4.10 admits the following corollary, which gives another way to express the decomposition $\mathcal{D}(S)$.

Corollary 4.11 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ be either a weak deterministic RVA, or a Muller RVA, recognizing a non-empty set $S \subseteq \mathbb{R}^n$ in base r. Suppose that*

- *for each $q \in Q$, we have $L(q) \neq \emptyset$;*
- *let $Q_\star = \{q \in Q \mid (\exists q' \in Q)(\delta(q', \star) = q)\}$. For each pair of distinct states $q, q' \in Q_\star$, we have $L(q) \neq L(q')$.*

For each set $(I, F) \in \mathcal{D}(S)$, there exists exactly one state $q \in Q_\star$ such that $F = \{\langle\vec{0} \star w\rangle_{r,n} \mid w \in L(q)\}$.

Proof This corollary directly follows from the definition of $\mathcal{D}(S)$, introduced in Lemma 4.5, and from Lemma 4.10. ■

Since Corollary 4.11 holds whatever the representation base r is, we conclude that the size of the set $Q_\star$ is independent from the representation base. Hence, we have the following result :

Corollary 4.12 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ be either a weak deterministic RVA, or a Muller RVA, recognizing a non-empty set $S \subseteq \mathbb{R}^n$ in base r. Suppose that*

- *for each $q \in Q$, we have $L(q) \neq \emptyset$;*
- *let $Q_\star = \{q \in Q \mid (\exists q' \in Q)(\delta(q', \star) = q)\}$. For each pair of distinct states $q, q' \in Q_\star$, we have $L(q) \neq L(q')$.*

The size of the set $Q_\star$ only depends on the set S.

Proof This is a direct corollary of Corollary 4.11. ■

We are now able to prove the main result of this section.

Theorem 4.13 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $\mathcal{A}$ be either a weak deterministic RVA, or a Muller RVA, recognizing a set $S \subseteq \mathbb{R}^n$ in base r. The set S can be decomposed into a finite union*

$$\bigcup_{i=1}^{m} (S_i^I + S_i^F),$$

where $m \in \mathbb{N}$ and, for each $i \in \{1, 2, \ldots, m\}$,

- *the set $S_i^I \subseteq \mathbb{Z}^n$ is recognizable by a NDD in every base in which S is recognizable;*
- *the set $S_i^F \subseteq [0, 1]^n$ is (resp. weakly) recognizable in every base in which S is (resp. weakly) recognizable.*

Proof If $S = \emptyset$, then the proof is immediate.

Otherwise, by Lemma 4.5, we have

$$S = \bigcup_{(S^I, S^F) \in \mathcal{D}(S)} (S^I + S^F).$$

Let $r' \in \mathbb{N}_{>1}$ be a base. Suppose that S is (resp. weakly) r'-recognizable. By Lemma 4.9, there exists a Muller RVA (resp. a weak deterministic RVA) $\mathcal{A}' = (Q, \Sigma, \delta, q_0, F)$ such that the following three conditions hold :

- $\mathcal{A}'$ recognizes S in base r';
- for each $q \in Q$, we have $L(q) \neq \emptyset$;
- let $Q_\star = \{q \in Q \mid (\exists q' \in Q)(\delta(q', \star) = q)\}$. For each pair of distinct states $q, q' \in Q_\star$, we have $L(q) \neq L(q')$.

We know by Corollary 4.11 that

$$\mathcal{D}(S) = \bigcup_{q \in Q_\star} (\{\langle w_1 \rangle_{r',n} \mid \delta^*(q_0, w_1\star) = q\}, \{\langle \vec{0} \star w_2 \}_{r',n} \mid w_2 \in L(q)\}).$$

Hence, this decomposition, as well as the components of the union

$$S = \bigcup_{(S^I, S^F) \in \mathcal{D}(S)} (S^I + S^F),$$

do not change with the representation base r'. If $|Q_\star| = m$, then we thus have

$$S = \bigcup_{i=1}^{m} (S_i^I + S_i^F)$$

where there exists a one-to-one mapping $b : \{1, 2, \ldots, m\} \to Q_\star$, and, for each $i \in \{1, 2, \ldots, m\}$,

$$S_i^I = \{\langle w_1 \rangle_{r',n} \mid \delta^*(q_0, w_1\star) = b(i)\}$$

and

$$S_i^F = \{\langle \vec{0} \star w_2 \rangle_{r',n} \mid w_2 \in L(b(i))\}.$$

For each $i \in \{1, 2, \ldots, m\}$, the set S_i^I is recognizable by a NDD in base r'. Indeed, a NDD recognizing it can be built from $\mathcal{A}'$ by marking as accepting states the states that belong to $\{q \in Q \mid \delta(q, \star) = b(i)\}$.

Finally, for each $i \in \{1, 2, \ldots, m\}$, the set S_i^F is recognizable by a (resp. weak) RVA in base r'. Indeed, if $w \in \{0, r'-1\}^n (\{0, 1, \ldots, r'-1\}^n)^*$ is an arbitrary word such that $\delta^*(q_0, w\star) = b(i)$, then we have

$$S_i^F = (S - \langle w \rangle_{r',n}) \cap [0,1]^n$$

since $\mathcal{A}'$ is deterministic. It follows from Theorem 2.65 (resp. Theorems 3.34 and 3.35) that S_i^F is (resp weakly) r'-recognizable. ■

The following example illustrates Theorem 4.13.

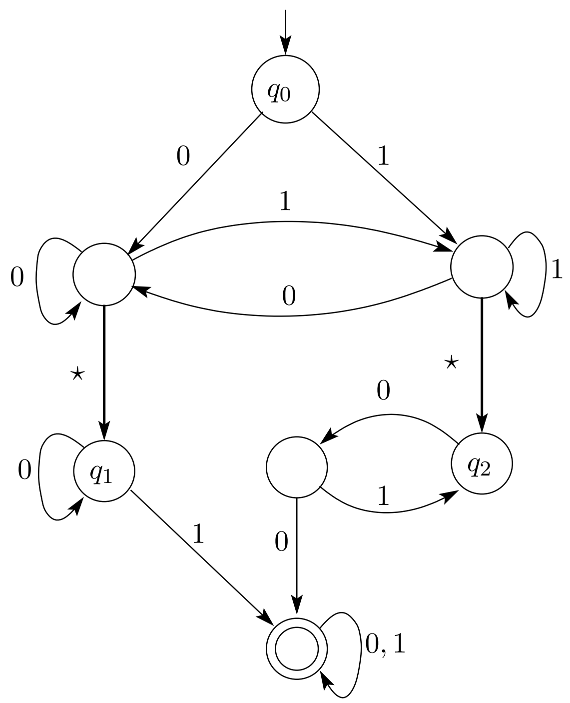

Figure 4.4: RVA representing $2\mathbb{Z} + \left]0, \frac{4}{3}\right[$ in base 2.

Example 4.14 *Consider the set* $S = 2\mathbb{Z} + \left]0, \frac{4}{3}\right[$*, introduced in Example 4.6 and depicted in Figure 4.1.*

Weak deterministic RVA $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ *and* $\mathcal{A}' = (Q', \Sigma, \delta', q'_0, F')$ *representing it in bases* 2 *and* 3 *are respectively shown in Figures 4.4 and 4.5.*

We have

$$\mathcal{D}(S) = \{(2\mathbb{Z},]0, 1]), (2\mathbb{Z} + 1, \left[0, \frac{1}{3}\right[)\}.$$

In this decomposition, the sets of integers are $2\mathbb{Z}$ *and* $2\mathbb{Z}+1$*. We can remark that*

$$2\mathbb{Z} = \{\langle w \rangle_2 \mid \delta^*(q_0, w\star) = q_1\} = \{\langle w \rangle_3 \mid \delta'^*(q'_0, w\star) = q_3\},$$

and that

$$2\mathbb{Z} + 1 = \{\langle w \rangle_2 \mid \delta^*(q_0, w\star) = q_2\} = \{\langle w \rangle_3 \mid \delta'^*(q'_0, w\star) = q_4\}.$$

Similarly, the sets of fractional parts that are in $\mathcal{D}(S)$ *are* $]0, 1]$ *and* $[0, 1/3[$*, and we have*

$$]0, 1] = \{\langle 0 \star w \rangle_2 \mid w \in L(q_1)\} = \{\langle 0 \star w \rangle_3 \mid w \in L(q_3)\},$$

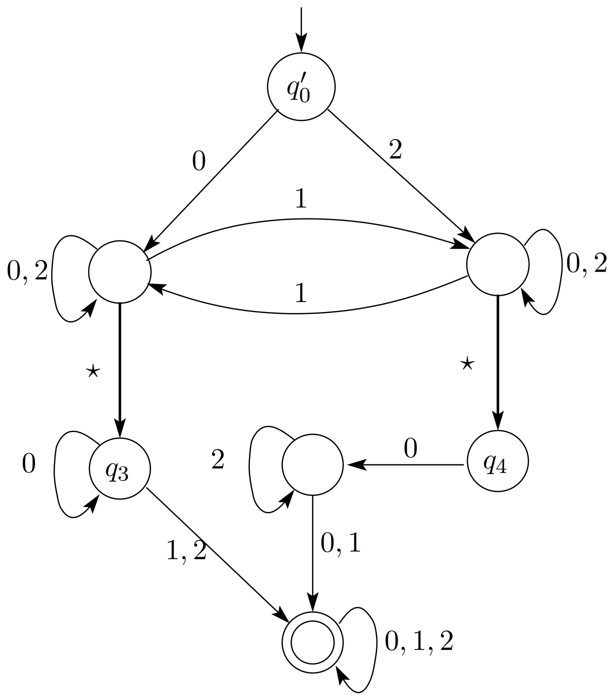

Figure 4.5: RVA representing $2\mathbb{Z} + \left]0, \frac{4}{3}\right[$ in base 3.

and

$$\left[0, \frac{1}{3}\right[= \{\langle 0 \star w\rangle_2 \mid w \in L(q_2)\} = \{\langle 0 \star w\rangle_3 \mid w \in L(q_4)\}.$$

◇

4.2.2 Reduction

The previous section was aimed at proving that each set represented by a (possibly weak deterministic) RVA admits a finite decomposition into integer and fractional parts such that this decomposition does not depend on the representation base. We now use this property, expressed in Theorem 4.13, for reducing the problems discussed in Section 4.1.

Theorem 4.15 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases that do not share the same set of prime factors, and $S \subseteq \mathbb{R}^n$ be simultaneously r- and s-recognizable. If the following property holds :*

> *For each set $S' \subseteq [0,1]^n$ simultaneously r- and s-recognizable, S' is definable in $\langle \mathbb{R}, +, <, 1\rangle$,*

then S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, <\rangle$.

Proof By Theorem 4.13, the set S can be decomposed into a finite union

$$\bigcup_{i=1}^{m} (S_i^I + S_i^F),$$

where $m \in \mathbb{N}$ and, for each $i \in \{1, 2, \dots, m\}$,

- the set $S_i^I \subseteq \mathbb{Z}^n$ is recognizable by NDDs in base r and s;
- the set $S_i^F \subseteq [0,1]^n$ is recognizable by RVA in base r and s.

By Cobham's and Semenov's theorems (Theorems 2.71 and 2.72), the set S_i^I is definable in $\langle \mathbb{Z}, +, <\rangle$ for each $i \in \{1, 2, \dots, m\}$.

By hypothesis, and for each $i \in \{1, 2, \dots, m\}$, the set S_i^F is definable in $\langle \mathbb{R}, \mathbb{Z}, +, <\rangle$.

Hence, since the union is finite, the set S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, <\rangle$. ■

Theorem 4.16 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be multiplicatively independent bases, and $S \subseteq \mathbb{R}^n$ be simultaneously weakly r- and weakly s-recognizable. If the following property holds :*

> *For each set $S' \subseteq [0,1]^n$ simultaneously weakly r- and weakly s-recognizable, S' is definable in $\langle \mathbb{R}, +, <, 1 \rangle$,*

then S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

Proof The proof is identical to the proof of Theorem 4.15. ■

In both Theorems 4.15 and 4.16, we reduced the problem of establishing that S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$ to the problem of establishing the definability of the sets S_i^F in $\langle \mathbb{R}, +, <, 1 \rangle$. Since we have $S_i^F \subseteq [0,1]^n$ for all i, the problems have thus been reduced from the domain $\mathbb{R}$ to the n-cube $[0,1]^n$.

Of course, Theorems 4.15 and 4.16 also hold when the sets S' are definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$, but we will be more precise and we will show that the two properties expressed in Theorems 4.15 and 4.16 hold, i.e.,

1. Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be multiplicatively independent bases, and $S \subseteq [0,1]^n$. If S is simultaneously weakly r- and weakly s-recognizable, then S is definable in $\langle \mathbb{R}, +, <, 1 \rangle$.
2. Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases that do not share the same set of prime factors, and $S \subseteq [0,1]^n$. If S is simultaneously r- and s-recognizable, then S is definable in $\langle \mathbb{R}, +, <, 1 \rangle$.

These problems are addressed in the two following chapters, first in the case of one-dimensional sets, then in the case of multi-dimensional ones.

Chapter 5

One-dimensional sets

5.1 Overview

In this chapter, we precisely characterize the (one-dimensional) subsets of $\mathbb{R}$ that are recognizable, either by weak deterministic RVA or by Muller RVA, in multiple bases.

We prove the following claim, stating that if a subset of $\mathbb{R}$ is simultaneously recognizable by weak deterministic RVA in two multiplicatively independent bases, then this set is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

Claim 5.1 *Let $r, s \in \mathbb{N}_{>1}$ be multiplicatively independent bases, and let $S \subseteq \mathbb{R}$. If S is simultaneously weakly r- and weakly s-recognizable, then S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.*

Since, by Corollary 3.33, recognizability in two multiplicatively dependent bases is equivalent to recognizability in only one of them, this result provides a complete characterization of the sets that are recognizable in multiple bases by weak deterministic RVA.

Then, we move to subsets of $\mathbb{R}$ recognizable by Muller RVA. We establish that, for every pair r, s of bases that share the same set of prime factors, there exists a set of real numbers simultaneously r- and s-recognizable, but not definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. Formally, we establish Claim 5.2.

Claim 5.2 *Let $r, s \in \mathbb{N}_{>1}$ be two bases that share the same set of prime factors. There exists a set $S \subseteq \mathbb{R}$ that is both r- and s-recognizable, and that is not definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.*

This shows that there does not exist a direct generalization of Cobham's theorem to Muller RVA.

Finally, we prove Claim 5.3, i.e., that if a set $S \subseteq \mathbb{R}$ is simultaneously recognizable in two bases that do not share the same set of prime factors, then S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

Claim 5.3 *Let $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors, and let $S \subseteq \mathbb{R}$. If S is simultaneously r- and s-recognizable, then S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.*

5.2 Reduction of the problems

In this section, we introduce, with respect to a set $S \subseteq \mathbb{R}$, the notion of *boundary points* of S. These points share a common topological characteristic : In every arbitrary small neighborhood close to them, there exist at least one point of S and one point of the complement $\mathbb{R} \setminus S$ of S. We show that if a r-recognizable set S admits finitely many boundary points, then it is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

Formally, boundary points are defined in the following way, where the notation $N_\varepsilon(x)$ for the neighborhood of $x \in \mathbb{R}$ with respect to $\varepsilon \in \mathbb{R}_{>0}$, is reused from Definition 3.2, with the distance $d(x, y) = |x - y|$.

Definition 5.4 *Let $S \subseteq \mathbb{R}$ and let $x \in \mathbb{R}$. The point x is a* boundary point *of S if and only if for all $\varepsilon \in \mathbb{R}_{>0}$, we have*

$$N_\varepsilon(x) \cap S \neq \emptyset$$

as well as

$$N_\varepsilon(x) \cap (\mathbb{R} \setminus S) \neq \emptyset.$$

□

The set $B(S)$ of boundary points of a set S can then be defined as follows.

Definition 5.5 *Let $S \subseteq \mathbb{R}$. The set $B(S)$ of* boundary points *of S is*

$$B(S) = \{x \in \mathbb{R} \mid (\forall \varepsilon \in \mathbb{R}_{>0})((N_\varepsilon(x) \cap S \neq \emptyset) \wedge (N_\varepsilon(x) \cap (\mathbb{R} \setminus S) \neq \emptyset))\}.$$

□

The following lemma proves that if r is a base and S is a r-recognizable set, then the set of boundary points of S is also r-recognizable.

Lemma 5.6 *Let $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq \mathbb{R}$. If S is r-recognizable, then the set $B(S)$ is r-recognizable as well.*

Proof Since S is r-recognizable, it is definable in $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$ by Theorem 2.65. By the same theorem, the set $B(S)$ is r-recognizable if and only if it is definable in $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$. It is actually the case, since a formula defining $B(S)$ in this theory is

$$\{x \in \mathbb{R} \mid (\forall \varepsilon \in \mathbb{R}_{>0})(\exists y, z \in \mathbb{R})(y \in F_S \wedge z \notin F_S \wedge |x - y| < \varepsilon \wedge |x - z| < \varepsilon)\},$$

where F_S is a formula defining S in $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$. ■

We can now prove that if a set has the properties of being r-recognizable and admitting a finite number of boundary points, then it is definable in the theory $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

Theorem 5.7 *Let $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq \mathbb{R}$. If S is r-recognizable, and if the set $B(S)$ is finite, then S is definable in $\langle \mathbb{R}, +, <, 1 \rangle$.*

Proof Since the set $B(S)$, i.e., the set of boundary points of S, is finite, the set S can be decomposed into a finite union of intervals such that the extremities of these intervals are the elements of $B(S)$. Formally, if $B(S) = \emptyset$, then S is either equal to $\mathbb{R}$ or to $\emptyset$. Otherwise, if $B(S) \neq \emptyset$, then suppose that $B(S) = \{x_1, x_2, \ldots, x_{|B(S)|}\}$, with $x_1 < x_2 < \ldots < x_{|B(S)|}$, and let $I(S)$ be the set of intervals with boundaries in $B(S)$, i.e.,

$$I(S) = \{]-\infty, x_1[, \{x_1\},]x_1, x_2[, \{x_2\}, \ldots, \{x_{|B(S)|}\},]x_{|B(S)|}, +\infty[\}.$$

There exists a set $U \subseteq I(S)$ such that

$$S = \bigcup_{u \in U} u.$$

By Lemma 5.6, $B(S)$ is r-recognizable. Since $B(S)$ is finite, one can find, for each point $x \in B(S)$, two rational numbers $a, b \in \mathbb{Q}$ such that $B(S) \cap [a, b] = \{x\}$. By Theorem 2.65, the set $\{x\}$ is then r-recognizable. By Theorem 2.22, the number x admits one ultimately periodic encoding. It follows by Theorem 2.51 that x is rational.

The set S can thus be expressed as a finite union of intervals with rational extremities. Hence, it is definable in $\langle \mathbb{R}, +, <, 1 \rangle$. ■

Theorem 5.7 is useful since it provides a reduction of the problem of establishing Claims 5.1 and 5.3 : In order to prove that S, simultaneously recognizable in two different bases, is definable in the theory $\langle \mathbb{R}, +, <, 1\rangle$, it suffices to show that S only admits a finite number of boundary points. This reduction will be used in the following sections : We will suppose that such a set admits infinitely many boundary points, and we will show that this assumption leads to a contradiction.

5.3 Product-stability properties

In this section, we focus on subsets of $[0, 1]$ that are simultaneously recognizable in two bases, either by weak deterministic or by Muller automata. As discussed in Section 5.2, we suppose in addition that the sets we consider admit infinitely many boundary points.

Under these hypotheses, we exhibit the existence of a set $S \subseteq [0, 1]$ that inherits those properties, and that presents additional *product-stability* properties which will be useful later. Intuitively, we show that the set S is, in a given way, invariant under the multiplication by the values of the bases in which S is recognizable.

The next lemma will be useful. It gives a strong link between the repetition of a subword in an encoding and the values of these encodings.

Lemma 5.8 *Let* $r \in \mathbb{N}_{>1}$ *be a base,* $w_1 \in \{0, r-1\}\{0, 1, \ldots, r-1\}^*$, $w_2, w_3 \in \{0, 1, \ldots, r-1\}^*$, *and* $w_4 \in \{0, 1, \ldots, r-1\}^\omega$. *For each* $k \in \mathbb{N}$, *we have :*

- *If* $\langle w_1 \star w_2 w_4\rangle_r > \langle w_1 \star w_2 w_3 w_4\rangle_r$, *then* $\langle w_1 \star w_2 w_3^k w_4\rangle_r > \langle w_1 \star w_2 w_3^{k+1} w_4\rangle_r$.
- *If* $\langle w_1 \star w_2 w_4\rangle_r = \langle w_1 \star w_2 w_3 w_4\rangle_r$, *then* $\langle w_1 \star w_2 w_3^k w_4\rangle_r = \langle w_1 \star w_2 w_3^{k+1} w_4\rangle_r$.
- *If* $\langle w_1 \star w_2 w_4\rangle_r < \langle w_1 \star w_2 w_3 w_4\rangle_r$, *then* $\langle w_1 \star w_2 w_3^k w_4\rangle_r < \langle w_1 \star w_2 w_3^{k+1} w_4\rangle_r$.

Proof First, note that if

$$\langle w_1 \star w_2 w_4\rangle_r \# \langle w_1 \star w_2 w_3 w_4\rangle_r,$$

with $\# \in \{<, =, >\}$, then

$$\langle 0 \star w_4\rangle_r \# \langle 0 \star w_3 w_4\rangle_r.$$

Indeed, we have

$$\begin{aligned}
& \langle w_1 \star w_2 w_3 w_4 \rangle_r - \langle w_1 \star w_2 w_4 \rangle_r \\
= \; & \frac{\langle w_1 w_2 \star w_3 w_4 \rangle_r}{r^{|w_2|}} - \frac{\langle w_1 w_2 \star w_4 \rangle_r}{r^{|w_2|}} \\
= \; & \frac{\langle w_1 w_2 \rangle_r + \langle 0 \star w_3 w_4 \rangle_r}{r^{|w_2|}} - \frac{\langle w_1 w_2 \rangle_r + \langle 0 \star w_4 \rangle_r}{r^{|w_2|}} \\
= \; & \frac{\langle 0 \star w_3 w_4 \rangle_r - \langle 0 \star w_4 \rangle_r}{r^{|w_2|}},
\end{aligned}$$

which implies, since $r^{|w_2|} > 0$, that the signs of the values $\langle w_1 \star w_2 w_3 w_4 \rangle_r - \langle w_1 \star w_2 w_4 \rangle_r$ and $\langle 0 \star w_3 w_4 \rangle_r - \langle 0 \star w_4 \rangle_r$ are identical.

Using this fact, and remarking that

$$\begin{aligned}
& \langle w_1 \star w_2 w_3^{k+1} w_4 \rangle_r - \langle w_1 \star w_2 w_3^k w_4 \rangle_r \\
= \; & \frac{\langle w_1 w_2 w_3^k \star w_3 w_4 \rangle_r}{r^{|w_2|+k|w_3|}} - \frac{\langle w_1 w_2 w_3^k \star w_4 \rangle_r}{r^{|w_2|+k|w_3|}} \\
= \; & \frac{\langle w_1 w_2 w_3^k \rangle_r + \langle 0 \star w_3 w_4 \rangle_r}{r^{|w_2|+k|w_3|}} - \frac{\langle w_1 w_2 w_3^k \rangle_r + \langle 0 \star w_4 \rangle_r}{r^{|w_2|+k|w_3|}} \\
= \; & \frac{\langle 0 \star w_3 w_4 \rangle_r - \langle 0 \star w_4 \rangle_r}{r^{|w_2|+k|w_3|}},
\end{aligned}$$

the same argument holds for concluding. ■

First, we consider a set $S \subseteq [0, 1]$ that is both r- and s- recognizable, or both weakly r- and weakly s-recognizable, for bases $r, s \in \mathbb{N}_{>1}$. We use the structure of an automaton recognizing in base r its set of boundary points to extract a converging sequence of boundary points, and then to build a set S' that admits a sequence of boundary points that converges to 0.

Lemma 5.9 *Let $r, s \in \mathbb{N}_{>1}$ be bases, and $S \subseteq [0, 1]$ be a set that is both (resp. weakly) r- and s-recognizable and that has infinitely many boundary points. There exists a set $S' \subseteq [0, 1]$ that is both (resp. weakly) r- and s-recognizable and that admits an infinite sequence of boundary points converging to* 0.

Proof Since weakly r-recognizability implies r-recognizability, the set S is r-recognizable. By Lemma 5.6, the set $B(S)$ of its boundary points is also r-recognizable. Suppose that $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ is a (deterministic) Muller RVA recognizing it. By assumption, the set $B(S) \subseteq [0, 1]$ is infinite. Hence, the language $L_1 \subseteq \{0, 1, \ldots, r-1\}^\omega$ defined by

$$L_1 = L(\delta^*(q_0, 0\star))$$

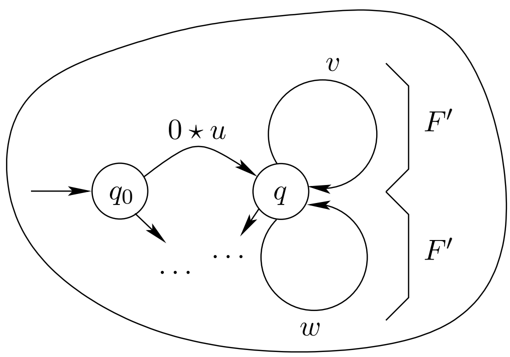

Figure 5.1: Illustration of Lemma 5.9 (1).

contains infinitely many infinite words.

Recall that a word w is accepted by the Muller automaton $\mathcal{A}$ if there exists a set $F' \in F$ equal to the set of states visited infinitely many times by the path reading w in $\mathcal{A}$. Since L_1 is infinite, there exists a set $F' \in F$ such that infinitely many words of L_1 have F' as the set of states appearing infinitely many times during their reading from the state $\delta^*(q_0, 0\star)$ in $\mathcal{A}$.

If F' contains more than one cycle visiting all the states of F', i.e., if there exist labels $v, w \in \{0, 1, \ldots, r-1\}^+$ and a state $q \in F'$ such that v is not a prefix of w, w is not a prefix of v, the set of states visited by the reading of v (resp. w) from q in $\mathcal{A}$ are both equal to F', and $\delta^*(q, v) = \delta^*(q, w) = q$, then all the words of the language

$$0 \star uv^* w^\omega,$$

where u is a word such that $\delta^*(q_0, 0 \star u) = q$, are accepted by $\mathcal{A}$. This situation is illustrated in Figure 5.1.

Otherwise, since L_1 is infinite, there exist states $q \in F', q' \notin F'$ and words $u, v, t, w \in \{0, 1, \ldots, r-1\}^*$ such that $v \neq \varepsilon$, $w \neq \varepsilon$, $\delta^*(q_0, 0 \star u) = q'$, $\delta^*(q', v) = q'$, $\delta^*(q', t) = q$, $\delta^*(q, w) = q$, and the set of states visited by the reading of w from q in $\mathcal{A}$ is equal to F'. Hence, all the words in the infinite language

$$0 \star uv^* tw^\omega$$

are accepted by $\mathcal{A}$, as shown if Figure 5.2.

In both cases, we thus extracted from $\mathcal{A}$ a language

$$L_2 = 0 \star uv^* tw^\omega,$$

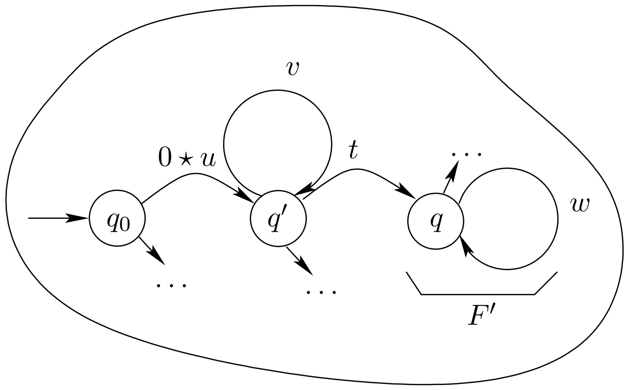

Figure 5.2: Illustration of Lemma 5.9 (2).

where u, t are words of $\{0, 1, \ldots, r-1\}^*$, v, w are words of $\{0, 1, \ldots, r-1\}^+$, and $0 \star utw^\omega \neq 0 \star uvtw^\omega$. Since both these two words end in w^ω, they cannot be dual encodings of the same number and, using Lemma 5.8, the boundary points encoded by the elements of L_2 are distinct.

We then define

$$y = \langle 0 \star uv^\omega \rangle_r$$

and, for each $k \in \mathbb{N}_{>0}$,

$$y_k = \langle 0 \star uv^k tw^\omega \rangle_r.$$

By Theorem 2.51, the numbers $y, y_1, y_2, y_3, \ldots$ are rationals.

The sequence $y_1, y_2, y_3, \ldots \in \mathbb{Q}^\omega$ forms an infinite sequence of distinct boundary points of S, converging to $y \in \mathbb{Q}$.

If we have $y_k > y$ for infinitely many k, then we define

$$S' = (S - y) \cap [0, 1].$$

Otherwise, we define

$$S' = (-S + y) \cap [0, 1].$$

From Theorems 2.65 and 3.35, the set S' is both (resp. weakly) r- and s-recognizable. Moreover, this set admits an infinite sequence of distinct boundary points that converges to 0. ■

Lemma 5.9 exhibits the existence of a r- and s-recognizable set S' that admits an infinite sequence of boundary points that converges to 0. From these characteristics, the next lemma is aimed at deriving properties of the languages representing S' in base r and s.

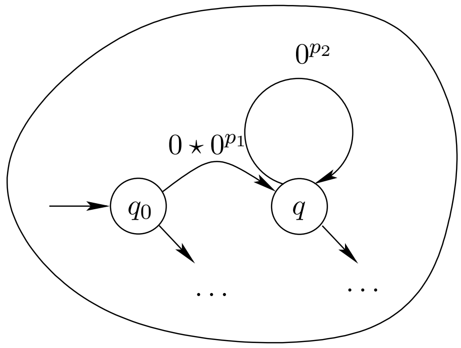

Figure 5.3: Illustration of Lemma 5.10.

Lemma 5.10 *Let $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq [0,1]$ be a set that is r-recognizable and that admits an infinite sequence of boundary points that converges to* 0*. There exist numbers $p_1 \in \mathbb{N}$ and $p_2 \in \mathbb{N}_{>0}$ such that, for all $w \in \{0, 1, \ldots, r-1\}^\omega$, we have*

$$\langle 0 \star 0^{p_1} w \rangle_r \in S \Leftrightarrow \langle 0 \star 0^{p_1+p_2} w \rangle_r \in S.$$

Proof Since S is r-recognizable, there exists a Muller RVA

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

recognizing S in base r. Since $\mathcal{A}$ has finitely many states, the path π of $\mathcal{A}$ reading 0^ω from the state $\delta^*(q_0, 0\star)$ visits two times the same state $q \in Q$: Let 0^{p_2}, with $p_2 \in \mathbb{N}_{>0}$, be a word such that $\delta^*(q, 0^{p_2}) = q$. Since the state q is visited by the path π, there exists a word 0^{p_1}, with $p_1 \in \mathbb{N}$, such that $\delta^*(0\star 0^{p_1}) = q$. Hence, the path π_0 of $\mathcal{A}$ reading $0\star 0^\omega$ is composed of a prefix labeled by $0\star$, followed by a finite path of length p_1, and finally followed by a cycle of length p_2, as illustrated in Figure 5.3.

It follows that a word of the form $0 \star 0^{p_1} w$, with $w \in \{0, 1, \ldots, r-1\}^\omega$, is accepted by $\mathcal{A}$ if and only if the word $0 \star 0^{p_1} 0^{p_2} w = 0 \star 0^{p_1+p_2}$ is accepted as well. Since $\mathcal{A}$ recognizes the r-encodings of S, we have $\langle 0 \star 0^{p_1} w \rangle_r \in S$ if and only if $\langle 0 \star 0^{p_1+p_2} w \rangle_r \in S$. ■

In the proof of Lemma 5.10, we discovered, in the automaton $\mathcal{A}$, the existence of a cycle rooted in q and labeled by a non-empty sequence 0^{p_2}. Intuitively, if the state reached after the reading of $0\star$ in $\mathcal{A}$ is q, then the accepted words are not influenced by the insertion of p_2 zero digits in the front of their encodings. Since the insertion of the digit 0 in a r-encoding of the form $0\star w$, in such a way that this encoding becomes $0\star 0w$, is equivalent to dividing the value of the number it represents by r, inserting those p_2 zero digits amounts to dividing the values of the represented numbers by r^{p_2}. In this case, the represented set is thus, in some sense, invariant by the multiplication of r^{p_2}, which leads to the following definition.

Definition 5.11 *Let $D \subseteq \mathbb{R}$ be a domain, and let $f \in \mathbb{R}_{>0}$. A set $S \subseteq D$ is* f-product-stable in the domain D *if and only if for all $x \in D$ such that $fx \in D$, we have*

$$x \in S \Leftrightarrow fx \in S.$$

□

We will follow this intuitive idea to build, from a r- and s-recognizable set S admitting an infinite sequence of boundary points converging to 0, a set being product-stable with respect to two values that respectively depend on the bases r and s. Furthermore, this new set will inherit the properties of S. The properties of product-stability in both bases r and s will be intensively used in the remaining of this chapter.

Lemma 5.12 *Let $S \subseteq [0,1]$ be a set that admits an infinite sequence of boundary points converging to 0, and let $v \in \mathbb{R}_{>0}$. The set $vS \cap [0,1]$ also admits an infinite sequence of boundary points converging to 0.*

Proof First, remark that if x is a boundary point of S, then vx is a boundary point of vS.

Since S admits an infinite sequence of distinct boundary points converging to 0, its subsequence containing the boundary points x such that $\frac{x}{v} \in]0,1[$ is still infinite and the points vx are boundary points of $vS \cap [0,1]$.

■

Theorem 5.13 *Let $r, s \in \mathbb{N}_{>1}$ be bases, and $S \subseteq [0,1]$ be a set that is both (resp. weakly) r- and s-recognizable and that admits an infinite sequence of boundary points converging to 0. There exist $i, j \in \mathbb{N}_{>0}$, and a set $S' \subseteq [0,1]$ that is both (resp. weakly) r- and s-recognizable, both r^i- and s^j-product-stable in $[0,1]$, and that admits an infinite sequence of boundary points converging to 0.*

Proof Let $\mathcal{A}_r$ and $\mathcal{A}_s$ be (resp. weak) RVA recognizing S in the respective bases r and s. By Lemma 5.10, there exist $p_{1,r}, p_{1,s} \in \mathbb{N}$, and $p_{2,r}, p_{2,s} \in \mathbb{N}_{>0}$ such that

$$\langle 0 \star 0^{p_{1,r}} w \rangle_r \in S \Leftrightarrow \langle 0 \star 0^{p_{1,r}+p_{2,r}} w \rangle_r \in S$$

and

$$\langle 0 \star 0^{p_{1,s}} w \rangle_s \in S \Leftrightarrow \langle 0 \star 0^{p_{1,s}+p_{2,s}} w \rangle_s \in S.$$

Let S_1 be the set

$$S_1 = r^{p_{1,r}} S \cap [0,1].$$

This set has the following properties :

- S_1 *admits an infinite sequence of boundary points converging to* 0. This is a consequence of the definition of S_1 and of Lemma 5.12.
- By Theorems 2.65 and 3.35, S_1 *is both (resp. weakly)* r*- and* s*-recognizable.*
- S_1 *is* $r^{p_{2,r}}$*-product-stable in* $[0,1]$. Indeed, for all $x \in [0,1]$, we have

$$\begin{aligned}
& x \in S_1 \\
\Leftrightarrow\ & (\exists w \in \{0,1,\ldots,r-1\}^\omega)(x = \langle 0 \star w \rangle_r \wedge \langle 0 \star w \rangle_r \in S_1) \\
\Leftrightarrow\ & (\exists w \in \{0,1,\ldots,r-1\}^\omega)(x = \langle 0 \star w \rangle_r \wedge \frac{\langle 0 \star w \rangle_r}{r^{p_{1,r}}} \in S) \\
\Leftrightarrow\ & (\exists w \in \{0,1,\ldots,r-1\}^\omega)(x = \langle 0 \star w \rangle_r \wedge \langle 0 \star 0^{p_{1,r}} w \rangle_r \in S) \\
\Leftrightarrow\ & (\exists w \in \{0,1,\ldots,r-1\}^\omega)(x = \langle 0 \star w \rangle_r \wedge \langle 0 \star 0^{p_{1,r}+p_{2,r}} w \rangle_r \in S) \\
\Leftrightarrow\ & (\exists w \in \{0,1,\ldots,r-1\}^\omega)(x = \langle 0 \star w \rangle_r \wedge \frac{\langle 0 \star 0^{p_{2,r}} w \rangle_r}{r^{p_{1,r}}} \in S) \\
\Leftrightarrow\ & (\exists w \in \{0,1,\ldots,r-1\}^\omega)(x = \langle 0 \star w \rangle_r \wedge \langle 0 \star 0^{p_{2,r}} w \rangle_r \in S_1) \\
\Leftrightarrow\ & \frac{x}{r^{p_{2,r}}} \in S_1.
\end{aligned}$$

Now, we define S_2 as the set

$$S_2 = s^{p_{1,s}} S_1 \cap [0,1].$$

S_2 also has also several recognizability and product-stability properties, and admits an infinite sequence of boundary points converging to 0 :

- By Lemma 5.12, the set S_2 inherits from S_1 the fact that *it admits an infinite sequence of boundary points converging to* 0.
- By Theorems 2.65 and 3.35, S_2 *is both (resp. weakly)* r*- and* s*-recognizable.*
- S_2 *is* $r^{p_{2,r}}$*-product-stable in* $[0,1]$. Indeed, for all $x \in [0,1]$, we have

$$\begin{aligned}
& x \in S_2 \\
\Leftrightarrow\ & \frac{x}{s^{p_{1,s}}} \in S_1 \\
\Leftrightarrow\ & \frac{x}{s^{p_{1,s}} r^{p_{2,r}}} \in S_1 \\
\Leftrightarrow\ & \frac{x}{r^{p_{2,r}}} \in S_2.
\end{aligned}$$

- *S_2 is $s^{p_{2,s}}$-product-stable in* $[0,1]$. For proving this property, recall that

$$S_2 = r^{p_{1,r}} s^{p_{1,s}} S \cap [0,1].$$

Hence, the set S_2 can alternatively be obtained by first defining

$$S_3 = s^{p_{1,s}} S \cap [0,1],$$

which is both (resp. weakly) r- and s-recognizable by Theorems 2.65 and 3.35. Then, one has

$$S_2 = r^{p_{1,r}} S_3 \cap [0,1].$$

By a similar reasoning in base s, we get that S_2 is $s^{p_{2,s}}$-product-stable in $[0,1]$.

To conclude, we define S', i and j by

$$S' = S_2,$$

$$i = p_{2,r},$$

and

$$j = p_{2,s}.$$

■

Since, by Theorems 2.66 and 3.32, (resp. weakly) r-recognizability is equivalent to (resp. weakly) r^ℓ-recognizability (for all $\ell \in \mathbb{N}_{>0}$), the bases r and s can be replaced by $r' = r^i$ and $s' = s^j$. Hence, we have the following corollary.

Corollary 5.14 *Let $r, s \in \mathbb{N}_{>1}$ be bases, and $S \subseteq [0,1]$ be a set that is both (resp. weakly) r- and s-recognizable and that admits an infinite sequence of boundary points converging to* 0. *There exist powers $r' = r^i$, $s' = s^j$ of r and s, with $i, j \in \mathbb{N}_{>0}$, and a set $S' \subseteq [0,1]$ that is both (resp. weakly) r'- and s'-recognizable, both r'- and s'-product-stable in* $[0,1]$, *and that admits an infinite sequence of boundary points converging to* 0.

Lemma 5.9 and Corollary 5.14 can now be combined to produce the main theorem of this section : Under the hypothesis of r- and s-recognizability of a set S, as well as the infiniteness of its set of boundary points, we infer the existence of a r'- and s'-recognizable set S' which is r'- and s'-product-stable, and that admits infinitely many boundary points.

Theorem 5.15 *Let $r, s \in \mathbb{N}_{>1}$ be multiplicatively independent bases, and $S \subseteq [0,1]$ be a set that is both (resp. weakly) r- and s- recognizable and that admits infinitely many boundary points. There exist powers $r' = r^i$, $s' = s^j$ of r and s, with $i, j \in \mathbb{N}_{>0}$, and a set $S' \subseteq [0,1]$ that is both (resp. weakly) r'- and s'-recognizable, both r'- and s'-product-stable in $[0,1]$, and that admits infinitely many boundary points.*

Proof This is a direct consequence of Lemma 5.9 and Corollary 5.14. ■

5.4 Generalization of Cobham's theorem to weak deterministic RVA

In this section, we use the observations of Section 5.3, as well as the reduction presented in Section 5.2, for proving that a set that is simultaneously weakly recognizable in two multiplicatively independent bases r and s is necessarily definable in the theory $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. The multiplicatively independence of the bases r and s will obviously be crucial in our arguments, since, by Corollary 3.33, a set is weakly recognizable in two multiplicatively dependent bases if and only if it is weakly recognizable in only one of them.

We begin by some technical lemmas. The first of them states that two bases r and s are multiplicatively independent if and only if $\log_s r$ is irrational.

Lemma 5.16 *Two numbers $r, s \in \mathbb{N}_{>1}$ are multiplicatively independent if and only if $\log_s r \in \mathbb{R} \setminus \mathbb{Q}$.*

Proof By definition, the numbers r and s multiplicatively independent if and only if $(\forall i, j \in \mathbb{N}_{>0})(r^i \neq s^j)$. This is true if and only if $(\forall i, j \in \mathbb{N}_{>0})(r \neq s^{\frac{j}{i}})$, or, equivalently, if and only if $(\forall i, j \in \mathbb{N}_{>0})(\log_s r \neq \frac{j}{i})$, that is, if and only if $\log_s r$ is negative or irrational.

But, since r and s are strictly greater than 1, we have $\log_s r > 0$. Hence, r and s are multiplicatively independent if and only if $\log_s r \in \mathbb{R} \setminus \mathbb{Q}$. ■

Now, let us prove that, for a given irrational number $x \in \mathbb{R} \setminus \mathbb{Q}$ and a real number $y \in [0,1]$, there exists a multiple mx of x, with $m \in \mathbb{Z}$, such that its fractional part $mx - \lfloor mx \rfloor$ is arbitrarily close to y. In other words, we will establish that the set of fractional parts of the multiples of x is dense in $[0,1]$. Actually, this is a consequence of Kronecker's approximation theorem [HW85], but we give here a direct proof.

Lemma 5.17 *Let $x \in \mathbb{R} \setminus \mathbb{Q}$, and $y \in \mathbb{R}$. For every $\varepsilon \in \mathbb{R}_{>0}$, there exists $m \in \mathbb{Z}$ such that*

$$(mx + y) - \lfloor mx + y \rfloor < \varepsilon.$$

Proof The set

$$S = \{mx + y - \lfloor mx + y \rfloor \mid m \in \mathbb{Z}\}$$

represents the set of fractional parts of the numbers $mx + y$, with $m \in \mathbb{Z}$. Hence, it is a subset of $[0, 1]$.

Moreover, S is infinite. Indeed, otherwise, there would exist $m_1, m_2 \in \mathbb{Z}$ such that $m_1 \neq m_2$ and the fractional parts of $(m_1 x + y)$ and $(m_2 x + y)$ are identical. Hence, the number $(m_1 x + y) - (m_2 x + y)$ would be an integer $k \in \mathbb{N}$, and we would have

$$x = \frac{k}{m_1 - m_2},$$

which contradicts the fact that x is irrational.

For each $\varepsilon \in \mathbb{R}_{>0}$, one can thus extract two values $m_1, m_2 \in \mathbb{Z}$ such that

$$0 < ((m_1 x + y) - \lfloor m_1 x + y \rfloor) - ((m_2 x + y) - \lfloor m_2 x + y \rfloor) < \varepsilon.$$

Let $m' = m_1 - m_2$. We then have

$$0 < m'x - \lfloor m'x \rfloor < \varepsilon. \tag{5.1}$$

We have to prove that there exists $m \in \mathbb{Z}$ such that the fractional part of $mx + y$, i.e., $(mx + y) - \lfloor mx + y \rfloor$, is strictly smaller than ε. Hence, the integer part of y does not have matter. Let $y_F = y - \lfloor y \rfloor$. It thus suffices to prove that there exists an integer $m \in \mathbb{Z}$ such that

$$(mx + y_F) - \lfloor mx + y_F \rfloor < \varepsilon.$$

For every $k \in \mathbb{N}$ such that $k(m'x - \lfloor m'x \rfloor) < 1$, we have, by Equation 5.1,

$$0 < km'x - \lfloor km'x \rfloor < k\varepsilon,$$

as illustrated in Figure 5.4.

By choosing

$$k = \left\lfloor \frac{y_F}{m'x - \lfloor m'x \rfloor} \right\rfloor,$$

this yields

$$y_F - (km'x - \lfloor km'x \rfloor) < \varepsilon.$$

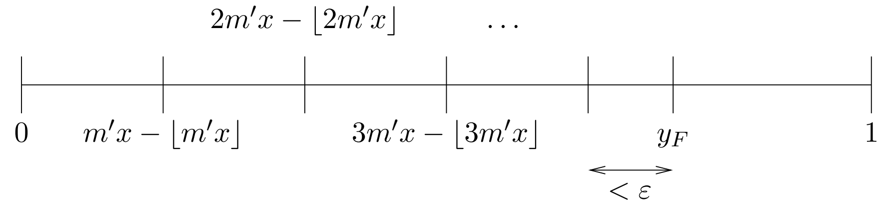

Figure 5.4: Illustration of Lemma 5.17.

We conclude by choosing

$$m = -km'.$$

■

The next lemma is aimed at proving that, when r and s are two multiplicatively independent bases, any open interval of $\mathbb{R}_{>0}$ contains some number of the form $\frac{r^i}{s^j}$, with $i, j \in \mathbb{N}_{>0}$. The proof is inspired from [Per90].

Lemma 5.18 *Let $r, s \in \mathbb{N}_{>1}$ be two multiplicatively independent numbers, and let $x, \varepsilon \in \mathbb{R}_{>0}$. There exist $i, j \in \mathbb{N}_{>0}$ such that*

$$0 < x < \frac{r^i}{s^j} < x + \varepsilon.$$

Proof We have to show that

$$x < \frac{r^i}{s^j} < x + \varepsilon$$

for a given pair of numbers $i, j \in \mathbb{N}_{>0}$. Defining $\varepsilon' = \log_s \left(1 + \frac{\varepsilon}{x}\right)$, this constraint becomes

$$x < \frac{r^i}{s^j} < xs^{\varepsilon'}.$$

It remains to prove that there exist $i, j \in \mathbb{N}_{>0}$ such that

$$\log_s(x) < i \log_s r - j < \log_s(x) + \varepsilon',$$

which can be rewritten as

$$j < i \log_s r - \log_s(x) < j + \varepsilon'.$$

Since r and s are multiplicatively independent, $\log_s r$ is irrational by Lemma 5.16.

The conclusion directly follows from Lemma 5.17. ■

We are now able to tackle the generalization of Cobham's theorem for weak deterministic automata.

Theorem 5.19 *Let $r, s \in \mathbb{N}_{>1}$ be multiplicatively independent bases, and $S \subseteq [0, 1]$. If S is simultaneously weakly r- and weakly s-recognizable, then S is definable in $\langle \mathbb{R}, +, <, 1\rangle$.*

Proof By Theorem 5.7, it suffices to prove that the set $B(S)$ of boundary points of S is finite.

By contradiction, suppose that $B(S)$ is infinite.

By Theorem 5.15, there exist powers $r' = r^i$, $s' = s^j$ of r and s, with $i, j \in \mathbb{N}_{>0}$, and a set $S' \subseteq [0, 1]$ that is both weakly r'- and weakly s'-recognizable, both r'- and s'-product-stable in $[0, 1]$, and such that $B(S')$ is infinite.

Remark that the bases r' and s' are multiplicatively independent : Otherwise, we would have $r'^{i'} = s'^{j'}$ for some $i', j' \in \mathbb{N}_{>0}$, that is, $r^{ii'} = s^{jj'}$; hence, r and s would be multiplicatively dependent.

Since S' is both r'- and s'-product-stable in $[0, 1]$ and S' admits infinitely many boundary points, there exist $x, y \in \,]0, 1]$ such that $x \in S'$ and $y \notin S'$. Moreover, for every $i'', j'' \in \mathbb{Z}$ such that $r'^{i''} s'^{j''} x \in \,]0, 1]$, we have $r'^{i''} s'^{j''} x \in S'$. Similarly, for every $i'', j'' \in \mathbb{Z}$ such that $r'^{i''} s'^{j''} y \in \,]0, 1]$, we have $r'^{i''} s'^{j''} y \notin S'$.

Let z be an arbitrary point in the open interval $]0, 1[$, and let $\varepsilon \in \mathbb{R}_{>0}$ be such that $z - \varepsilon > 0$ and $z + \varepsilon < 1$.

By Lemma 5.18, and for each $\alpha \in \{x, y\}$, there exist $i''', j''' \in \mathbb{N}_{>0}$ such that

$$\frac{z - \varepsilon}{\alpha} < \frac{r^{i'''}}{s^{j'''}} < \frac{z + \varepsilon}{\alpha},$$

which can be rewritten as

$$z - \varepsilon < \frac{r^{i'''}}{s^{j'''}} \alpha < z + \varepsilon.$$

By the definition of ε, we thus have

$$0 < z - \varepsilon < \frac{r^{i'''}}{s^{j'''}} \alpha < z + \varepsilon < 1.$$

Hence, every neighborhood $N_\varepsilon(x)$ of x contains one point of S' as well as one from $[0,1] \setminus S'$. This property implies that S' satisfies the dense oscillating sequence property introduced in Definition 3.12, and therefore, by Theorem 3.38, cannot be recognized by a weak deterministic RVA. This is a contradiction with the fact that S' is weakly r'- (and weakly s'-) recognizable. Hence, S is definable in $\langle \mathbb{R}, +, <, 1 \rangle$. ■

By combining Theorem 5.19 with Theorem 4.16, we move from the domain $[0,1]$ to the full domain $\mathbb{R}$, and we get a proof of Claim 5.1.

Theorem 5.20 *Let $r, s \in \mathbb{N}_{>1}$ be multiplicatively independent bases, and $S \subseteq \mathbb{R}$. If S is simultaneously weakly r- and weakly s-recognizable, then S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.*

Proof This is a direct corollary of Theorems 4.16 and 5.19. ■

5.5 Counter-example for Muller RVA

We now show that Theorem 5.20 does not directly generalize to non-weak recognizability. Indeed, a set can be recognizable in two multiplicatively independent bases without being definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. This property is established in the following theorem.

In this theorem, restating Claim 5.2, we consider bases sharing the same set of prime factors. Remark that there exist multiplicatively independent bases having this property : For instance, the bases 6 and 12 are multiplicatively independent and share the same set of prime factors.

Theorem 5.21 *Let $r, s \in \mathbb{N}_{>1}$ be two bases that share the same set of prime factors. There exists a set $S \subseteq \mathbb{R}$ that is both r- and s-recognizable, and that is not definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.*

Proof Let $\{f_1, f_2, \ldots, f_n\}$ be the set of prime factors of r and s. A counter-example is provided by the set

$$S = \left\{ \frac{p}{f_1^{i_1} f_2^{i_2} \ldots f_n^{i_n}} \mid p \in \mathbb{Z}, i_1, i_2, \ldots, i_n \in \mathbb{N} \right\}.$$

By Corollary 2.54, and whatever the base $t \in \{r, s\}$ is, the set S is encoded by the language

$$L_t = \{0, t-1\}\{0, 1, \ldots, t-1\}^* \star \{0, 1, \ldots, t-1\}^*(0^\omega \cup (t-1)^\omega),$$

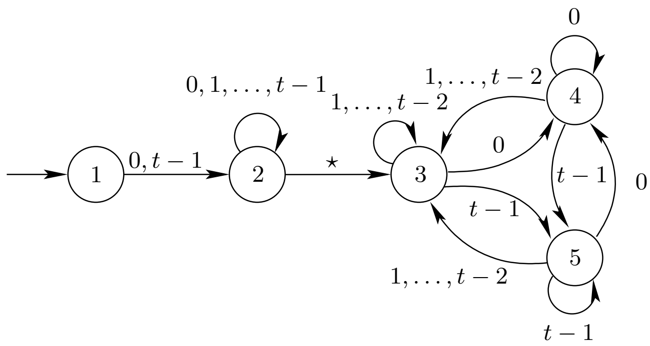

Figure 5.5: Muller RVA, with acceptance condition $\{\{4\},\{5\}\}$, accepting the language $\{0,t-1\}\{0,1,\ldots,t-1\}^* \star \{0,1,\ldots,t-1\}^*(0^\omega \cup (t-1)^\omega)$.

i.e., the set S contains the numbers that admit dual encodings.

The language L_t is accepted by the Muller automaton depicted in Figure 5.5, where the acceptance condition is $\{\{4\},\{5\}\}$. Hence, S is both r- and s-recognizable.

Suppose, by contradiction, that S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. Then, it is weakly t-recognizable in any base $t \in \{r, s\}$ thanks to Theorem 3.31. By Theorem 3.38 and since S satisfies the dense oscillating sequence property, this leads to a contradiction. ■

5.6 Bases with different sets of prime factors and sum-stability properties

In this section, we consider two bases $r, s \in \mathbb{N}_{>1}$ that do not share the same set of prime factors. Since this property implies that r and s are multiplicatively independent, we know by Theorem 5.20 that any subset of $\mathbb{R}$ that is simultaneously weakly r- and weakly s-recognizable must be definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

Our goal is to show that a subset of $\mathbb{R}$ that is both r- and s-recognizable is necessarily definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. Recall that, as shown is Section 5.5, this result does not extend to pairs of bases that are multiplicatively independent but that share the same sets of prime factors.

In the same way as we did for weak recognizability, we will use Theorems 4.15 and 5.7, and suppose the existence of a set $S \subseteq [0, 1]$ being both r-

and s- recognizable and admitting infinitely many boundary points. Starting from this hypothesis, we will show that there exists a set $S' \subseteq \mathbb{R}$ having in addition product-stability properties, as well as *sum-stability* ones (this notion will be formally defined in the sequel). Next, we will exploit those properties to obtain strong results about the structure of the set of numbers t for which S' is t-sum-stable.

These results will be used in Section 5.7 for showing that such a set cannot exist, i.e., by Theorems 4.15 and 5.7, that each set S simultaneously r- and s-recognizable is necessarily definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

We thus consider two bases $r, s \in \mathbb{N}_{>1}$ with different sets of prime factors, and a set $S \subseteq [0, 1]$ that is both r- and s-recognizable, and that admits infinitely many boundary points. We begin by reusing the results of Section 5.3. By Theorem 5.15, which is valid for weak recognizability and also for recognizability, there exist bases $r', s' \in \mathbb{N}_{>1}$ with different sets of prime factors (actually, r' and s' are respectively powers of r and s), and a set $S' \subseteq [0, 1]$ that is both r'- and s'-recognizable, both r'- and s'-product-stable in $[0, 1]$, and that admits infinitely many boundary points.

We thus suppose that such a set S' exists. Our first strategy consists in exploiting Cobham's theorem so as to derive additional properties of this set.

The initial step is to build a set that coincides with S' over $[0, 1]$, shares the same recognizability and product-stability properties, and contains numbers with non-trivial integer parts. For this purpose, we begin by proving the following lemma, expressing that, if a set $S \subseteq [0, 1]$ is r-recognizable, then the set $\{r^k x \mid x \in S \wedge k \in \mathbb{N}\}$ is r-recognizable as well.

Lemma 5.22 *Let $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq [0, 1]$ be a r-recognizable set. The set $S' = \{r^k x \mid x \in S \wedge k \in \mathbb{N}\}$ is r-recognizable as well.*

Proof Since S is r-recognizable, it is accepted by a (possibly non-deterministic) Büchi RVA

$$\mathcal{A} = (Q, \Sigma, \Delta, q_0, F),$$

whose transitions are supposed to have labels of lengths not greater than 1. This assumption can be done without loss of generality since, for each transition (q, w, q') such that $|w| > 1$, one can add $|w| - 1$ new states $q_1, q_2, \ldots, q_{|w|-1}$ and replace (q, w', q') by the sequence of transitions

$$(q, \sigma_1, q_1), (q_1, \sigma_2, q_2), (q_2, \sigma_3, q_3), \ldots, (q_{|w|-1}, \sigma_{|w|}, q'),$$

where $\sigma_1, \sigma_2, \ldots, \sigma_{|w|} \in \Sigma$ are such that $w = \sigma_1 \sigma_2 \ldots \sigma_{|w|}$.

Let $Q_0 \subset Q$ be the set of states of $\mathcal{A}$ that are reached by the reading of a word of $0^+\star$, i.e.,

$$Q_0 = \{q \in Q \mid (\exists k \in \mathbb{N}_{>0})((q_0, 0^k\star, q) \in \Delta^*)\},$$

let $Q_1 \subset Q$ be the set of states that can be reached, non necessarily directly, from the states in Q_0, i.e.,

$$Q_1 = \{q \in Q \mid (\exists k \in \mathbb{N}_{>0})(\exists w \in \{0, 1, \ldots, r-1\}^*)((q_0, 0^k \star w, q) \in \Delta^*)\},$$

and let $Q_2 \subset Q$ be the set of states that are reached by the reading of a word of 0^* from a state of Q_0, i.e.,

$$Q_2 = \{q \in Q \mid (\exists q' \in Q_0)(\exists k \in \mathbb{N})((q', 0^k, q) \in \Delta^*)\}.$$

A non-deterministic Büchi RVA $\mathcal{A}'$ representing S' can intuitively be built from $\mathcal{A}$ by delaying arbitrarily the reading of the symbol $\star$. This construction can be achieved by copying the set of states Q_1 and, from each state of the copy, creating an outgoing transition labeled by $\star$, leading to the corresponding state of the original copy. The automaton $\mathcal{A}'$ is composed of those two copies of Q_1, and of two additional states that are aimed at ensuring that valid encodings are read.

Furthermore, due to the necessity of handling dual encodings, if $0 \in S$ (resp. $1 \in S$), then the languages $0^+10^* \star 0^\omega$ (resp. $(r-1)^+ \star (r-1)^\omega$) have to be accepted by $\mathcal{A}'$.

Following this idea, we get

$$\mathcal{A}' = (Q', \Sigma, \Delta', s_1, F'),$$

where

- $Q' = (Q_1 \times \{1, 2\}) \cup \{s_1, s_2\} \cup Q_3 \cup Q_4$ (s_1 and s_2 are new states);
- if $0 \in S$ (resp. $0 \notin S$), then $Q_3 = \{s_3, s_4\}$, where s_3 and s_4 are new states (resp. $Q_3 = \emptyset$);
- if $1 \in S$ (resp. $1 \notin S$), then $Q_4 = \{s_5, s_6\}$, where s_5 and s_6 are new states (resp. $Q_4 = \emptyset$);
- $\Delta' = \Delta_1 \cup \Delta_2 \cup \Delta_3$;
- if $0 \in S$ (resp. $0 \notin S$), then

  $$\Delta_2 = \{(s_1, r-1, s_3), (s_3, r-1, s_3), (s_3, \star, s_4), (s_4, r-1, s_4)\}$$

  (resp. $\Delta_2 = \emptyset$);

- if $1 \in S$ (resp. $1 \notin S$), then

$$\Delta_3 = \{(s_2, 1, s_5), (s_5, 0, s_5), (s_5, \star, s_6), (s_6, 0, s_6)\}$$

(resp. $\Delta_3 = \emptyset$);

- Δ_1 is defined by

$$\begin{array}{rcl} \Delta_1 & = & \{(s_1, 0, s_2), (s_2, 0, s_2)\} \\ & & \cup \bigcup_{q \in Q_2} \{(s_2, \varepsilon, q)\} \\ & & \cup \bigcup_{(q,\sigma,q') \in \Delta \wedge q,q' \in Q_1} \{((q,1), \sigma, (q',1)), ((q,2), \sigma, (q',2))\} \\ & & \cup \bigcup_{q \in Q_1} \{((q,1), \star, (q,2))\}; \end{array}$$

- $F' = ((F \cap Q_1) \times \{2\}) \cup F_1 \cup F_2$;
- if $0 \in S$ (resp. $0 \notin S$), then $F_1 = \{s_4\}$ (resp. $F_1 = \emptyset$);
- if $1 \in S$ (resp. $1 \notin S$), then $F_2 = \{s_6\}$ (resp. $F_2 = \emptyset$).

This construction is illustrated in Figure 5.6.

Since, for words $w_1 \in \{0, 1, \ldots, r-1\}^*$, $w_2 \in \{0, 1, \ldots, r-1\}^\omega$, and a symbol $\sigma \in \{0, 1, \ldots, r-1\}$, we have

$$r\langle 0w_1 \star \sigma w_2\rangle_r = \langle 0w_1\sigma \star w_2\rangle_r,$$

the automaton $\mathcal{A}'$ accepts r-encodings of the set $S' = \{r^k x \mid x \in S \wedge k \in \mathbb{N}\}$. Moreover, thanks to the states s_2, s_3, s_4, s_5, s_6 and their outgoing transitions, it accepts all their encodings. Hence S' is r-recognizable. ∎

Using Lemma 5.22, we now are able to prove the initial step of this section.

Lemma 5.23 *Let $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors, and $S \subseteq [0,1]$ be a set that is both r- and s-recognizable, both r- and s-product-stable in $[0,1]$, and that admits infinitely many boundary points. There exists a set $S' \subseteq \mathbb{R}_{\geq 0}$ that is both r- and s-recognizable, both r- and s-product-stable in $\mathbb{R}_{\geq 0}$, and that admits infinitely many boundary points.*

Proof Let $S' = \{r^k x \mid x \in S \wedge k \in \mathbb{N}\}$.

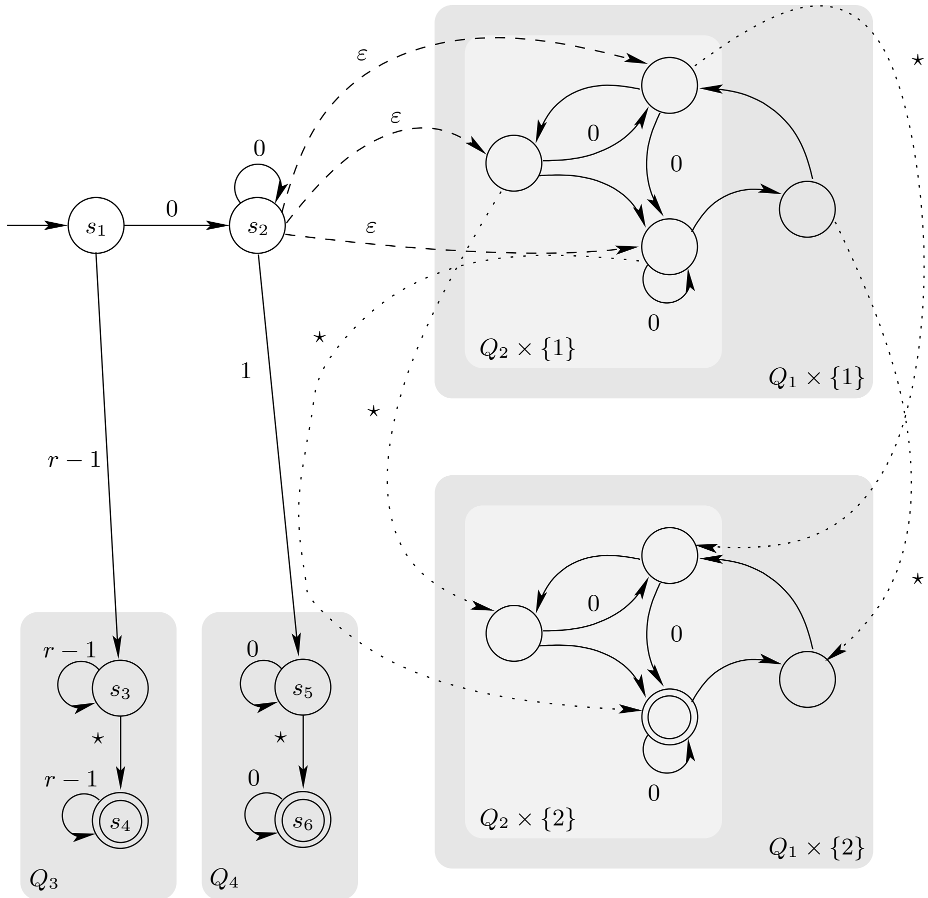

Figure 5.6: Construction of Lemma 5.22.

This set is clearly r-product-stable in $\mathbb{R}_{\geq 0}$. Indeed, let $v \in \mathbb{R}_{\geq 0}$. We have $v \in S'$ if and only if $(\exists v' \in S)(\exists k \in \mathbb{N})(v = r^k v')$. Hence, $v \in S'$ if and only if $rv = r^{k+1}v' \in S'$, by definition of S'.

By Lemma 5.22, S' is r-recognizable.

We have $S' \cap [0,1] = S$. Indeed, on the one hand, let $x \in S' \cap [0,1]$. We have $(\exists k \in \mathbb{N})(\exists x' \in S)(x = r^k x')$. Since $S \subseteq [0,1]$ is r-product-stable in $[0,1]$, we have $x = r^k x' \in S$. On the other hand, let $x \in S$. We have $x \in S' \cap [0,1]$ since $x = r^0 x$. As a corollary, the set S' admits infinitely many boundary points.

Remark that we have $S' = \{r^i s^j x \mid x \in S \wedge i, j \in \mathbb{Z}\}$: On the one hand, S' is trivially included in $\{r^i s^j x \mid x \in S \wedge i, j \in \mathbb{Z}\}$. On the other hand, let $x' \in \{r^i s^j x \mid x \in S \wedge i, j \in \mathbb{Z}\}$. There thus exist $x'' \in S$ and $i, j \in \mathbb{Z}$ such that $x' = r^i s^j x''$. Let $k \in \mathbb{N}$ be sufficiently large to have $i - k < 0$ and $r^{i-k} s^j x'' \in [0,1]$. Since S is r-product-stable in $[0,1]$, we have $r^{i-k} x'' \in S$. By the s-product-stability of S in $[0,1]$, we thus also have $r^{i-k} s^j x'' \in S$. Hence, we get $x' \in S'$ since x' can be expressed as $x' = r^k r^{i-k} s^j x''$.

The set S' can therefore be expressed as $S' = \{s^k x \mid x \in S \wedge k \in \mathbb{N}\}$. By the same reasoning as in base r, this set is s-recognizable, as well as s-product-stable in $\mathbb{R}_{\geq 0}$. ■

The next step consists in starting from the set $S' \subseteq \mathbb{R}_{\geq 0}$ we obtained in Lemma 5.23, and building a set being, in some sense, invariant by translation. Precisely, we will show the existence of a set which is 1-sum-stable in $\mathbb{R}_{>0}$. This notion is introduced in the next definition.

Definition 5.24 *Let $D \subseteq \mathbb{R}$ be a domain, and let $t \in \mathbb{R}$. A set $S \subseteq D$ is* t-sum-stable in the domain D *if and only if for all $x \in D$ such that $x+t \in D$, we have*

$$x \in S \Leftrightarrow x + t \in S.$$

□

Lemma 5.25 *Let $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors, and $S \subseteq [0,1]$ be a set that is both r- and s-recognizable, both r- and s-product-stable in $[0,1]$, and that admits infinitely many boundary points. There exists a set $S' \subseteq \mathbb{R}_{>0}$ that is both r- and s-recognizable, both r- and s-product-stable in $\mathbb{R}_{>0}$, 1-sum-stable in $\mathbb{R}_{>0}$, and that admits infinitely many boundary points.*

Proof By Lemma 5.23, there exists a set $S' \subseteq \mathbb{R}_{\geq 0}$ that is both r- and s-recognizable, both r- and s-product-stable in $\mathbb{R}_{\geq 0}$, and that admits infinitely many boundary points.

Let $S'' = S' \setminus \{0\}$. This set is r- and s-product-stable in $\mathbb{R}_{>0}$, and admits infinitely many boundary points. Moreover, by Theorem 2.65, S'' is both r- and s-recognizable.

By Theorem 4.13, the set S'' can be decomposed into a finite union

$$S'' = \bigcup_{i=1}^{m} (S_i^I + S_i^F),$$

where $m \in \mathbb{N}$ and, for each $i \in \{1, 2, \ldots, m\}$, the set $S_i^I \subseteq \mathbb{Z}$ is both r- and s-recognizable by NDDs, and the set $S_i^F \subseteq [0, 1]$ is both r- and s-recognizable.

Since S'' is a subset of $\mathbb{R}_{>0}$, the sets S_i^I can, without loss of generality, be supposed to be subsets of $\mathbb{N}$ instead of $\mathbb{Z}$. By Theorem 2.71, for each $i \in \{1, 2, \ldots, m\}$, the set S_i^I is definable in $\langle \mathbb{N}, + \rangle$. Since, by Corollary 2.5, such a set is ultimately periodic, there exists $k_i \in \mathbb{N}_{>0}$ for which

$$(\forall x \in \mathbb{N}_{\geq k_i})(x \in S_i^I \Leftrightarrow x + k_i \in S_i^I).$$

By defining $k = \text{lcm}_{i \in \{1,2,\ldots,m\}} k_i$, we obtain

$$(\forall x \in \mathbb{R}_{\geq k})(x \in S'' \Leftrightarrow x + k \in S'').$$

Let S''' be the set

$$S''' = \frac{1}{k} S''.$$

We have

$$(\forall x \in \mathbb{R}_{\geq 1})(x \in S''' \Leftrightarrow x + 1 \in S'''). \tag{5.2}$$

Let us show that S''' satisfies the expected properties.

By Theorem 2.65, S''' is both r- and s- recognizable.

S''' is r- and s-product-stable in $\mathbb{R}_{>0}$. Indeed, let $t \in \{r, s\}$, and $x \in \mathbb{R}_{>0}$ We have $x \in S''' \Leftrightarrow kx \in S'' \Leftrightarrow tkx \in S'' \Leftrightarrow tx \in S'''$.

S''' is 1-sum-stable in $\mathbb{R}_{>0}$. Indeed, for every $x \in \mathbb{R}_{\geq 1}$, we have $x \in S''' \Leftrightarrow x + 1 \in S'''$. For $x \in]0, 1[$, let $k' \in \mathbb{N}$ such that $r^{k'} x \geq 1$. Since S''' is r-product-stable, and by Equation 5.2, we have

$$x \in S''' \Leftrightarrow r^{k'} x \in S''' \Leftrightarrow r^{k'} x + \underbrace{1 + 1 + \ldots + 1}_{r^{k'}} \in S''' \Leftrightarrow x + 1 \in S'''.$$

Finally, S''' admits infinitely many boundary points, since this property holds for S''. ■

In the final steps of this section, we consider a set $S \subseteq \mathbb{R}_{>0}$ that satisfies the properties expressed in Lemma 5.25, i.e., a set $S \subseteq \mathbb{R}_{>0}$ that is both r- and s-recognizable, both r- and s-product-stable in $\mathbb{R}_{>0}$ and 1-sum-stable in $\mathbb{R}_{>0}$, and we proceed by characterizing the numbers $t \in \mathbb{R}$ for which S is t-sum-stable in $\mathbb{R}_{>0}$. The hypothesis on the prime factors of r and s is explicitly used in the sequel.

Definition 5.26 *Let $D \subseteq \mathbb{R}$ be a domain, and let $S \subseteq D$. The set of numbers $t \in \mathbb{R}$ such that S is t-sum-stable in D is denoted by $T_D(S)$.* □

Lemma 5.27 *Let $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq \mathbb{R}_{>0}$ be a r-recognizable set. The set $T_{\mathbb{R}_{>0}}(S)$ is r-recognizable as well.*

Proof Since S is r-recognizable, it is definable in $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$ by Theorem 2.65. Since $T_{\mathbb{R}_{>0}}(S)$ can be defined by

$$T_{\mathbb{R}_{>0}(S)} = \{t \in \mathbb{R} \mid (\forall x \in \mathbb{R}) \\ ((x > 0 \wedge x + t > 0) \Rightarrow (x \in S \Leftrightarrow x + t \in S))\},$$

it is definable in $\langle \mathbb{R}, \mathbb{Z}, +, <, X_r \rangle$ as well, and is therefore r-recognizable. ■

Lemma 5.28 *Let $S \subseteq \mathbb{R}_{>0}$. For every $t, u \in T_{\mathbb{R}_{>0}}(S)$ and $a \in \mathbb{Z}$, we have $at \in T_{\mathbb{R}_{>0}}(S)$ and $t + u \in T_{\mathbb{R}_{>0}}(S)$.*

Proof If $t \in T_{\mathbb{R}_{>0}}(S)$, then t can be added to (resp. subtracted from) $x \in \mathbb{R}_{>0}$, without modifying the fact that $x \in S$ or $x \notin S$, as far as $x + t > 0$ (resp. $x - t > 0$). This operation can be repeated an arbitrary number of times a.

Let $t, u \in T_{\mathbb{R}_{>0}}(S)$. If either $t, u \in \mathbb{R}_{>0}$ or $t, u \in \mathbb{R}_{<0}$, then $t + u$ belongs trivially to $T_{\mathbb{R}_{>0}}(S)$. If $x \in \mathbb{R}_{>0}$, $t \in \mathbb{R}_{>0}$, $u \in \mathbb{R}_{<0}$, and $x + t + u \in \mathbb{R}_{>0}$, then we have $x \in S \Leftrightarrow x + t \in S \Leftrightarrow (x + t) + u \in S$. Similarly, if $x \in \mathbb{R}_{>0}$, $t \in \mathbb{R}_{<0}$, $u \in \mathbb{R}_{>0}$, and $x + t + u \in \mathbb{R}_{>0}$, then we have $x \in S \Leftrightarrow x + u \in S \Leftrightarrow (x + u) - t \in S$. ■

Lemma 5.29 *Let $S \subseteq \mathbb{R}_{>0}$ be a 1-sum-stable set in $\mathbb{R}_{>0}$, and $a \in \mathbb{Z}$. We have $a \in T_{\mathbb{R}_{>0}}(S)$.*

Proof Since S is 1-sum-stable is $\mathbb{R}_{>0}$, we have $1 \in T_{\mathbb{R}_{>0}}(S)$. The conclusion follows by Lemma 5.28. ■

Lemma 5.30 *Let $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq \mathbb{R}_{>0}$ be a r-product-stable set in $\mathbb{R}_{>0}$. The set $T_{\mathbb{R}_{>0}}(S)$ is r-product-stable in $\mathbb{R}$.*

Proof Let $x \in \mathbb{R}$. We have to prove that $x \in T_{\mathbb{R}_{>0}}(S)$ if and only if $rx \in T_{\mathbb{R}_{>0}}(S)$.

On the one hand, suppose that $x \in T_{\mathbb{R}_{>0}}(S)$, and let $x' \in \mathbb{R}_{>0}$ such that $x' + rx \in \mathbb{R}_{>0}$. Since $\frac{x' + rx}{r} \in \mathbb{R}_{>0}$, since S is r-product-stable in $\mathbb{R}_{>0}$ and since $x \in T_{\mathbb{R}_{>0}}(S)$, we have

$$x' \in S \Leftrightarrow \frac{x'}{r} \in S \Leftrightarrow \frac{x'}{r} + x \in S \Leftrightarrow x' + rx \in S.$$

Hence, $rx \in T_{\mathbb{R}_{>0}}(S)$.

Similarly, on the other hand, suppose that $rx \in T_{\mathbb{R}_{>0}}(S)$, and let $x' \in \mathbb{R}_{>0}$ such that $x' + x \in \mathbb{R}_{>0}$. Since $r(x' + x) \in \mathbb{R}_{>0}$, since S is r-product-stable in S, and since $rx \in T_{\mathbb{R}_{>0}}(S)$, we have

$$x' \in S \Leftrightarrow rx' \in S \Leftrightarrow rx' + rx \in S \Leftrightarrow x' + x \in S.$$

Hence, $x \in T_{\mathbb{R}_{>0}}(S)$. ■

Lemma 5.31 *Let $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq \mathbb{R}_{>0}$ be a set that is r-product-stable in $\mathbb{R}_{>0}$ and 1-sum-stable in $\mathbb{R}_{>0}$. For every $k \in \mathbb{Z}$, we have $r^k \in T_{\mathbb{R}_{>0}}(S)$.*

Proof Since S is 1-sum-stable in $\mathbb{R}_{>0}$, we have $1 \in T_{\mathbb{R}_{>0}}(S)$. The conclusion follows by Lemma 5.30. ■

Intuitively, being able to add or subtract r^k from a number $x \in \mathbb{R}$, for any k, makes it possible to change in an arbitrary way finitely many digits in its r-encodings, without influencing the fact that x belongs or not to S. Our next step will be to show that this property can be extended to all digits of r-encodings.

Lemma 5.32 *Let $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors, and $S \subseteq \mathbb{R}_{>0}$ be a set that is both r- and s-recognizable, both r- and s-product-stable in $\mathbb{R}_{>0}$, and 1-sum-stable in $\mathbb{R}_{>0}$. There exist $l, m \in \mathbb{N}_{>0}$ such that, for every $k \in \mathbb{N}_{>0}$, we have*

$$\frac{m}{r^{lk} - 1} \in T_{\mathbb{R}_{>0}(S)}.$$

Proof Suppose, without loss of generality, that there exists a prime factor of s that does not divide r (if it is not the case, then it suffices to swap the bases r and s). Formally, let $\{q_1, q_2, \ldots, q_n\}$ be the set of prime factors of r. We suppose that there exist $j_1, j_2, \ldots, j_n \in \mathbb{N}$ such that

$$s = q_1^{j_1} q_2^{j_2} \ldots q_n^{j_n} Q,$$

with $Q \in \mathbb{N}_{\geq 2}$ and $\gcd(Q, r) = 1$.

Lemma 5.31 implies that $s^k \in T_{\mathbb{R}_{>0}}(S)$ for all $k \in \mathbb{Z}$. In particular, we have $\dfrac{1}{s^k} \in T_{\mathbb{R}_{>0}}(S)$ for all $k \in \mathbb{N}$.

By Theorems 2.56 and 2.57, the length $|v_k|$ of the smallest word $v_k \in \{0, 1, \ldots, r-1\}^+$ such that there exist $u_k \in \{0, 1, \ldots, r-1\}^*$ and $\dfrac{1}{s^k} = \langle 0 \star u_k v_k^\omega \rangle_r$, is the smallest integer $|v_k| \in \mathbb{N}_{>0}$ such that $r^{|v_k|} \equiv_{Q^k} 1$. In other words, $|v_k|$ is the smallest integer such that $r^{|v_k|} - 1$ is a multiple of Q^k. Since $Q \geq 2$, it follows that the lengths $|v_k|$ are unbounded with respect to k. Hence, the set $T_{\mathbb{R}_{>0}}(S)$ contains rational numbers with infinitely many distinct ultimate periods.

By Lemma 5.27, $T_{\mathbb{R}_{>0}}(S)$ is r-recognizable. Hence, there exists a Muller RVA $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ recognizing $T_{\mathbb{R}_{>0}}(S)$ in base r. Suppose that $F = \{F_1, F_2, \ldots, F_{|F|}\}$. Since an infinite number of encodings of rationals ending with distinct periods are accepted by $\mathcal{A}$, there exists a subset $F_i \in F$, with $i \in \{1, 2, \ldots, |F|\}$, such that an infinite number of paths reading such encodings in $\mathcal{A}$ end up in exactly F_i. In particular, there exist $u_1, u_2 \in \{0, r-1\}\{0, 1, \ldots, r-1\}^*$, $v_1, v_2 \in \{0, 1, \ldots, r-1\}^*$, and $w_1, w_2 \in \{0, 1, \ldots, r-1\}^+$ such that w_1^ω is not a suffix of w_2^ω, both words $u_1 \star v_1 w_1^\omega$ and $u_2 \star v_2 w_2^\omega$ are accepted by $\mathcal{A}$, and the paths π_1 and π_2 of $\mathcal{A}$ reading them end up in exactly F_i.

There then exists a word $w_1' \in \{0, 1, \ldots, r-1\}^*$ such that the set of states visited by π_1 after the reading of $u_1 \star v_1 w_1'$ is exactly F_i. There also exist symbols $\sigma_0, \sigma_1, \sigma_2, \ldots \in \{0, 1, \ldots, r-1\}$ such that

$$u_1 \star v_1 w_1^\omega = u_1 \star v_1 w_1' (\sigma_0 \sigma_1 \ldots \sigma_{|w_1|-1})^\omega,$$

where $\sigma_0 \sigma_1 \ldots \sigma_{|w_1|-1}$ is a cyclic permutation of w_1, i.e., there exists $s \in \{0, 1, \ldots, |w_1| - 1\}$ such that $w_1 = \sigma_s \sigma_{s+1} \ldots \sigma_{|w_1|-1} \sigma_0 \sigma_1 \ldots \sigma_{s-1}$.

Let q be an arbitrary state of F_i. At each time the path π visits q after the reading of $u_1 \star v_1 w_1'$, the symbol read directly after this visit corresponds to a position of the word $\sigma_0 \sigma_1 \ldots \sigma_{|w_1|-1}$. Since there are finitely many such positions, and since q is visited infinitely many times by π, there exist two distinct occurrences of q in π, located after the reading of $u_1 \star v_1 w_1'$, and such that the symbols read directly after the visits of these respective occurrences of π correspond to identical positions in $\sigma_0\sigma_1\ldots\sigma_{|w_1|-1}$. Hence, by the deterministic nature of $\mathcal{A}$, the path π ends up cycling between these occurrences of q, and $u_1 \star v_1 w_1^\omega$ ends with $(w_1'')^\omega$, where w_1'' is the word that is read between these occurrences. Moreover, the set of states visited by π between them is exactly the set F_i.

A similar reasoning holds for the word $u_2 \star v_2 w_2^\omega$, and we get that the path π_2 ends with a cycle rooted in q, labeled with a word $w_2'' \in \{0,1,\ldots,r-1\}^+$.

The periods w_1'' and w_2'' can be repeated arbitrarily, hence, by reading $|w_2''|$ times the word w_1'', and $|w_1''|$ times the word w_2'', we can assume without loss of generality that $|w_1''| = |w_2''|$. Moreover, we can assume without loss of generality that $\langle 0w_1''\rangle_r > \langle 0w_2''\rangle_r$, otherwise w_1^ω would be a suffix of w_2^ω. Besides, there exist $v \in \{0, r-1\}\{0,1,\ldots,r-1\}^*$ and $w \in \{0,1,\ldots,r-1\}^*$ such that the path reading $v \star w$ ends in q. From the structure of $\mathcal{A}$, it follows that for every $k \in \mathbb{N}$, the word $v \star w((w_1'')^k w_2'')^\omega$ is accepted by $\mathcal{A}$.

For each $k \in \mathbb{N}$, we thus have

$$\langle v \star w((w_1'')^k w_2'')^\omega\rangle_r \in T_{\mathbb{R}_{>0}}(S).$$

Developing, we get

$$\frac{d_k + \langle vw\rangle_r}{r^{|w|}} \in T_{\mathbb{R}_{>0}}(S),$$

with

$$d_k = \langle 0 \star ((w_1'')^k w_2'')^\omega\rangle_r. \tag{5.3}$$

By Lemma 5.28, we have

$$d_k + \langle vw\rangle_r \in T_{\mathbb{R}_{>0}}(S).$$

Using one more time this lemma, we get

$$d_k \in T_{\mathbb{R}_{>0}}(S).$$

We have

$$r^{l(k+1)} d_k = \langle 0(w_1'')^k w_2'' \star ((w_1'')^k w_2'')^\omega\rangle_r, \tag{5.4}$$

where $l = |w_1''| = |w_2''|$.

Combining (5.3) and (5.4), we get

$$d_k = \frac{\langle 0(w_1'')^k w_2''\rangle_r}{r^{l(k+1)} - 1}.$$

We now express d_k in terms of $\langle 0w_1''\rangle_r$, $\langle 0w_2''\rangle_r$, and k :

$$\begin{aligned} d_k &= \frac{\langle 0(w_1'')^k w_2''\rangle_r}{r^{l(k+1)}-1} \\ &= \frac{\langle 0w_2''\rangle_r + \langle 0w_1''\rangle_r \sum_{i=1}^{k} r^{li}}{r^{l(k+1)}-1} \\ &= \frac{\langle 0w_2''\rangle_r - \langle 0w_1''\rangle_r}{r^{l(k+1)}-1} + \frac{\langle 0w_1''\rangle_r}{r^l-1}. \end{aligned}$$

The next step will consist in getting rid of the second term of this expression. By Lemma 5.29, we have $\langle 0w_1''\rangle_r \in T_{\mathbb{R}_{>0}}(S)$. By Lemma 5.28, we then have for all $k \in \mathbb{N}$

$$(r^l-1)d_k - \langle 0w_1''\rangle_r = \frac{m}{r^{l(k+1)}-1} \in T_{\mathbb{R}_{>0}}(S),$$

where $m = (r^l-1)(\langle 0w_2''\rangle_r - \langle 0w_1''\rangle_r)$ is such that $m \in \mathbb{N}_{>0}$. For all $k \in \mathbb{N}_{>0}$, we thus have

$$\frac{m}{r^{lk}-1} \in T_{\mathbb{R}_{>0}}(S).$$

■

We can now prove the final result of this section.

Theorem 5.33 *Let $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors, and $S \subseteq \mathbb{R}_{>0}$ be a set that is both r- and s-recognizable, both r- and s-product-stable in $\mathbb{R}_{>0}$, 1-sum-stable in $\mathbb{R}_{>0}$, and that admits infinitely many boundary points. There exists a set $S' \subseteq \mathbb{R}_{>0}$ that is both r- and s-recognizable, that admits infinitely many boundary points, and such that $T_{\mathbb{R}_{>0}}(S') = \mathbb{R}$.*

Proof By Lemma 5.32, there exist $l, m \in \mathbb{N}_{>0}$ such that, for every $k \in \mathbb{N}_{>0}$, we have $\dfrac{m}{r^{lk}-1} \in T_{\mathbb{R}_{>0}}(S)$.

Let $S' = \dfrac{1}{m}S$.

First, remark that S' admits infinitely many boundary points, and that, since S is both r- and s-recognizable, Theorem 2.65 implies that S' is both r- and s-recognizable as well.

We have $T_{\mathbb{R}_{>0}}(S') = \dfrac{1}{m}T_{\mathbb{R}_{>0}}(S)$. Indeed, on the one hand, suppose that $x \in T_{\mathbb{R}_{>0}}(S')$ and let us show that $mx \in T_{\mathbb{R}_{>0}}(S)$. Let $x' \in \mathbb{R}_{>0}$ such that $x' + mx \in \mathbb{R}_{>0}$. Since $\dfrac{x'+mx}{m} \in \mathbb{R}_{>0}$, and since $x \in T_{\mathbb{R}_{>0}}(S')$, we have

$$x' \in S \Leftrightarrow \frac{x'}{m} \in S' \Leftrightarrow \frac{x'}{m} + x \in S' \Leftrightarrow x' + mx \in S.$$

Hence, $mx \in T_{\mathbb{R}_{>0}}(S)$. Similarly, on the other hand, suppose that $x \in T_{\mathbb{R}_{>0}}(S)$, and let us show that $\frac{x}{m} \in T_{\mathbb{R}_{>0}}(S')$. Let $x' \in \mathbb{R}_{>0}$ such that $x' + \frac{x}{m} \in \mathbb{R}_{>0}$. Since $m(x' + \frac{x}{m}) \in \mathbb{R}_{>0}$, and since $x \in T_{\mathbb{R}_{>0}}(S)$, we have

$$x' \in S' \Leftrightarrow mx' \in S \Leftrightarrow mx' + x \in S \Leftrightarrow x' + \frac{x}{m} \in S'.$$

Hence, $\frac{x}{m} \in T_{\mathbb{R}_{>0}}(S')$. We then have

$$\frac{1}{r^{lk} - 1} \in T_{\mathbb{R}_{>0}}(S') \text{ for all } k \in \mathbb{N}_{>0}. \tag{5.5}$$

Lemma 5.31 implies that we have $\frac{1}{r^k} \in T_{\mathbb{R}_{>0}}(S)$ for all $k \in \mathbb{N}_{>0}$. By Lemma 5.28, we thus have $\frac{m}{r^k} \in T_{\mathbb{R}_{>0}}(S)$. Hence, we have

$$\frac{1}{r^k} \in T_{\mathbb{R}_{>0}}(S') \text{ for all } k \in \mathbb{N}_{>0}. \tag{5.6}$$

By Lemma 5.30, $T_{\mathbb{R}_{>0}}(S)$ is r-product-stable in $\mathbb{R}$. Hence, $T_{\mathbb{R}_{>0}}(S')$ is r-product-stable in $\mathbb{R}$ as well. Indeed, let $x \in \mathbb{R}$. We have

$$x \in T_{\mathbb{R}_{>0}}(S') \Leftrightarrow mx \in T_{\mathbb{R}_{>0}}(S) \Leftrightarrow rmx \in T_{\mathbb{R}_{>0}}(S) \Leftrightarrow rx \in T_{\mathbb{R}_{>0}}(S').$$

By Lemma 5.27, $T_{\mathbb{R}_{>0}}(S)$ is r-recognizable. Hence, $T_{\mathbb{R}_{>0}}(S')$ is r-recognizable as well by Theorem 2.65. By Theorem 2.69, it is then sufficient to establish that $T_{\mathbb{R}_{>0}}(S') \cap \mathbb{Q} = \mathbb{Q}$.

Let $x \in \mathbb{Q}$. It follows from Corollary 2.59 that this number has a r-encoding of the form $v \star wu^\omega$, where $v \in \{0, r-1\}\{0, 1, \ldots, r-1\}^*$, $w \in \{0, 1, \ldots, r-1\}^*$, and $u \in \{0, 1, \ldots, r-1\}^+$. By defining $u' = u^l$, we have $|u'| = l|u|$, and

$$x = \langle v \star w(u')^\omega \rangle_r.$$

Hence, we have

$$r^{|w|}x = \langle vw \star (u')^\omega \rangle_r, \tag{5.7}$$

and

$$r^{|w|+|u'|}x = \langle vwu' \star (u')^\omega \rangle_r. \tag{5.8}$$

Combining (5.7) and (5.8), we get

$$\begin{aligned} x &= \frac{\langle vwu'\rangle_r - \langle vw\rangle_r}{r^{|w|}(r^{|u'|}-1)} \\ &= \frac{r^{|u'|}\langle vw\rangle_r + \langle 0u'\rangle_r - \langle vw\rangle_r}{r^{|w|}(r^{|u'|}-1)} \\ &= \frac{\langle vw\rangle_r(r^{|u'|}-1) + \langle 0u'\rangle_r}{r^{|w|}(r^{|u'|}-1)} \\ &= \frac{\langle vw\rangle_r}{r^{|w|}} + \frac{\langle 0u'\rangle_r}{r^{|w|}(r^{|u'|}-1)}. \end{aligned}$$

Using Equation 5.6, and by Lemma 5.28, we have

$$\frac{\langle vw\rangle_r}{r^{|w|}} \in T_{\mathbb{R}_{>0}}(S').$$

By Equation 5.5, since $T_{\mathbb{R}_{>0}}(S')$ is r-product-stable in $\mathbb{R}$, and since $|u'| = l|u|$, we have

$$\frac{1}{r^{|w|}(r^{|u'|}-1)} \in T_{\mathbb{R}_{>0}}(S').$$

We also have

$$\frac{\langle 0u'\rangle_r}{r^{|w|}(r^{|u'|}-1)} \in T_{\mathbb{R}_{>0}}(S')$$

by Lemma 5.28.

By Lemma 5.28, we thus have $x \in T_{\mathbb{R}_{>0}}(S')$. Since x was an arbitrary rational number, we have $T_{\mathbb{R}_{>0}}(S') \cap \mathbb{Q} = \mathbb{Q}$. ∎

5.7 Generalization of Cobham's theorem to Muller RVA

The results obtained in Sections 5.2, 5.3 and 5.6 can now be combined to obtain a generalization of Cobham's theorem to Muller RVA. The following theorem holds for subsets of $[0, 1]$ simultaneously recognizable in two bases that have different sets of prime factors.

Theorem 5.34 *Let $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors, and let $S \subseteq [0, 1]$. If S is simultaneously r- and s-recognizable, then S is definable in $\langle \mathbb{R}, +, <, 1\rangle$.*

Proof We know by Theorem 5.7 that if the set $B(S)$, containing the boundary points of S, is finite, then S is definable in $\langle \mathbb{R}, +, <, 1\rangle$. Let us prove that it is the case.

By contradiction, suppose that $B(S)$ is infinite, i.e., that S admits infinitely many boundary points.

Since bases with different sets of prime factors are necessarily multiplicatively independent, Theorem 5.15 holds, and implies the existence of powers $r' = r^i$, $s' = s^j$, with $i, j \in \mathbb{N}_{>0}$, and of a set $S' \subseteq [0, 1]$ that is both r'- and s'-recognizable, both r'- and s'-product-stable in $[0, 1]$, and that admits infinitely many boundary points.

Remark that the base r' (resp. s') has the same set of prime factors as r (resp. s). Hence, r' and s' have different sets of prime factors.

By Lemma 5.25, there exists a set $S'' \subseteq \mathbb{R}_{>0}$ that is both r'- and s'-recognizable, both r'- and s'-product-stable in $\mathbb{R}_{>0}$, 1-sum-stable in $\mathbb{R}_{>0}$, and that admits infinitely many boundary points.

Using Theorem 5.33 applied to S'', we obtain a set $S''' \subseteq \mathbb{R}_{>0}$ that is both r- and s-recognizable, that admits infinitely many boundary points, and such that the set of numbers t for which S''' is t-sum stable in $\mathbb{R}_{>0}$, i.e., $T_{\mathbb{R}_{>0}}(S''')$, is equal to $\mathbb{R}$.

Since S''' admits infinitely many boundary points, it contains at least one element $x \in \mathbb{R}_{>0}$. Let $x' \in \mathbb{R}_{>0}$ be an arbitrary point. We have $x' - x \in T_{\mathbb{R}_{>0}}(S''')$. Hence, by the definition of $T_{\mathbb{R}_{>0}}(S''')$, we have $x + (x' - x) = x' \in S'''$, which implies $S''' = \mathbb{R}_{>0}$. This is a contradiction with the fact that S''' admits infinitely many boundary points. Hence, the assumption stating that $B(S)$ is infinite is false. As a consequence, S is definable in $\langle \mathbb{R}, +, <, 1\rangle$. ■

Claim 5.3 then follows, since the domain including the set S of Theorem 5.34 can be extended to $\mathbb{R}$ by using Theorem 4.15.

Theorem 5.35 *Let $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors, and let $S \subseteq \mathbb{R}$. If S is simultaneously r- and s-recognizable, then S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, <\rangle$.*

Proof This is a direct corollary of Theorems 4.15 and 5.34. ■

5.8 Summary

In this chapter, we studied the one-dimensional sets simultaneously recognizable by Muller RVA, or by weak deterministic RVA, in multiple bases, and we obtained a complete characterization of such sets. In this section, we recall this characterization.

Let $r, s \in \mathbb{N}_{>1}$ be two bases, and let $S \subseteq \mathbb{R}$ be a one-dimensional set of reals.

- If r and s are multiplicatively dependent, then, by Corollary 3.33, S is weakly r-recognizable if and only if it is weakly s-recognizable.
- If r and s are multiplicatively independent, and if S is simultaneously weakly r- and weakly s-recognizable, then, by Theorem 5.20, S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

This characterization is similar to the one known for the integer domain [Cob69, BHMV94].

For Muller automata, or, equivalently, (possibly non deterministic) Büchi automata, we have demonstrated that multiplicatively independence of the bases is not a strong enough condition. Precisely, we have the following characterization :

- If r and s are multiplicatively dependent, then, by Corollary 2.68, S is r-recognizable if and only if it is s-recognizable.
- If r and s are multiplicatively independent, but share the same set of prime factors, then, by Theorem 5.21, there exists a set $S \subseteq \mathbb{R}$ that is both r- and s-recognizable, and that is not definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.
- If r and s are multiplicatively independent and, in addition, have different sets of prime factors, then, by Theorem 5.35, S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

As a consequence, we have thus established that the sets of real numbers that are recognized by infinite-word automata in all encoding bases are exactly those that are recognizable by weak deterministic automata in all encoding bases. This result brings us a theoretical justification to the use of weak deterministic automata for representing sets of reals and integer numbers : If recognizability by automata has to be achieved regardless of the representation base, then the representable sets are exactly those that can be recognized by weak deterministic automata.

Chapter 6

Multi-dimensional sets

6.1 Overview

In Chapter 5, we established that a necessary and sufficient condition for a one-dimensional set $S \subseteq \mathbb{R}$ to be definable in the theory $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$ is that S is simultaneously recognized by weak deterministic RVA in two multiplicatively independent bases. When dealing with general RVA, we also proved that this condition does not hold anymore; in this case, the bases must have different sets of prime factors.

In this chapter, we generalize these results to multi-dimensional sets, i.e., to subsets of $\mathbb{R}^n$, with $n \in \mathbb{N}_{>0}$, and we naturally consider the same conditions on the bases in which a set S is recognizable, or weakly recognizable. Precisely, the aim of this chapter is to prove the following claim.

Claim 6.1 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors (resp. multiplicatively independent bases), and $S \subseteq \mathbb{R}^n$. If S is (resp. weakly) r- and (resp. weakly) s-recognizable, then S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.*

By Theorems 4.15 and 4.16, it suffices to establish that the subsets of $[0,1]^n$ that have those properties are definable in $\langle \mathbb{R}, +, <, 1 \rangle$, that is, to prove the next claim.

Claim 6.2 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors (resp. multiplicatively independent bases), and $S \subseteq [0,1]^n$. If S is (resp. weakly) r- and (resp. weakly) s-recognizable, then S is definable in $\langle \mathbb{R}, +, <, 1 \rangle$.*

An identical proof of Claim 6.2 will handle both cases of recognizability, and of weakly recognizability. Its sketch is illustrated in Figure 6.11, page 134.

6.2 Induction

In order to establish Claim 6.2, we proceed by induction. The base case, $n = 1$, has been solved in the previous chapter. It thus remains to prove the inductive case, i.e.,

Claim 6.3 *Let $n \in \mathbb{N}_{>1}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors (resp. multiplicatively independent bases), and let $S \subseteq [0,1]^n$ be a both (resp. weakly) r- and (resp. weakly) s-recognizable set. If, for all $m \in \{1, 2, \ldots, n-1\}$, and for each set $S' \subseteq [0,1]^m$ that is both (resp. weakly) r- and (resp. weakly) s-recognizable, S' is definable in $\langle \mathbb{R}, +, <, 1\rangle$, then S is definable in $\langle \mathbb{R}, +, <, 1\rangle$.*

In the next section, we reduce these problems to simpler ones.

6.3 Product-stable sets

In order to be able to prove that the recognizability of a subset of $[0,1]^n$ in multiple bases leads to its definability in $\langle \mathbb{R}, +, <, 1\rangle$, we need to establish a link between the arithmetical properties of this set, and the structure of automata recognizing it. Those arithmetical properties will consist in generalizations of product-stability ones, introduced in Section 5.3 when dealing with one-dimensional sets. We will reduce the problem of establishing the definability of a set $S \subseteq [0,1]^n$ in $\langle \mathbb{R}, +, <, 1\rangle$ to the same problem for sets that have the characteristic of being product-stable.

Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq [0,1]^n$ be a set recognized in base r either by a Muller, or by a weak deterministic RVA. Let $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ be such an automaton, and $q \in Q$. Recall that $L(q)$ denotes the language accepted from q in $\mathcal{A}$. Similarly, if q is a state reachable from $\delta^*(q_0, w\star)$, with $w \in \{0, r-1\}^n(\{0, 1, \ldots, r-1\}^n)^*$, then $S(q)$ denotes the set of vectors encoded by $\vec{0} \star L(q)$, as formalized in the following definition.

Definition 6.4 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq [0,1]^n$ be a set recognized by an automaton $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ that is either weak deterministic, or a Muller automaton. Let $q \in Q$ be such that there exist $w_1 \in \{0, r-1\}^n(\{0,1,\ldots,r-1\}^n)^*$, $w_2 \in (\{0,1,\ldots,r-1\}^n)^*$ such that $\delta^*(q_0, w_1 \star w_2)$ is defined and equal to q. The set $S(q) \subseteq [0,1]^n$ is defined by*

$$S(q) = \{\langle \vec{0} \star w\rangle_{r,n} \mid w \in L(q)\}.$$

□

The next lemma states that the sets $S(q)$ are (resp. weakly) recognizable in all bases in which S is (resp. weakly) recognizable.

Lemma 6.5 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases, and $S \subseteq [0,1]^n$ be a set recognized by a Muller (resp. weak deterministic) RVA $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ in base r. Let $q \in Q$ be a state such that there exist $w_1 \in \{0, r-1\}^n(\{0,1,\ldots,r-1\}^n)^*$ and $w_2 \in (\{0,1,\ldots,r-1\}^n)^*$ such that $\delta^*(q_0, w_1 \star w_2)$ is defined and equal to q. If S is (resp. weakly) s-recognizable, then $S(q)$ is (resp. weakly) s-recognizable as well.*

Proof Recall that $\mathcal{A}$ has a deterministic transition relation. Hence, the language $L(q)$ can be expressed as

$$L(q) = \{w \in (\{0,1,\ldots,r-1\}^n)^\omega \mid w_1 \star w_2 w \in L(\mathcal{A})\}.$$

We then get

$$S(q) = \left\{\vec{x} \in [0,1]^n \mid \frac{\langle w_1 w_2\rangle_{r,n} + \vec{x}}{r^{|w_2|}} \in S\right\}.$$

From this relation and Theorem 2.65 (resp. Theorem 3.35), we obtain that $S(q)$ is (resp. weakly) recognizable in all bases for which S is (resp. weakly) recognizable. ■

Let us introduce a notion of product-stability, now adapted for subsets of $[0,1]^n$.

Definition 6.6 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $\vec{v} \in [0,1]^n \cap \mathbb{Q}^n$ be a* pivot*, and $f \in \mathbb{R}_{\geq 1}$ be a factor. A set $S \subseteq [0,1]^n$ is* f-product-stable *with respect to the pivot $\vec{v}$ if and only if*

$$(\forall \vec{x} \in [0,1]^n - \vec{v}) \left(\vec{x} \in S - \vec{v} \Leftrightarrow \frac{1}{f}\vec{x} \in S - \vec{v}\right).$$

□

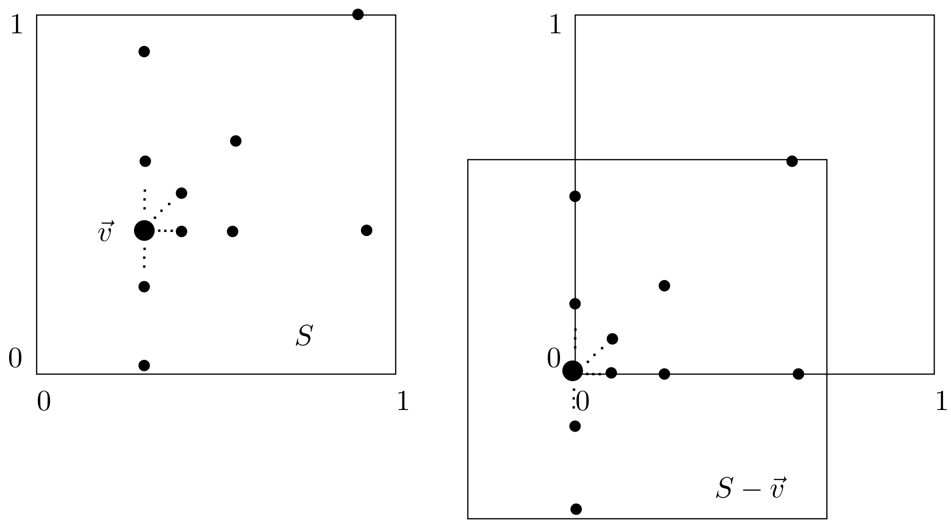

Figure 6.1: A $\frac{5}{2}$-product-stable set with respect to $\vec{v}$.

Intuitively, that a set is f-product-stable with respect to the pivot $\vec{v}$ means that the set does not change when it is magnified by the zoom factor f around the pivot $\vec{v}$. Figure 6.1 illustrates the notion of product-stability.

The next lemma states that this property is preserved by transformations of the form $\vec{x} \mapsto C\vec{x} + \vec{a}$, with $C \in \mathbb{Q}^{n \times n}$ and $\vec{a} \in \mathbb{Q}^n$, provided that $(C[0,1]^n + \vec{a}) \subseteq [0,1]^n$, and the new pivot $\vec{v'} = C\vec{v} + \vec{a}$ belongs to $[0,1]^n$.

Lemma 6.7 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $\vec{v} \in [0,1]^n \cap \mathbb{Q}^n$ be a pivot, $f \in \mathbb{R}_{\geq 1}$ be a factor, $S \subseteq [0,1]^n$ be a f-product-stable set with respect to $\vec{v}$, $C \in \mathbb{Q}^{n \times n}$, and $\vec{a} \in \mathbb{Q}^n$. If $C[0,1]^n + \vec{a} \subseteq [0,1]^n$ and $C\vec{v} + \vec{a} \in [0,1]^n$, then $CS + \vec{a}$ is f-product-stable with respect to $C\vec{v} + \vec{a}$.*

Proof Suppose $C[0,1]^n + \vec{a} \subseteq [0,1]^n$ and $C\vec{v} + \vec{a} \in [0,1]^n$. Let us prove that $CS + \vec{a}$ is f-product-stable with respect to $C\vec{v} + \vec{a}$, i.e., that

$$(\forall \vec{x} \in [0,1]^n - (C\vec{v} + \vec{a})) \left(\vec{x} \in C(S - \vec{v}) \Leftrightarrow \frac{1}{f}\vec{x} \in C(S - \vec{v}) \right),$$

since we have $CS + \vec{a} - (C\vec{v} + \vec{a}) = C(S - \vec{v})$.

Let $\vec{x} \in [0,1]^n - (C\vec{v} + \vec{a})$. We successively have

$$\begin{aligned}
& \vec{x} \in C(S - \vec{v}) \\
\Leftrightarrow \quad & (\exists \vec{x'} \in S - \vec{v})(\vec{x} = C\vec{x'}) \\
\Leftrightarrow \quad & (\exists \vec{x'} \in S - \vec{v}) \left(\frac{1}{f}\vec{x'} \in S - \vec{v} \wedge \vec{x} = C\vec{x'} \right) \\
\Leftrightarrow \quad & (\exists \vec{x'} \in S - \vec{v}) \left(\frac{1}{f}\vec{x'} \in S - \vec{v} \wedge \frac{1}{f}\vec{x} = C\left(\frac{1}{f}\vec{x'}\right) \right) \\
\Leftrightarrow \quad & \frac{1}{f}\vec{x} \in C(S - \vec{v}).
\end{aligned}$$

■

In the following lemma, we give the expected link between the structure of automata recognizing sets of real vectors, and the notion of product-stability : We establish that if a state q, reachable after a transition labeled by $\star$, is the origin of a cycle, then the set $S(q)$ is product-stable.

Lemma 6.8 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq [0,1]^n$ be a set recognized by a Muller (resp. weak deterministic) RVA $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ in base r. Let $q \in Q$ be a state such that there exist $w_1 \in \{0, r-1\}^n(\{0,1,\dots,r-1\}^n)^*$ and $w_2 \in (\{0,1,\dots,r-1\}^n)^*$ such that $\delta^*(q_0, w_1 \star w_2)$ is defined and equal to q. If, in addition, there exists $w \in (\{0,1,\dots,r-1\}^n)^+$ such that $\delta^*(q,w) = q$, then $S(q)$ is $r^{|w|}$-product-stable with respect to the pivot $\langle \vec{0} \star w^\omega \rangle_{r,n}$.*

Proof Since $\delta^*(q,w) = q$, we have $w' \in L(q) \Leftrightarrow ww' \in L(q)$. It follows that

$$\vec{x} \in S(q) \Leftrightarrow \vec{x} \in [0,1]^n \wedge \frac{\vec{x} + \langle \vec{0}w \rangle_{r,n}}{r^{|w|}} \in S(q).$$

Remark that the transformation $\vec{x} \mapsto \dfrac{\langle \vec{0}w \rangle_{r,n} + \vec{x}}{r^{|w|}}$ admits the fixed point $\langle \vec{0} \star w^\omega \rangle_{r,n} \in [0,1]^n$. Translating $S(q)$ so as to move this fixed point onto $\vec{0}$, we get

$$\begin{aligned}
& \vec{x} \in S(q) - \langle \vec{0} \star w^\omega \rangle_{r,n} \\
\Leftrightarrow \quad & \vec{x} \in [0,1]^n - \langle \vec{0} \star w^\omega \rangle_{r,n} \wedge \\
& \quad \frac{\vec{x} + \langle \vec{0} \star w^\omega \rangle_{r,n} + \langle \vec{0}w \rangle_{r,n}}{r^{|w|}} \in S(q).
\end{aligned} \tag{6.1}$$

Since $\dfrac{\langle \vec{0} \star w^\omega \rangle_{r,n} + \langle \vec{0}w \rangle_{r,n}}{r^{|w|}} = \langle \vec{0} \star w^\omega \rangle_{r,n}$, Equation 6.1 becomes

$$\begin{aligned}
& \vec{x} \in S(q) - \langle \vec{0} \star w^\omega \rangle_{r,n} \\
\Leftrightarrow \quad & \vec{x} \in [0,1]^n - \langle \vec{0} \star w^\omega \rangle_{r,n} \wedge \frac{\vec{x}}{r^{|w|}} + \langle \vec{0} \star w^\omega \rangle_{r,n} \in S(q).
\end{aligned}$$

Hence, we have

$$\begin{aligned} & \vec{x} \in S(q) - \langle \vec{0} \star w^{\omega} \rangle_{r,n} \\ \Leftrightarrow \quad & \vec{x} \in [0,1]^n - \langle \vec{0} \star w^{\omega} \rangle_{r,n} \wedge \frac{\vec{x}}{r^{|w|}} \in S(q) - \langle \vec{0} \star w^{\omega} \rangle_{r,n}, \end{aligned}$$

which matches with Definition 6.6. It follows that $S(q)$ is $r^{|w|}$-product-stable with respect to the pivot $\langle \vec{0} \star w^{\omega} \rangle_{r,n}$. ■

In summary, if $\mathcal{A}$ recognizes the set $S \subseteq [0,1]^n$ in base r, then each reachable state q of the fractional part of $\mathcal{A}$ recognizes a set $S(q) \subseteq [0,1]^n$ that is (resp. weakly) recognizable in all bases for which S is (resp. weakly) recognizable. Furthermore, if the transition relation admits a cycle labeled by $w \in (\{0,1,\ldots,r-1\}^n)^+$ and rooted at q, then the set $S(q)$ is $r^{|w|}$-product-stable with respect to the pivot $\langle \vec{0} \star w^{\omega} \rangle_{r,n}$.

In the next step, we show that any recognizable set can be decomposed into a combination of product-stable sets that can be considered individually.

Lemma 6.9 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq [0,1]^n$ be a set (resp. weakly) r-recognizable. There exist $k \in \mathbb{N}$, $\vec{\imath_1}, \vec{\imath_2}, \ldots, \vec{\imath_k} \in \mathbb{N}^n$, $\ell_1, \ell_2, \ldots, \ell_k \in \mathbb{N}_{>0}$, $\vec{v_1}, \vec{v_2}, \ldots \vec{v_k} \in [0,1]^n \cap \mathbb{Q}^n$, and $S_1, S_2, \ldots, S_k \subseteq [0,1]^n$ such that, for each $j \in \{1,2,\ldots,k\}$, S_j is r^{ℓ_j}-product-stable with respect to the pivot $\vec{v_j}$ and is (resp. weakly) recognizable in all bases in which S is (resp. weakly) recognizable, and*

$$S = \bigcup_{j \in \{1,2,\ldots,k\}} \frac{S_j + \vec{\imath_j}}{r^{\ell_j}}.$$

Proof Let $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ be a Muller (resp. weak deterministic) RVA recognizing S in base r.

Let Q_1 be the set of states q of $\mathcal{A}$ reachable from $\delta^*(q_0, \vec{0}\star)$ and that are origin of at least one cycle, i.e.,

$$\begin{aligned} Q_1 = \{q \in Q \mid (\exists w_1 \in (\{0,1,\ldots,r-1\}^n)^*)(\exists w_2 \in (\{0,1,\ldots,r-1\}^n)^+) \\ (\delta^*(q_0, \vec{0} \star w_1) = q = \delta^*(q, w_2))\}. \end{aligned}$$

Since S is a subset of $[0,1]^n$, and RVA accept all the encodings of the vectors they represent, each vector $\vec{x} \in S$ admits at least one r-encoding having $\vec{0}\star$ as prefix; hence, there exists an accepting path $\pi_{\vec{x}}$ of $\mathcal{A}$, reading an encoding $\vec{0} \star w$ such that $\langle \vec{0} \star w \rangle_{r,n} = \vec{x}$, and eventually visiting a state in Q_1.

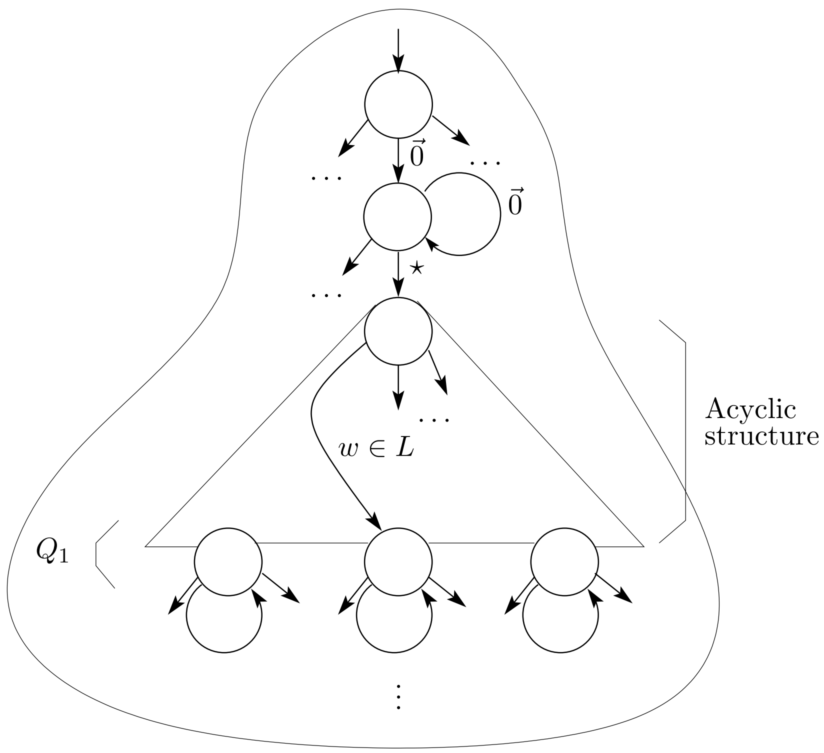

Figure 6.2: Illustration of Lemma 6.9.

Let L be the language of words $w \in (\{0, 1, \ldots, r-1\}^n)^*$ labeling finite paths π with origin $\delta^*(q_0, \vec{0}\star)$ and destination $q \in Q_1$ such that $q' \notin Q_1$ for every state q' distinct from q visited by π, and there is only one occurrence of q in π. The situation is illustrated in Figure 6.2.

For each word $w \in L$, let $q_w \in Q_1$ be the state $\delta^*(q_0, \vec{0} \star w)$. $\mathcal{A}$ admits a cycle rooted at q_w. Hence, by Lemma 6.8, and thanks to Theorem 2.51, there exist $\vec{v_w} \in \mathbb{Q}^n \cap [0,1]^n$ and $\ell_w \in \mathbb{N}_{>0}$ such that $S(q_w)$ is r^{ℓ_w}-product-stable with respect to the pivot $\vec{v_w}$. By Lemma 6.5, $S(q_w)$ is (resp. weakly) recognizable in all bases for which S is (resp. weakly) recognizable. Moreover, we have

$$S = \bigcup_{w \in L} \frac{S(q_w) + \langle \vec{0} w \rangle_{r,n}}{r^{\ell_w}}.$$

Since L is finite, the conclusion follows. ■

Lemma 6.9 provides an argument for reducing the problem of proving Claim 6.3 : It suffices to establish that this claim holds for subsets of $[0,1]^n$ that are r^ℓ-product-stable ($\ell \in \mathbb{N}_{>0}$) with respect to a pivot $\vec{v} \in [0,1]^n \cap \mathbb{Q}^n$, i.e., to prove the following claim.

Claim 6.10 *Let $n \in \mathbb{N}_{>1}$ be a dimension, and $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors (resp. multiplicatively independent bases), and let $S \subseteq [0,1]^n$ be a both (resp. weakly) r- and (resp. weakly) s-recognizable set. Suppose that there exist $\ell \in \mathbb{N}_{>0}$ and $\vec{v} \in [0,1]^n \cap \mathbb{Q}^n$ such that S is r^ℓ-product-stable with respect to the pivot $\vec{v}$. If, for all $m \in \{1, 2, \ldots, n-1\}$, and for each set $S' \subseteq [0,1]^m$ that is both (resp. weakly) r- and (resp. weakly) s-recognizable, S' is definable in $\langle \mathbb{R}, +, <, 1\rangle$, then S is definable in $\langle \mathbb{R}, +, <, 1\rangle$.*

The following theorem is aimed at formalizing that reduction by proving that the r^ℓ-product-stability of Claim 6.10 is unnecessary. In other words, we prove that Claim 6.10 implies Claim 6.3.

Theorem 6.11 *Claim 6.10 implies Claim 6.3.*

Proof Assume that Claim 6.10 holds. Let S be a set satisfying the hypothesis of Claim 6.3. S is (resp. weakly) r-recognizable. Hence, by Lemma 6.9, there exist $k \in \mathbb{N}$, $\vec{i_1}, \vec{i_2}, \ldots, \vec{i_k} \in \mathbb{N}^n$, $\ell_1, \ell_2, \ldots, \ell_k \in \mathbb{N}_{>0}$, $\vec{v_1}, \vec{v_2}, \ldots \vec{v_k} \in [0,1]^n \cap \mathbb{Q}^n$, and $S_1, S_2, \ldots, S_k \subseteq [0,1]^n$ such that, for each $j \in \{1, 2, \ldots, k\}$, S_j is r^{ℓ_j}-product-stable with respect to the pivot $\vec{v_j}$ and is (resp. weakly) recognizable in all bases in which S is (resp. weakly) recognizable, and

$$S = \bigcup_{j \in \{1,2,\ldots,k\}} \frac{S_j + \vec{i_j}}{r^{\ell_j}}.$$

The sets S_j satisfy the hypothesis of Claim 6.10. It follows that they are definable in $\langle \mathbb{R}, +, <, 1\rangle$. Hence, the union

$$\bigcup_{j \in \{1,2,\ldots,k\}} \frac{S_j + \vec{i_j}}{r^{\ell_j}} = S$$

is definable in $\langle \mathbb{R}, +, <, 1\rangle$ as well. ■

In the remaining of this chapter, we prove that Claim 6.10 holds.

6.4 Multiple product stabilities

We consider a set $S \subseteq [0,1]^n$ that is either r- and s-recognizable in two bases r and s that do not share the same set of prime factors, or weakly r- and weakly s-recognizable in two multiplicatively independent bases r and s. As discussed in Section 6.3, we suppose that S is r^ℓ-product-stable with respect to a pivot $\vec{v} \in [0,1]^n \cap \mathbb{Q}^n$, with $\ell \in \mathbb{N}_{>0}$. Moreover, we make the assumption that for all $m \in \{1, 2, \ldots, n-1\}$, each set $S' \subseteq [0,1]^m$ that is either r- and s-recognizable, in the case of general recognizability, or weakly r- and weakly s- recognizable, in the case of weak recognizability, is necessarily definable in $\langle \mathbb{R}, +, <, 1 \rangle$.

The goal of the current section, as well as the following one, will be to establish that these hypotheses imply that S is definable in $\langle \mathbb{R}, +, <, 1 \rangle$.

In this section, we prove that the (resp. weakly) s-recognizability of S implies that, in addition to being r^ℓ-product-stable by hypothesis, S is $s^{\ell'}$-product-stable as well, for some $\ell' \in \mathbb{N}_{>0}$. We first consider the case where S is r^ℓ-product-stable with respect to the pivot $\vec{0}$.

Lemma 6.12 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases, and $S \subseteq [0,1]^n$ be a set that is both (resp. weakly) r- and s-recognizable. Suppose that there exists $\ell \in \mathbb{N}_{>0}$ such that S is r^ℓ-product-stable with respect to the pivot $\vec{0}$. There exists $\ell' \in \mathbb{N}_{>0}$ such that S is $s^{\ell'}$-product-stable with respect to the pivot $\vec{0}$.*

Proof Let $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ be a Muller RVA recognizing S in base s. We assume, without loss of generality, that $\mathcal{A}$ has a complete transition relation. Since $\mathcal{A}$ has finitely many states, the path reading $\vec{0}\star\vec{0}^\omega$ is ultimately cycling, i.e, there exist $q \in Q$, $m \in \mathbb{N}$ and $\ell' \in \mathbb{N}_{>0}$ such that $\delta^*(q_0, \vec{0} \star \vec{0}^m) = q = \delta^*(q, \vec{0}^{\ell'})$. Figure 6.3 illustrates this property.

By Definition 6.4 and since $w \in L(q) \Leftrightarrow \vec{0}^{\ell'} w \in L(q)$, we have, for all $\vec{x} \in [0,1]^n$, $\vec{x} \in S(q) \Leftrightarrow \dfrac{\vec{x}}{s^{\ell'}} \in S(q)$. It follows that $S(q)$ is $s^{\ell'}$-product-stable with respect to the pivot $\vec{0}$.

Let $\vec{x} \in [0,1]^n$. We have

$$\begin{aligned}
\vec{x} \in S(q) \quad &\Leftrightarrow \quad (\exists w \in (\{0, 1, \ldots, s-1\}^n)^\omega) \\
&\qquad (\langle \vec{0} \star w \rangle_s = \vec{x} \wedge \langle \vec{0} \star \vec{0}^m w \rangle_s \in S) \\
&\Leftrightarrow \quad \frac{1}{s^m} \vec{x} \in S.
\end{aligned}$$

Using this relation and the r^ℓ-product-stability hypothesis on S, one can establish that the set $S(q)$ is r^ℓ-product-stable as well. Indeed, let $\vec{x} \in [0,1]^n$.

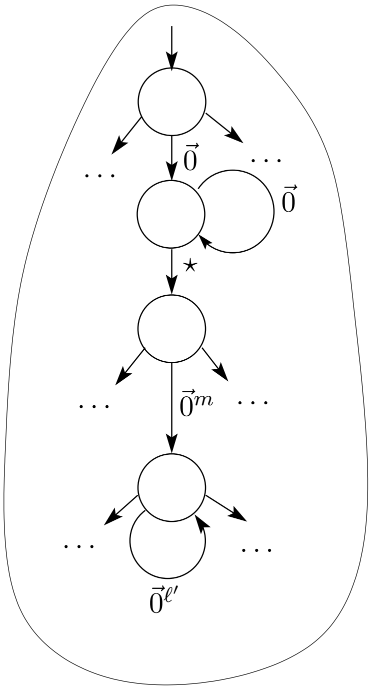

Figure 6.3: Illustration of Lemma 6.12.

We have

$$\begin{aligned}\vec{x} \in S(q) &\Leftrightarrow \frac{1}{s^m}\vec{x} \in S \\ &\Leftrightarrow \frac{1}{s^m r^\ell}\vec{x} \in S \\ &\Leftrightarrow \frac{1}{r^\ell}\vec{x} \in S(q),\end{aligned}$$

since $\frac{1}{r^\ell}\vec{x} \in [0,1]^n$.

The set $S(q)$ is thus both r^ℓ- and $s^{\ell'}$-product-stable, with respect to the same pivot $\vec{0}$. Let us show that these properties imply that S itself is both r^ℓ- and $s^{\ell'}$-product-stable with respect to $\vec{0}$. By hypothesis, S is r^ℓ-product-stable with respect to $\vec{0}$. Let $\vec{x}$ be an arbitrary vector of $[0,1]^n$, and let $k \in \mathbb{N}$ be large enough to have $r^{\ell k} \geq s^m$. We thus have $\frac{s^m}{r^{\ell k}} \leq 1$ and $\frac{s^m}{r^{\ell k}}\vec{x} \in [0,1]^n$, and we get

$$\begin{aligned}\vec{x} \in S &\Leftrightarrow \frac{1}{r^{\ell k}}\vec{x} \in S \\ &\Leftrightarrow \frac{s^m}{r^{\ell k}}\vec{x} \in S(q).\end{aligned} \tag{6.2}$$

Since $S(q)$ is $s^{\ell'}$-product-stable with respect to $\vec{0}$, we obtain

$$\begin{aligned}\frac{s^m}{r^{\ell k}}\vec{x} \in S(q) &\Leftrightarrow \frac{s^m}{r^{\ell k}s^{\ell'}}\vec{x} \in S(q) \\ &\Leftrightarrow \frac{1}{r^{\ell k}s^{\ell'}}\vec{x} \in S \\ &\Leftrightarrow \frac{1}{s^{\ell'}}\vec{x} \in S.\end{aligned} \tag{6.3}$$

By combining Equations 6.2 and 6.3, the $s^{\ell'}$-product-stability of S with respect to the pivot $\vec{0}$ follows. ■

We now consider the general case where S is r^ℓ-product-stable with respect to a pivot $\vec{v} \in [0,1]^n \cap \mathbb{Q}^n$. The proof of the following lemma decomposes the set S according to the 2^n possible positions of vectors in $[0,1]^n$ with respect to $\vec{v}$, in order to apply Lemma 6.12.

Lemma 6.13 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases, and $S \subseteq [0,1]^n$ be a set that is both (resp. weakly) r- and s-recognizable. Suppose that there exist $\ell \in \mathbb{N}_{>0}$ and $\vec{v} \in [0,1]^n \cap \mathbb{Q}^n$ such that S is r^ℓ-product-stable with respect to the pivot $\vec{v}$. There exists $\ell' \in \mathbb{N}_{>0}$ such that S is $s^{\ell'}$-product-stable with respect to the pivot $\vec{v}$.*

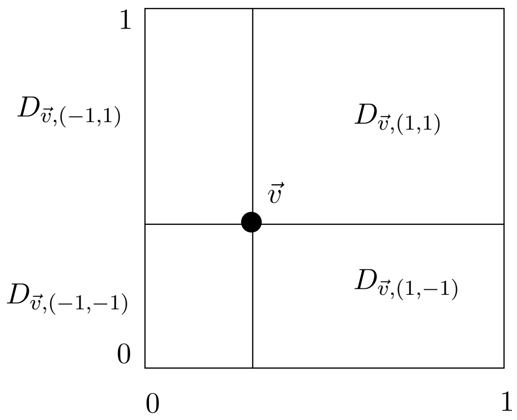

Figure 6.4: Sets $D_{\vec{v},\vec{a}}$, with $\vec{a} \in \{-1,1\}^2$.

Proof We consider the pivot $\vec{v}$. For each vector $\vec{a} \in \{-1,1\}^n$, let $M_{\vec{a}}$ be the matrix

$$M_{\vec{a}} = \mathrm{diag}(\vec{a}),$$

and $D_{\vec{v},\vec{a}}$ be the set

$$D_{\vec{v},\vec{a}} = (\vec{v} + M_{\vec{a}}(\mathbb{R}_{\geq 0})^n) \cap [0,1]^n.$$

The n-cube $[0,1]^n$ is equal to the union of the sets $D_{\vec{v},\vec{a}}$. More precisely, each set $D_{\vec{v},\vec{a}}$ is a Cartesian product of intervals

$$D_{\vec{v},\vec{a}} = I_1 \times I_2 \times \cdots \times I_n,$$

where for all $i \in \{1,2,\ldots,n\}$,

$$I_i = \begin{cases} [0, \vec{v}[i]] & \text{if } \vec{a}[i] = -1; \\ [\vec{v}[i], 1] & \text{if } \vec{a}[i] = 1. \end{cases}$$

Figure 6.4 illustrates the decomposition of $[0,1]^n$ into the sets $D_{\vec{v},\vec{a}}$ in the case $n = 2$.

Let $A_{\vec{v}}$ be the set of vectors $\vec{a} \in \{-1,1\}^n$ such that $D_{\vec{v},\vec{a}}$ has a non-zero volume, i.e., such that $f_{\vec{a}}(\vec{v}[i]) > 0$ for all $i \in \{1,2,\ldots,n\}$, where $f_{\vec{a}}(\vec{v}[i])$ denotes the length of the interval I_i, that is,

$$f_{\vec{a}}(\vec{v}[i]) = \begin{cases} \vec{v}[i] & \text{if } \vec{a}[i] = -1; \\ 1 - \vec{v}[i] & \text{if } \vec{a}[i] = 1. \end{cases}$$

We have

$$\bigcup_{\vec{a} \in A_{\vec{v}}} D_{\vec{v},\vec{a}} = [0,1]^n, \tag{6.4}$$

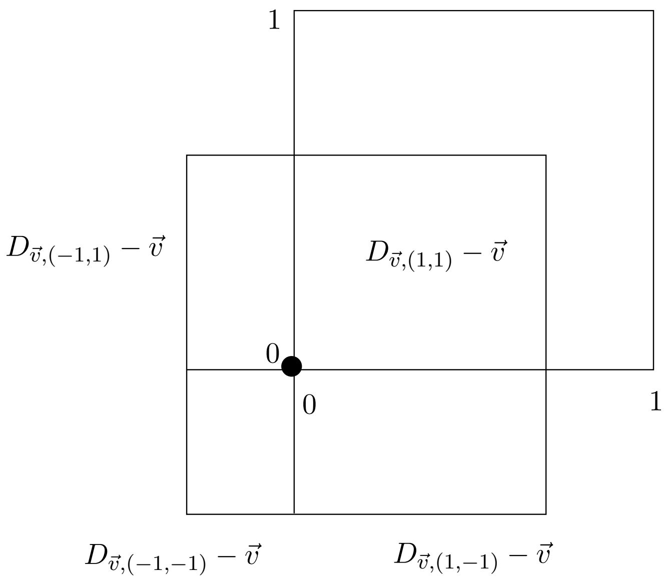

Figure 6.5: Sets $D_{\vec{v},\vec{a}} - \vec{v}$, with $\vec{a} \in \{-1,1\}^2$.

since each set $D_{\vec{v},\vec{a}}$ that has a zero volume is included into at least one set with non-zero volume. Indeed, let $\vec{a} \notin A$, and suppose that $D_{\vec{v},\vec{a}} = I_1 \times I_2 \times \cdots \times I_n$. We have

$$D_{\vec{v},\vec{a}} \subseteq I'_1 \times I'_2 \times \cdots \times I'_n,$$

where, for all $i \in \{1,2,\ldots,n\}$, I'_i is defined by

$$I'_i = \begin{cases} [0,1] & \text{if } I_i = \{0\} \text{ or } I_1 = \{1\}; \\ I_i & \text{otherwise.} \end{cases}$$

For each $\vec{a} \in A$, the set $M_{\vec{a}}(D_{\vec{v},\vec{a}} - \vec{v})$ takes the form of the Cartesian product $[0, f_{\vec{a}}(\vec{v}[1])] \times [0, f_{\vec{a}}(\vec{v}[2])] \times \cdots \times [0, f_{\vec{a}}(\vec{v}[n])]$, as shown in Figures 6.5 and 6.6 that respectively depict the sets $D_{\vec{v},\vec{a}} - \vec{v}$ and $M_{\vec{a}}(D_{\vec{v},\vec{a}} - \vec{v})$. Since $\vec{a} \in A$, we have $f_{\vec{a}}(\vec{v}[i]) > 0$ for each $i \in \{1,2,\ldots,n\}$ and we can thus map the elements of $D_{\vec{a}}$ onto $[0,1]^n$ by defining the transformation $\vec{x} \mapsto C_{\vec{v},\vec{a}} M_{\vec{a}}(\vec{x} - \vec{v})$, where

$$C_{\vec{v},\vec{a}} = \operatorname{diag}\left(\frac{1}{f_{\vec{a}}(\vec{v}[1])}, \frac{1}{f_{\vec{a}}(\vec{v}[2])}, \ldots, \frac{1}{f_{\vec{a}}(\vec{v}[n])}\right).$$

For each $\vec{a} \in A$, we now study the properties of the set

$$S_{\vec{v},\vec{a}} = C_{\vec{v},\vec{a}} M_{\vec{a}}((S \cap D_{\vec{v},\vec{a}}) - \vec{v}).$$

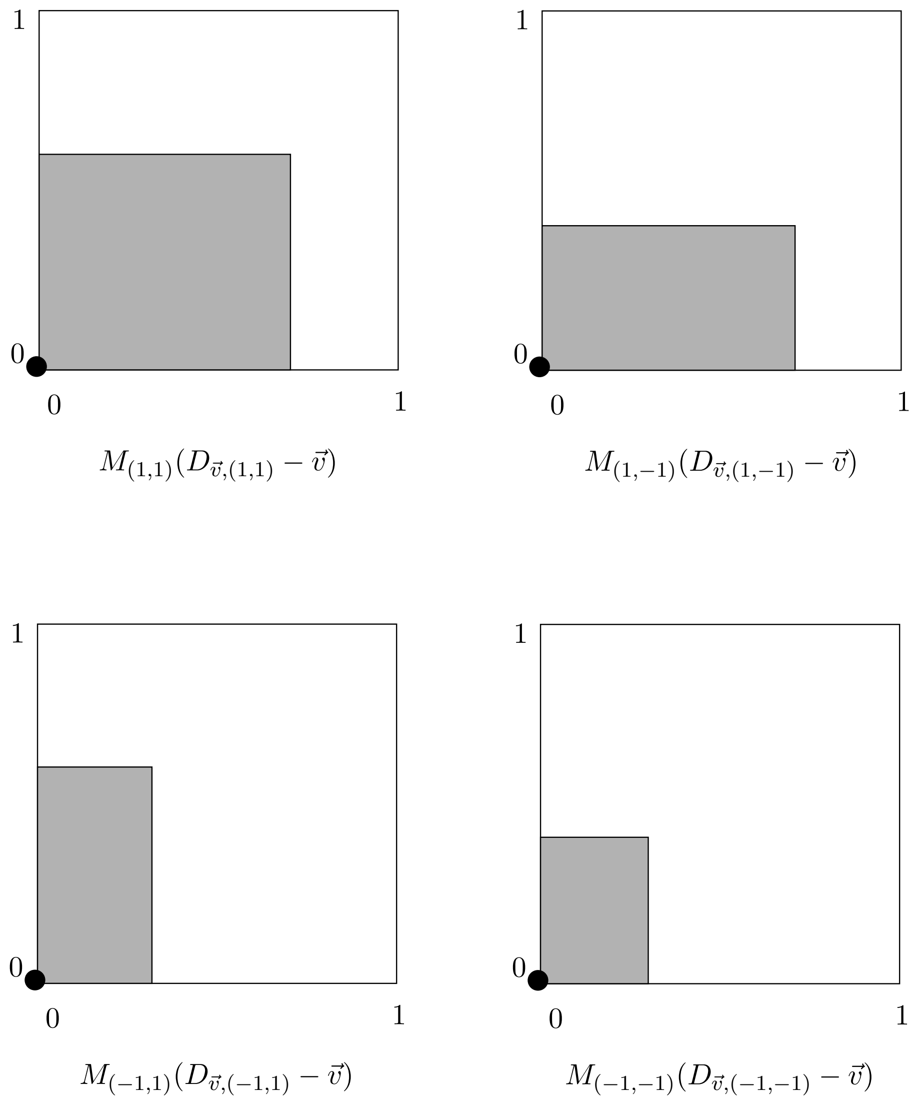

Figure 6.6: Sets $M_{\vec{a}}(D_{\vec{v},\vec{a}} - \vec{v})$, with $\vec{a} \in \{-1,1\}^2$.

- Since $D_{\vec{v},\vec{a}}$ is a Cartesian product of intervals with rational boundaries, it is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$ and is then weakly r- and weakly s-recognizable by Theorem 3.31. Hence, from Theorems 2.65, 3.34 and 3.35, and since the matrices $C_{\vec{v},\vec{a}}$ and $M_{\vec{a}}$ are non-singular, the set $S_{\vec{v},\vec{a}}$ is (resp. weakly) r- and s-recognizable.

- $S_{\vec{v},\vec{a}}$ is r^ℓ-product-stable with respect to the pivot $\vec{0}$. Indeed, let $\vec{x} \in [0,1]^n$. We have

$$\begin{array}{rcl} \vec{x} \in S_{\vec{v},\vec{a}} & \Leftrightarrow & \vec{x} \in C_{\vec{v},\vec{a}} M_{\vec{a}}((S \cap D_{\vec{v},\vec{a}}) - \vec{v}) \\ & \Leftrightarrow & M_{\vec{a}}^{-1} C_{\vec{v},\vec{a}}^{-1} \vec{x} \in (S \cap D_{\vec{v},\vec{a}}) - \vec{v}. \end{array}$$

  Since S is r^ℓ-product-stable with respect to $\vec{v}$ and since $M_{\vec{a}}^{-1} C_{\vec{v},\vec{a}}^{-1} \vec{x} \in [0,1]^n$, we get

$$\begin{array}{rcl} \vec{x} \in S_{\vec{v},\vec{a}} & \Leftrightarrow & \frac{1}{r^\ell} M_{\vec{a}}^{-1} C_{\vec{v},\vec{a}}^{-1} \vec{x} \in (S \cap D_{\vec{v},\vec{a}}) - \vec{v} \\ & \Leftrightarrow & M_{\vec{a}}^{-1} C_{\vec{v},\vec{a}}^{-1} \frac{1}{r^\ell} \vec{x} \in (S \cap D_{\vec{v},\vec{a}}) - \vec{v} \\ & \Leftrightarrow & \frac{1}{r^\ell} \vec{x} \in S_{\vec{v},\vec{a}}. \end{array}$$

- Thanks to the two previous properties, Lemma 6.12 can be applied on $S_{\vec{v},\vec{a}}$. Hence, there exists $\ell'_{\vec{v},\vec{a}} \in \mathbb{N}_{>0}$ such that $S_{\vec{v},\vec{a}}$ is $s^{\ell'_{\vec{v},\vec{a}}}$-product-stable with respect to the pivot $\vec{0}$.

A corollary of the $s^{\ell'_{\vec{v},\vec{a}}}$-product-stability of $S_{\vec{v},\vec{a}}$ with respect to the pivot $\vec{0}$ is that $S \cap D_{\vec{v},\vec{a}}$ is $r^{\ell'_{\vec{v},\vec{a}}}$-product-stable with respect to the pivot $\vec{v}$. Indeed, let $\vec{x} \in [0,1]^n - \vec{v}$. We successively have

$$\begin{array}{rcl} \vec{x} \in (S \cap D_{\vec{v},\vec{a}}) - \vec{v} & \Leftrightarrow & \vec{x} \in M_{\vec{a}}^{-1} C_{\vec{v},\vec{a}}^{-1} S_{\vec{v},\vec{a}} \\ & \Leftrightarrow & \frac{1}{s^{\ell'_{\vec{v},\vec{a}}}} \vec{x} \in M_{\vec{a}}^{-1} C_{\vec{v},\vec{a}}^{-1} S_{\vec{v},\vec{a}} \\ & \Leftrightarrow & \frac{1}{s^{\ell'_{\vec{v},\vec{a}}}} \in (S \cap D_{\vec{v},\vec{a}}). \end{array}$$

From Equation 6.4, it follows that

$$S = \bigcup_{\vec{a} \in A} (S \cap D_{\vec{v},\vec{a}}),$$

which implies that S is $s^{\ell'}$-product-stable with respect to $\vec{v}$, where $\ell' = \mathrm{lcm}_{\vec{a} \in A}(\ell'_{\vec{v},\vec{a}})$. ■

6.5 Exploiting multiple product stabilities

In the previous section, we established than any (resp. weakly) r- and s-recognizable subset of $[0,1]^n$ such that S is r^ℓ-product-stable ($\ell \in \mathbb{N}_{>0}$) with respect to a pivot $\vec{v} \in [0,1]^n \cap \mathbb{Q}^n$ is, in addition, $s^{\ell'}$-product-stable ($\ell' \in \mathbb{N}_{>0}$) with respect to the same pivot $\vec{v}$.

Recall that our inductive hypotheses state that r and s are bases with different sets of prime factors (resp. multiplicatively independent bases), and that the simultaneously (resp. weakly) r- and s-recognizable sets $S' \subseteq [0,1]^m$, with $m \in \{1,2,\ldots,n-1\}$, are definable in $\langle \mathbb{R},+,<,1\rangle$. In this section, we combine those hypotheses with the results of the previous section for proving that the sets $S \subseteq [0,1]^n$ that are r^ℓ-product stable, with $\ell \in \mathbb{N}_{>0}$, as well as simultaneously r- and s-recognizable, are definable in $\langle \mathbb{R},+,<,1\rangle$.

Let $\vec{v}$ and $\vec{x}$ be two distinct vectors in $[0,1]^n \cap \mathbb{Q}^n$. The set $h_{\vec{v}}(\vec{x})$ is defined as the line segment containing $\vec{x}$, and whose extremities are $\vec{v}$ and a point of a face of the n-cube $[0,1]^n$, as formalized in the following definition.

Definition 6.14 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $\vec{v},\vec{x} \in [0,1]^n \cap \mathbb{Q}^n$ be such that $\vec{v} \neq \vec{x}$. The set $h_{\vec{v}}(\vec{x})$ is*

$$h_{\vec{v}}(\vec{x}) = \{\vec{v} + \lambda(\vec{x} - \vec{v}) \in [0,1]^n \mid \lambda \in \mathbb{R}_{>0}\}.$$

□

This definition implies that $\vec{v} \notin h_{\vec{v}}(\vec{x})$. In the case $n = 2$, and for particular values $\vec{v}$ and $\vec{x}$, the set $h_{\vec{v}}(\vec{x})$ is depicted in Figure 6.7.

The next lemma is aimed at establishing that, for all distinct vectors $\vec{v},\vec{x}$ in $[0,1]^n \cap \mathbb{Q}^n$, the set $h_{\vec{v}}(\vec{x})$ is definable in $\langle \mathbb{R},+,<,1\rangle$.

Lemma 6.15 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $\vec{v},\vec{x} \in [0,1]^n \cap \mathbb{Q}^n$ be such that $\vec{v} \neq \vec{x}$. The set $h_{\vec{v}}(\vec{x})$ is definable in $\langle \mathbb{R},+,<,1\rangle$.*

Proof Let $l_{\vec{v}}(\vec{x})$ be the line $\{\vec{v} + \lambda(\vec{x} - \vec{v}) \mid \lambda \in \mathbb{R}\}$, and let $\{i_1, i_2, \ldots, i_k\} \subseteq \{1,2,\ldots,n\}$ be the set of indexes such that $i \in \{i_1,i_2,\ldots,i_k\}$ if and only if $\vec{x}[i] \neq \vec{v}[i]$. The line $l_{\vec{v}}(\vec{x})$ can be expressed as the intersection of linear

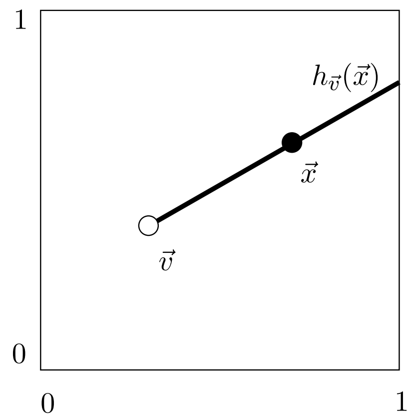

Figure 6.7: Set $h_{\vec{v}}(\vec{x})$, in dimension 2.

equalities

$$l_{\vec{v}}(\vec{x}) = \bigcap_{i \in \{1,2,\ldots,n\} \setminus \{i_1, i_2, \ldots, i_k\}} \{\vec{x'} \in \mathbb{R}^n \mid \vec{x'}[i] = \vec{v}[i]\} \cap \left\{ \vec{x'} \in \mathbb{R}^n \mid \frac{\vec{x'}[i_1] - \vec{v}[i_1]}{\vec{x}[i_1] - \vec{v}[i_1]} = \frac{\vec{x'}[i_2] - \vec{v}[i_2]}{\vec{x}[i_2] - \vec{v}[i_2]} = \cdots = \frac{\vec{x'}[i_k] - \vec{v}[i_k]}{\vec{x}[i_k] - \vec{v}[i_k]} \right\}.$$

Note that the coefficients of these equalities are rationals.

The set $h_{\vec{v}}(\vec{x})$ is the intersection of $l_{\vec{v}}(\vec{x})$ and the Cartesian product of intervals

$$I_1 \times I_2 \times \cdots \times I_n,$$

where, for all $i \in \{1, 2, \ldots, n\}$,

$$I_i = \begin{cases} [0, \vec{v}[i]] & \text{if } \vec{x}[i] \leq \vec{v}[i]; \\ [\vec{v}[i], 1] & \text{otherwise,} \end{cases}$$

from which the point $\vec{v}$ has to be removed.

Remark that the set $I_1 \times I_2 \times \cdots \times I_n$ can be expressed as a finite Boolean combination of linear inequalities with rational coefficients. Indeed, for each $i \in \{1, 2, \ldots, n\}$, let

$$S_i = \{\vec{x'} \in \mathbb{R}^n \mid 0 \leq \vec{x'}[i] \leq \vec{v}[i]\}$$

if $\vec{x}[i] \leq \vec{v}[i]$ and

$$S_i = \{\vec{x'} \in \mathbb{R}^n \mid \vec{v}[i] \leq \vec{x'}[i] \leq 1\}$$

otherwise. We have

$$I_1 \times I_2 \times \cdots \times I_n = \bigcap_{i \in \{1,2,\ldots,n\}} S_i.$$

Since we have

$$h_{\vec{v}}(\vec{x}) = (l_{\vec{v}}(\vec{x}) \cap (I_1 \times I_2 \cdots \times I_n)) \setminus \{\vec{v}\},$$

the set $h_{\vec{v}}(\vec{x})$ is a finite Boolean combination of linear constraints with rational coefficients, which implies that $h_{\vec{v}}(\vec{x})$ is definable in $\langle \mathbb{R}, +, <, 1 \rangle$. ■

We now prove the following lemma, stating that, under the above hypotheses, each set $h_{\vec{v}}(\vec{x})$, with $\vec{x} \neq \vec{v}$, is either included in S or in $[0,1]^n \setminus S$.

Lemma 6.16 *Let $n \in \mathbb{N}_{>1}$ be a dimension, and $r, s \in \mathbb{N}_{>1}$ be bases that do not share the same set of prime factors (resp. multiplicatively independent bases). Let $S \subseteq [0,1]^n$ be a set that is both (resp. weakly) r- and (resp. weakly) s-recognizable. Suppose that there exist $\ell \in \mathbb{N}_{>0}$ and $\vec{v} \in [0,1]^n \cap \mathbb{Q}^n$ such that S is r^ℓ-product-stable with respect to the pivot $\vec{v}$. Suppose that, for all $m \in \{1, 2, \ldots, n-1\}$, and for each set $S' \subseteq [0,1]^m$ that is both (resp. weakly) r- and (resp. weakly) s-recognizable, S' is definable in $\langle \mathbb{R}, +, <, 1 \rangle$. For each $\vec{x} \in [0,1]^n \cap \mathbb{Q}^n$ such that $\vec{x} \neq \vec{v}$, we have either $h_{\vec{v}}(\vec{x}) \subseteq S$, or $h_{\vec{v}}(\vec{x}) \cap S = \emptyset$.*

Proof By Lemma 6.13, there exists $\ell' \in \mathbb{N}_{>0}$ such that S is $s^{\ell'}$-product-stable with respect to the pivot $\vec{v}$.

Consider $\vec{x} \in [0,1]^n \cap \mathbb{Q}^n$ such that $\vec{x} \neq \vec{v}$. By Lemma 6.15, the set $h_{\vec{v}}(\vec{x})$ is definable in $\langle \mathbb{R}, +, <, 1 \rangle$. As a consequence of Theorem 3.31, is thus weakly recognizable in all bases. From Theorems 2.65 and 3.34, it follows that the set

$$S' = S \cap h_{\vec{v}}(\vec{x})$$

is (resp. weakly) r- and (resp. weakly) s-recognizable. Besides, S' is both r^ℓ and $s^{\ell'}$-product-stable with respect to the pivot $\vec{v}$; indeed, let $\vec{x'} \in [0,1]^n - \vec{v}$

and $(t, \ell'') \in \{(r, \ell), (s, \ell')\}$. Since S is itself $t^{\ell''}$-product-stable, we have

$$\begin{aligned}
& \vec{x'} \in (S \cap h_{\vec{v}}(\vec{x})) - \vec{v} \\
\Leftrightarrow\ & \vec{x'} \in S - \vec{v} \wedge \vec{x'} \in h_{\vec{v}}(\vec{x}) - \vec{v} \\
\Leftrightarrow\ & \frac{1}{t^{\ell''}}\vec{x'} \in S - \vec{v} \wedge \vec{x'} \in \{\lambda(\vec{x} - \vec{v}) \in [0,1]^n - \vec{v} \mid \lambda \in \mathbb{R}_{>0}\} \\
\Leftrightarrow\ & \frac{1}{t^{\ell''}}\vec{x'} \in S - \vec{v} \wedge \frac{1}{t^{\ell''}}\vec{x'} \in \{\lambda(\vec{x} - \vec{v}) \in [0,1]^n - \vec{v} \mid \lambda \in \mathbb{R}_{>0}\} \\
\Leftrightarrow\ & \frac{1}{t^{\ell''}}\vec{x'} \in (S \cap h_{\vec{v}}(\vec{x})) - \vec{v}.
\end{aligned}$$

Let $C \in \mathbb{Q}^{n\times n}$ be the matrix whose inverse is

$$C^{-1} = [\ (\vec{x} - \vec{v}) \quad \vec{v_1} \quad \vec{v_2} \quad \cdots \quad \vec{v_{n-1}}\ ],$$

where $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_{n-1}} \in \mathbb{Q}^n$ are such that the vectors $(\vec{x} - \vec{v}), \vec{v_1}, \vec{v_2}, \ldots, \vec{v_{n-1}}$ are linearly independent. For instance, $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_{n-1}}$ can be chosen by selecting, among $\{(1, 0, \ldots, 0), (0, 1, \ldots, 0), (0, 0, \ldots, 1)\}$, $n - 1$ vectors that are linearly independent with $\vec{x} - \vec{v}$. These $n-1$ vectors necessarily exist since the vectors $\{(1, 0, \ldots, 0), (0, 1, \ldots, 0), (0, 0, \ldots, 1)\}$ are linearly independent and since the vector $\vec{x} - \vec{v} \neq \vec{0}$ is linearly dependent with at most one of them.

We have $C^{-1}(1, 0, \ldots, 0) = \vec{x} - \vec{v}$. Hence, for all $\lambda \in \mathbb{R}$, we have $\lambda C(\vec{x} - \vec{v}) = (\lambda, 0, \ldots, 0)$. Since for all $\vec{x'} \in h_{\vec{v}}(\vec{x})$ we have $C(\vec{x'} - \vec{v}) = C((\vec{v} + \lambda(\vec{x} - \vec{v})) - \vec{v})$, with $\lambda \in \mathbb{R}_{>0}$, we also have $C(\vec{x'} - \vec{v}) = C(\lambda(\vec{x} - \vec{v})) = (\lambda, 0, \ldots, 0)$, which is illustrated in Figure 6.8, that depicts the set $C(h_{\vec{v}}(\vec{x}) - \vec{v})$ in dimension 2. Note that $\vec{0} \notin C(h_{\vec{v}}(\vec{x}) - \vec{v})$ since $\vec{v} \notin h_{\vec{v}}(\vec{x})$.

By Theorems 2.65, 3.34 and 3.35, the set

$$S'' = C(S' - \vec{v}) \cap [0,1]^n$$

is (resp. weakly) r- and (resp. weakly) s-recognizable. Moreover, S'' is both r^{ℓ}- and $s^{\ell'}$-product-stable with respect to the pivot $\vec{0}$ since, for all $\vec{x'} \in [0,1]^n$, we have

$$\begin{aligned}
\vec{x'} \in S'' \ &\Leftrightarrow\ \vec{x'} \in C(S' - \vec{v}) \cap [0,1]^n \\
&\Leftrightarrow\ (\exists \vec{x''} \in S' - \vec{v})(\vec{x'} = C\vec{x''} \wedge \vec{x'} \in [0,1]^n).
\end{aligned}$$

For all $(t, \ell'') \in \{(r, \ell), (s, \ell')\}$, the set S' is $t^{\ell''}$-product-stable with respect to $\vec{v}$. Hence, we obtain

$$\begin{aligned}
\vec{x'} \in S'' \ &\Leftrightarrow\ (\exists \vec{x''} \in (S' - \vec{v}))(\vec{x'} = t^{\ell''}C\vec{x''} \wedge \vec{x'} \in [0,1]^n) \\
&\Leftrightarrow\ \frac{1}{t^{\ell''}}\vec{x'} \in S''.
\end{aligned}$$

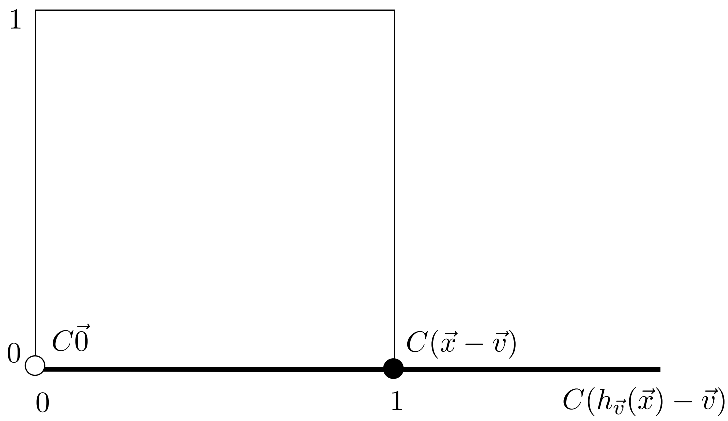

Figure 6.8: Set $C(h_{\vec{v}}(\vec{x})-\vec{v})$, in dimension 2.

Note that the set S'' can be decomposed into

$$S'' = S''' \times \{0\}^{n-1},$$

with $S''' \subseteq [0,1]$. Applying Theorem 2.65 (resp. Theorem 3.36), the set S''' is (resp. weakly) r- and (resp. weakly) s-recognizable. Since S''' is of dimension $1 < n$, our hypotheses state that it is definable in the theory $\langle \mathbb{R}, +, <, 1\rangle$. In other words, S''' is equal to a finite union of intervals with rational boundaries, as discussed in Section 2.2.2.

In addition, we know that S''' is both r^ℓ- and $s^{\ell'}$-product-stable with respect to the pivot 0. Since r and s are multiplicatively independent, r^ℓ and $s^{\ell'}$ are multiplicatively independent as well. By Lemma 5.18, the set

$$\left\{ \frac{r^{\ell i}}{s^{\ell' j}} \mid i, j \in \mathbb{N}_{>0} \right\}$$

is dense in $\mathbb{R}_{>0}$. It follows that if $1 \in S'''$ then $S''' =]0,1]$ and if $1 \notin S'''$, then $S''' = \emptyset$.

To conclude, it remains to prove that $S' = S \cap h_{\vec{v}}(\vec{x})$ is either empty, or equal to $h_{\vec{v}}(\vec{x})$. Let $\vec{x'} \in [0,1]^n$. For all $k \in \mathbb{N}$, we have

$$\begin{aligned}
& \vec{x'} \in S' \\
\Leftrightarrow\ & \vec{x'} - \vec{v} \in S' - \vec{v} \\
\Leftrightarrow\ & \frac{1}{r^{\ell k}}(\vec{x'} - \vec{v}) \in S' - \vec{v} \\
\Leftrightarrow\ & C\left(\frac{1}{r^{\ell k}}(\vec{x'} - \vec{v})\right) \in C(S' - \vec{v}).
\end{aligned}$$

If k is chosen to be sufficiently large, we have $C\left(\frac{1}{r^{\ell k}}(\vec{x}-\vec{v})\right) \in [0,1]^n$. Hence, we get

$$\begin{aligned} & \vec{x'} \in S' \\ \Leftrightarrow\ & C\left(\frac{1}{r^{\ell k}}(\vec{x}-\vec{v})\right) \in S'' \\ \Leftrightarrow\ & 1 \in S'''. \end{aligned}$$

Since S''' is either equal to $]0,1]$ or to $\emptyset$, the conclusion follows. ■

Intuitively, Lemma 6.16 hints at the fact that the set S has a conical structure. We formalize this property by the following definition.

Definition 6.17 *Let $n \in \mathbb{N}_{>0}$ be a dimension. A set $S \subseteq [0,1]^n$ is a* bounded conical set *of apex $\vec{v} \in [0,1]^n$ if and only if for all $\vec{x} \in [0,1]^n, f \in]0,1]$, we have*

$$\vec{x} \in S \Leftrightarrow \vec{v} + f(\vec{x}-\vec{v}) \in S.$$

□

In other words, a bounded conical set is entirely determined by its apex, and its intersection with the faces of the n-cube $[0,1]^n$. It follows that, in order to establish that S is definable in $\langle \mathbb{R}, +, <, 1\rangle$, it roughly suffices to show that this intersection is definable in the same theory, and that S is a bounded conical set.

The notion of face of the n-cube $[0,1]^n$ is defined as follows.

Definition 6.18 *Let $n \in \mathbb{N}_{>0}$ be a dimension, let $\lambda \in \{0,1\}$ and let $i \in \{1,2,\ldots,n\}$. The* face $F_{\lambda,i}$ *of the n-cube $[0,1]^n$ is defined by*

$$F_{\lambda,i} = \{\vec{x} \in [0,1]^n \mid \vec{x}[i] = \lambda\}.$$

□

The sets $F_{\lambda,i}$ are depicted in Figure 6.9 in the case $n = 2$.

The next lemma establishes that, under the conditions of Lemma 6.16, the intersection of S with the faces of the n-cube $[0,1]^n$ is definable in $\langle \mathbb{R}, +, <, 1\rangle$.

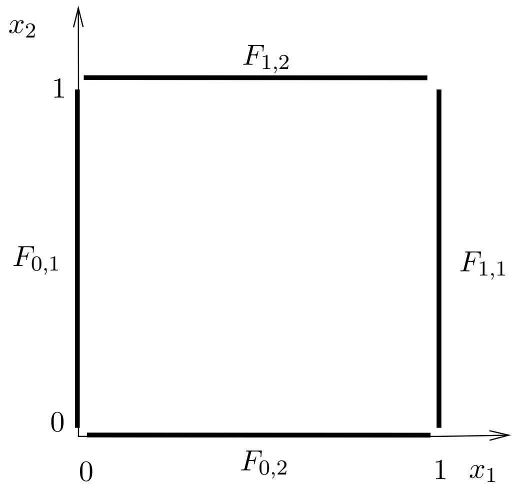

Figure 6.9: Faces of $[0,1]^2$.

Lemma 6.19 *Let $n \in \mathbb{N}_{>1}$ be a dimension, and $r, s \in \mathbb{N}_{>1}$ be bases that do not share the same set of prime factors (resp. multiplicatively independent bases). Let $S \subseteq [0,1]^n$ be a set that is both (resp. weakly) r- and (resp. weakly) s-recognizable. Suppose that there exist $\ell \in \mathbb{N}_{>0}$ and $\vec{v} \in [0,1]^n \cap \mathbb{Q}^n$ such that S is r^ℓ-product-stable with respect to the pivot $\vec{v}$. Suppose that, for all $m \in \{1,2,\ldots,n-1\}$, and for each set $S' \subseteq [0,1]^m$ that is both (resp. weakly) r- and (resp. weakly) s-recognizable, S' is definable in $\langle \mathbb{R}, +, <, 1\rangle$. For every $i \in \{1,2,\ldots,n\}$ and $\lambda \in \{0,1\}$, the set $S \cap F_{\lambda,i}$ is definable in $\langle \mathbb{R}, +, <, 1\rangle$.*

Proof Let $i \in \{1,2,\ldots,n\}$ and $\lambda \in \{0,1\}$.

We first build the permutation matrix $C \in \{0,1\}^{n\times n}$ such that $C\vec{x} = (\vec{x}[1],\ldots,\vec{x}[i-1],\vec{x}[i+1],\ldots,x[n],x[i])$ for any $\vec{x} \in \mathbb{R}^n$. Formally, C is the matrix

$$[\ \vec{v_1} \quad \vec{v_2} \quad \cdots \quad \vec{v_n}\ ],$$

where, for all $j \in \{1,2,\ldots,n\}$,

$$\vec{v_j} = \begin{cases} (\underbrace{0,\ldots,0}_{j-1},1,\underbrace{0,\ldots,0}_{n-j}) & \text{if } j < i; \\ (\underbrace{0,\ldots,0}_{n-1},1) & \text{if } j = i; \\ (\underbrace{0,\ldots,0}_{j-2},1,\underbrace{0,\ldots,0}_{n-j+1}) & \text{if } j > i. \end{cases}$$

The set $F_{\lambda,i}$ is clearly definable in $\langle \mathbb{R}, +, <, 1\rangle$. By Theorem 3.31, it is thus weakly r- and weakly s-recognizable. It follows from Theorem 2.65 (resp. Theorems 3.34 and 3.35) that the set

$$S' = C(S \cap F_{\lambda,i})$$

is (resp weakly) r- and (resp. weakly) s-recognizable as well.

Moreover, we have $S' = S'' \times \{\lambda\}$, with $S'' \subseteq [0,1]^{n-1}$. By Theorem 2.65 (resp. Theorem 3.36), the set S'' is (resp. weakly) r- and (resp. weakly) s-recognizable as well. Since S'' is of dimension $n-1$, the assumptions of the lemma imply that it is definable in $\langle \mathbb{R}, +, <, 1\rangle$.

Since C^{-1} exists and belongs to $\{0,1\}^{n\times n}$, it then follows that the set

$$S \cap F_{\lambda,i} = C^{-1}S' = C^{-1}(S'' \times \{\lambda\})$$

is definable in $\langle \mathbb{R}, +, <, 1\rangle$ as well. ■

We now address the problem of establishing that Lemmas 6.16 and 6.19 imply that the set S is itself definable in $\langle \mathbb{R}, +, <, 1\rangle$. For this purpose, we will prove that S coincides with a bounded conical set, in the sense of Definition 6.17, that is definable in $\langle \mathbb{R}, +, <, 1\rangle$. Our proof requires additional lemmas that show that any subset of $\mathbb{R}^{n-1}$ definable in $\langle \mathbb{R}, +, <, 1\rangle$ can be extended into a cone in dimension n, without influencing its definability in $\langle \mathbb{R}, +, <, 1\rangle$. The operation consisting in mapping a subset of $\mathbb{R}^{n-1}$ onto a cone in $\mathbb{R}^n$ is formally described by the following definition.

Definition 6.20 *Let $n \in \mathbb{N}_{>1}$ be a dimension, let $S \subseteq \mathbb{R}^{n-1}$, let $i \in \{1, 2, \ldots, n\}$, let $\lambda \in \mathbb{Q}$, and let $\vec{v} \in \mathbb{Q}^n$. The set* $\mathrm{Cone}(S, i, \lambda, \vec{v}) \subseteq [0,1]^n$ *is defined by*

$$\begin{aligned}\mathrm{Cone}(S, i, \lambda, \vec{v}) = \\ &\{\vec{x} \in \mathbb{R}^n \mid (\exists f \in \,]0,1])(\exists \vec{x'} \in S) \\ &(\vec{x} = f((\vec{x'}[1], \ldots, \vec{x'}[i-1], \lambda, \vec{x'}[i], \ldots, \vec{x'}[n-1]) - \vec{v}) + \vec{v})\}.\end{aligned}$$

□

Figure 6.10 illustrates Definition 6.20 in the case $n = 2$.

In the following lemma, we consider a set S described by a linear constraint with rational coefficients, and, in the particular case $\vec{v} = \vec{0}$ and $\lambda \neq 0$, we prove that the set $\mathrm{Cone}(S, i, \lambda, \vec{v})$ is an intersection of two linear constraints with rational coefficients. The proof is adapted from a construction of Chernikova's algorithm [LV92].

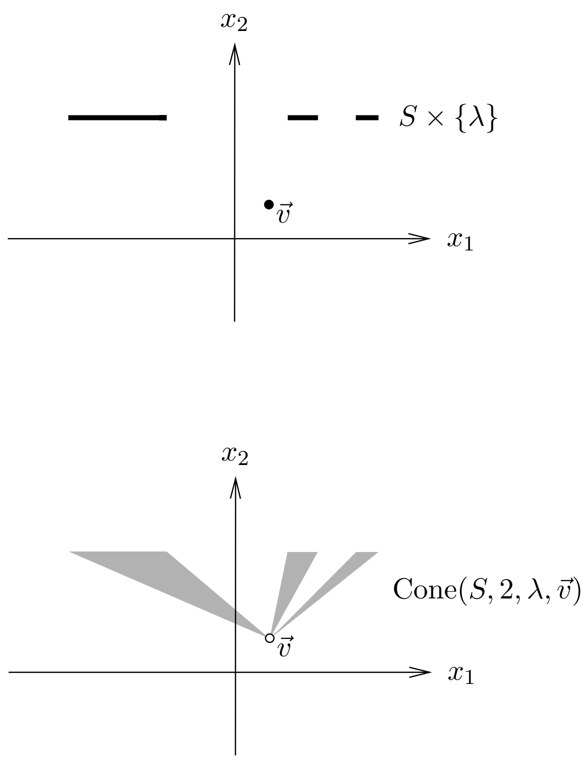

Figure 6.10: Example of set $\mathrm{Cone}(S, i, \lambda, \vec{v})$, in dimension 2.

Lemma 6.21 *Let $n \in \mathbb{N}_{>1}$ be a dimension, let $\vec{a} \in \mathbb{Q}^{n-1}$, let $b \in \mathbb{Q}$, let $i \in \{1, 2, \ldots, n\}$, and let $\lambda \in \mathbb{Q}_{\neq 0}$. Let S be the set $\{\vec{x} \in \mathbb{R}^{n-1} \mid \vec{a}\vec{x} \leq b\}$. There exists $\vec{a'} \in \mathbb{Q}^n$ such that*

$$\begin{aligned}\mathrm{Cone}(S, i, \lambda, \vec{0}) = \{\vec{x} \in \mathbb{R}^n \mid \vec{a'}\vec{x} \leq 0\} \\ \cap \begin{cases} \{\vec{x} \in \mathbb{R}^n \mid 0 < \vec{x}[i] \leq \lambda\} & \text{if } \lambda > 0; \\ \{\vec{x} \in \mathbb{R}^n \mid \lambda \leq \vec{x}[i] < 0\} & \text{if } \lambda < 0. \end{cases}\end{aligned}$$

Proof Let $\vec{a'} = \left(\vec{a}[1], \ldots, \vec{a}[i-1], -\dfrac{b}{\lambda}, \vec{a}[i], \ldots, \vec{a}[n-1]\right)$. Let us prove that the vector $\vec{a'}$ is the expected vector.

On the one hand, let $\vec{x} \in \mathrm{Cone}(S, i, \lambda, 0)$. By Definition 6.20, there exist $f \in \,]0, 1]$ and $\vec{x'} \in \mathbb{R}^{n-1}$ such that

$$\vec{x} = f(\vec{x'}[1], \ldots, \vec{x'}[i-1], \lambda, \vec{x'}[i], \ldots, \vec{x'}[n-1])$$

and $\vec{a}\vec{x'} \leq b$. We then have

$$\vec{a'}\vec{x} = \vec{a}\vec{x'} - \frac{b}{\lambda}\lambda \leq b - b = 0.$$

Moreover, since $\vec{x}[i] = f\lambda$ and $f \in \,]0, 1]$, we have $0 < \vec{x}[i] \leq \lambda$ (resp. $\lambda \leq \vec{x}[i] < 0$) if $\lambda > 0$ (resp. $\lambda < 0$).

On the other hand, let $\vec{x} \in \mathbb{R}^n$ such that $\vec{a'}\vec{x} \leq 0$ and $0 < \vec{x}[i] \leq \lambda$ (resp. $\lambda \leq \vec{x}[i] < 0$) if $\lambda > 0$ (resp. $\lambda < 0$). Let $\vec{x'} = \dfrac{\lambda}{\vec{x}[i]}\vec{x}$, and let $\vec{x''} = (\vec{x'}[1], \ldots, \vec{x'}[i-1], \vec{x'}[i+1], \ldots \vec{x'}[n])$. We thus have

$$\vec{a'}\vec{x'} = \vec{a}\vec{x''} - \frac{b}{\lambda}\lambda,$$

and

$$\vec{a}\vec{x''} = \vec{a'}\vec{x'} + b \leq b.$$

Since $\vec{x} = \dfrac{\vec{x}[i]}{\lambda}(\vec{x''}[1], \ldots, \vec{x''}[i-1], \lambda, \vec{x''}[i], \ldots, \vec{x''}[n-1])$, and $\dfrac{\vec{x}[i]}{\lambda} \in \,]0, 1]$, we then get $\vec{x} \in \mathrm{Cone}(S, i, \lambda, \vec{0})$. ■

Thanks to the previous lemma, we can now prove that the definability in $\langle \mathbb{R}, +, <, 1 \rangle$ of a set $S \subseteq [0, 1]^n$ implies the definability in the same theory of the set $\mathrm{Cone}(S, i, \lambda, \vec{0})$.

Lemma 6.22 *Let $n \in \mathbb{N}_{>1}$ be a dimension, let $S \subseteq \mathbb{R}^{n-1}$ be a set that is definable in $\langle \mathbb{R}, +, <, 1 \rangle$, let $i \in \{1, 2, \ldots, n\}$, and let $\lambda \in \mathbb{Q}_{\neq 0}$. The set* $\mathrm{Cone}(S, i, \lambda, \vec{0})$ *is definable in $\langle \mathbb{R}, +, <, 1 \rangle$ as well.*

Proof As noticed in Section 2.2.2, the theory $\langle \mathbb{R}, +, <, 1\rangle$ admits the elimination of quantifiers [FR75]. Hence, the set S can be expressed as a finite Boolean combination $\mathcal{B}$ of linear constraints with rational coefficients. It follows from Lemma 6.21 that, for each of the sets S' represented by these constraints, the set $\mathrm{Cone}(S', i, \lambda, \vec{0})$ is an intersection of three linear constraints with rational coefficients. Since the set $\mathrm{Cone}(S, i, \lambda, \vec{0})$ is equal to the Boolean combination $\mathcal{B}$ applied to the sets $\mathrm{Cone}(S', i, \lambda, \vec{0})$, where the complement operations have to be considered with respect to the set

$$\{\vec{x} \in \mathbb{R}^n \mid 0 < \vec{x}[i] \leq \lambda\}$$

if $\lambda > 0$ and to the set

$$\{\vec{x} \in \mathbb{R}^n \mid \lambda \leq \vec{x}[i] < 0\}$$

if $\lambda < 0$, we have that the set $\mathrm{Cone}(S, i, \lambda, \vec{0})$ is itself a finite Boolean combination of linear constraints with rational coefficients. Hence, it is definable in the theory $\langle \mathbb{R}, +, <, 1\rangle$. ∎

Lemma 6.22 can be generalized for handling vectors $\vec{v} \in \mathbb{Q}$, provided that the value of their ith components are not equal to λ.

Lemma 6.23 *Let $n \in \mathbb{N}_{>1}$ be a dimension, let $S \subseteq [0,1]^{n-1}$ be a set that is definable in $\langle \mathbb{R}, +, <, 1\rangle$, let $i \in \{1, 2, \ldots, n\}$, let $\vec{v} \in \mathbb{Q}^n$, and let $\lambda \in \mathbb{Q}_{\neq \vec{v}[i]}$. The set $\mathrm{Cone}(S, i, \lambda, \vec{v})$ is definable in $\langle \mathbb{R}, +, <, 1\rangle$ as well.*

Proof Since S is definable in $\langle \mathbb{R}, +, <, 1\rangle$, the set $S - \vec{v}$ is definable in $\langle \mathbb{R}, +, <, 1\rangle$ as well. We have

$$\mathrm{Cone}(S, i, \lambda, \vec{v}) - \vec{v} = \mathrm{Cone}(S - \vec{v}, i, \lambda - \vec{v}[i], 0),$$

which is definable in $\langle \mathbb{R}, +, <, 1\rangle$ by Lemma 6.22. It follows that the set $\mathrm{Cone}(S, i, \lambda, \vec{v})$ is definable in $\langle \mathbb{R}, +, <, 1\rangle$. ∎

Thanks to Lemmas 6.16, 6.19 and 6.23, we are now able to prove the main result of this section, i.e., that a set S satisfying the hypotheses of Lemmas 6.16 and 6.19 is necessarily definable in $\langle \mathbb{R}, +, <, 1\rangle$. In other words, we can restate Claim 6.10 as a theorem.

Theorem 6.24 *Let $n \in \mathbb{N}_{>1}$ be a dimension, and $r, s \in \mathbb{N}_{>1}$ be bases that do not share the same set of prime factors (resp. multiplicatively independent bases). Let $S \subseteq [0,1]^n$ be a set that is both (resp. weakly) r- and (resp. weakly) s-recognizable. Suppose that there exist $\ell \in \mathbb{N}_{>0}$ and $\vec{v} \in [0,1]^n \cap \mathbb{Q}^n$ such that S is r^ℓ-product-stable with respect to the pivot $\vec{v}$. Suppose that,* *for all* $m \in \{1, 2, \ldots, n-1\}$*, and for each set* $S' \subseteq [0,1]^m$ *that is both (resp. weakly) r- and (resp. weakly) s-recognizable,* S' *is definable in* $\langle \mathbb{R}, +, <, 1 \rangle$*. The set* S *is a bounded conical set definable in* $\langle \mathbb{R}, +, <, 1 \rangle$*.*

Proof Given a set $S' \subseteq \mathbb{R}^n$, an index $i \in \{1, 2, \ldots, n\}$ and a value $\lambda \in \mathbb{R}$, we denote by $\mathrm{Section}(S', i, \lambda)$ the set

$$\{\vec{x} \in \mathbb{R}^{n-1} \mid (\vec{x}[1], \ldots, \vec{x}[i-1], \lambda, \vec{x}[i], \ldots, \vec{x}[n-1]) \in S'\}.$$

Let S' be the set

$$S' = \bigcup_{i \in \{1,2,\ldots,n\}} \bigcup_{\lambda \in \{0,1\}} \mathrm{Cone}(\mathrm{Section}((S \cap F_{\lambda,i}), i, \lambda), i, \lambda, \vec{v}) \cup \begin{cases} \{\vec{v}\} & \text{if } \vec{v} \in S; \\ \emptyset & \text{otherwise.} \end{cases} \tag{6.5}$$

If follows from Lemma 6.16 that $S' \cap \mathbb{Q}^n = S \cap \mathbb{Q}^n$. It then follows from Theorem 2.69 that

$$S' = S.$$

Hence, it remains to establish that S' is definable in $\langle \mathbb{R}, +, <, 1 \rangle$.

By Lemma 6.19, the set $S \cap F_{\lambda,i}$ is definable in $\langle \mathbb{R}, +, <, 1 \rangle$ for all $\lambda \in \{0, 1\}$ and $i \in \{1, 2, \ldots, n\}$. It follows that the set $\mathrm{Section}((S \cap F_{\lambda,i}), i, \lambda)$ is definable in $\langle \mathbb{R}, +, <, 1 \rangle$ as well.

Let $i \in \{1, 2, \ldots, n\}$ and $\lambda \in \{0, 1\}$.

If $\vec{v}[i] \neq \lambda$, then the set $\mathrm{Cone}(\mathrm{Section}((S \cap F_{\lambda,i}), i, \lambda), i, \lambda, \vec{v})$ is definable in $\langle \mathbb{R}, +, <, 1 \rangle$ by Lemma 6.23. Otherwise, it follows from Definition 6.20 and from Lemma 6.16 that $\mathrm{Cone}(\mathrm{Section}((S \cap F_{\lambda,i}), i, \lambda), i, \lambda, \vec{v}) = S \cap F_{\lambda,i}$, which is definable in $\langle \mathbb{R}, +, <, 1 \rangle$.

As a consequence, each element of the union (6.5) is definable in the theory $\langle \mathbb{R}, +, <, 1 \rangle$. The same result holds for the set $\{\vec{v}\}$ since $\vec{v} \in \mathbb{Q}^n$. The set S' is thus a finite union of sets definable in $\langle \mathbb{R}, +, <, 1 \rangle$. Hence, it is definable in that theory. Moreover, since S' is a union of bounded conical sets (of apexes $\vec{v}$), it is itself a bounded conical set (of apex $\vec{v}$). ■

Taking into account the reduction provided in Theorem 6.11, we know that the r^ℓ-product-stability hypothesis of Theorem 6.24 is not needed. Hence, we have the following immediate corollary, stating that Claim 6.3 is true.

Corollary 6.25 *Let $n \in \mathbb{N}_{>1}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors (resp. multiplicatively independent bases), and $S \subseteq [0,1]^n$ be a set that is both (resp. weakly) r- and (resp. weakly) s-recognizable. If, for all $m \in \{1, 2, \ldots, n-1\}$ and for each set $S' \subseteq [0,1]^m$ that is both (resp. weakly) r- and (resp. weakly) s-recognizable, S' is definable in $\langle \mathbb{R}, +, <, 1\rangle$, then S is definable in $\langle \mathbb{R}, +, <, 1\rangle$.*

6.6 Generalizations of Semenov's theorem to RVA

In this section, we combine the results of Sections 6.3 and 6.5 to establish generalizations of Semenov's theorem to weak deterministic RVA and Muller RVA.

In the following theorem, restating Claim 6.2, we consider subsets of $[0,1]^n$ that are simultaneously (resp. weakly) r- and s-recognizable, where r and s are bases with different sets of prime factors (resp. multiplicatively independent bases).

Theorem 6.26 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors (resp. multiplicatively independent bases), and $S \subseteq [0,1]^n$. If S is simultaneously (resp. weakly) r- and (resp. weakly) s-recognizable, then S is definable in $\langle \mathbb{R}, +, <, 1\rangle$.*

Proof The proof is by induction on the dimension n.

By Theorems 5.19 and 5.34, the base case $n = 1$ holds. Moreover, the inductive case holds by Corollary 6.25. ∎

Thanks to the reduction developed in Chapter 4, we can now prove Claim 6.1, that considers subsets of $\mathbb{R}^n$ instead of subsets of $[0,1]^n$.

Theorem 6.27 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r, s \in \mathbb{N}_{>1}$ be bases with different sets of prime factors (resp. multiplicatively independent bases), and $S \subseteq \mathbb{R}^n$. If S is simultaneously (resp. weakly) r- and (resp. weakly) s-recognizable, then S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, <\rangle$.*

Proof This is an immediate corollary of Theorems 4.15 (resp. 4.16) and 6.26. ∎

6.7 Summary

The main theorem of this chapter, i.e, Theorem 6.27, leads to a complete characterization of the multi-dimensional sets simultaneously recognizable by Muller RVA (or by weak deterministic RVA) in multiple bases. In this section, we give the details of this characterization, extending the results of Section 5.8.

Let $n \in \mathbb{N}_{>0}$ be a dimension, let $r, s \in \mathbb{N}_{>1}$ be two bases, and let $S \subseteq \mathbb{R}^n$.

- If r and s are multiplicatively dependent, then, by Corollary 3.33, S is weakly r-recognizable if and only if it is weakly s-recognizable.
- If r and s are multiplicatively independent, and if S is simultaneously weakly r- and weakly s-recognizable, then, by Theorem 6.27, S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

When moving from weak recognizability to (full) recognizability, we have the following results :

- If r and s are multiplicatively dependent, then, by Corollary 2.68, S is r-recognizable if and only if it is s-recognizable.
- If r and s are multiplicatively independent, but share the same set of prime factors, then, by Theorem 5.21, there exists a set $S \subseteq \mathbb{R}$ that is both r- and s-recognizable, and that is not definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.
- If r and s are multiplicatively independent and have different sets of prime factors, then by Theorem 6.27, S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

Figure 6.11 schematically synthesizes the approach we followed to prove these results.

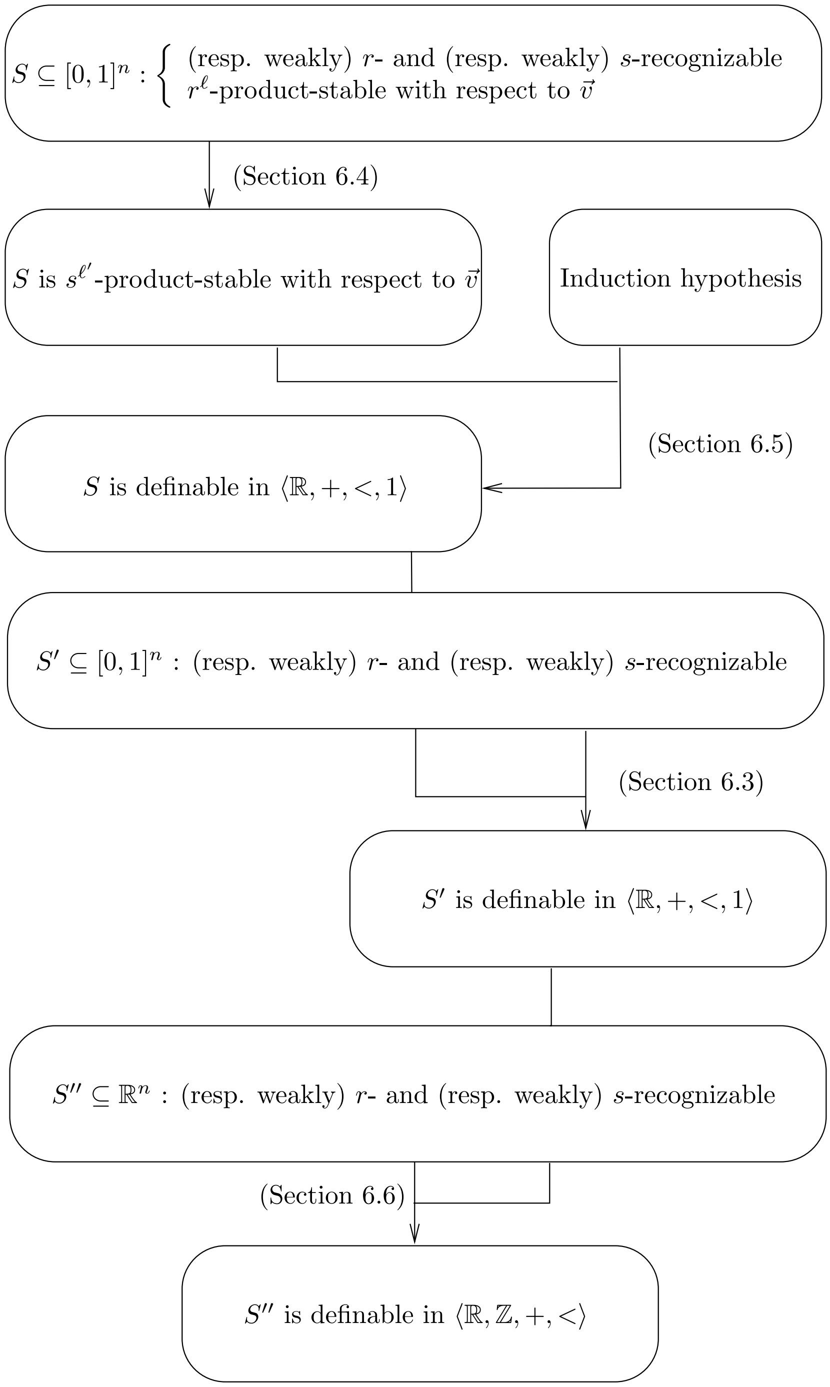

Figure 6.11: Summary of the results of Chapter 6.

Chapter 7

Internal structure of RVA

In Chapters 5 and 6, we have established that, for any $n \in \mathbb{N}_{>0}$, the subsets of $\mathbb{R}^n$ that are either recognizable, or weakly recognizable, in two sufficiently different bases, are exactly the sets definable in the theory $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. Moreover, it turned out that the recognizability, or weak recognizability, of a subset of $[0,1]^n$ in two sufficiently different bases implies its definability in the theory $\langle \mathbb{R}, +, <, 1 \rangle$. Taking into account the results of Chapter 4, we get the following characterization of the theory $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. This characterization was also obtained in [BFL08].

Theorem 7.1 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $S \subseteq \mathbb{R}^n$. The set S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$ if and only if it can be decomposed into a finite union*

$$\bigcup_{i=1}^{m} (S_i^I + S_i^F),$$

where $m \in \mathbb{N}$ and, for each $i \in \{1, 2, \ldots, m\}$,

- *the set $S_i^I \subseteq \mathbb{Z}^n$ is definable in $\langle \mathbb{Z}, +, < \rangle$;*
- *the set $S_i^F \subseteq [0,1]^n$ is definable in $\langle \mathbb{R}, +, <, 1 \rangle$.*

Proof If S is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$, then it is weakly r-recognizable for each base $r \in \mathbb{N}_{>1}$, by Theorem 3.31. In particular, there exist multiplicatively independent bases $r_1, r_2 \in \mathbb{N}_{>1}$ such that S is both weakly r_1- and weakly r_2-recognizable. It follows from Theorem 4.13 that S is equal to a finite union

$$S = \bigcup_{i=1}^{m} (S_i^I + S_i^F),$$

where $m \in \mathbb{N}$ and, for each $i \in \{1, 2, \ldots, m\}$, the set $S_i^I \subseteq \mathbb{Z}^n$ is both r_1- and r_2-recognizable (by NDDs), and the set $S_i^F \subseteq [0, 1]^n$ is both weakly r_1- and weakly r_2-recognizable. By Theorem 2.72, the sets S_i^I are definable in $\langle \mathbb{Z}, +, < \rangle$. In the same way, as a consequence of Theorem 6.26, the sets S_i^F are definable in $\langle \mathbb{R}, +, <, 1 \rangle$.

The other direction of the proof is immediate : Since each set S_i^I (resp. S_i^F) is definable in $\langle \mathbb{Z}, +, < \rangle$ (resp. $\langle \mathbb{R}, +, <, 1 \rangle$), the set $(S_i^I + S_i^F)$ is definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$, and so is the finite union $\bigcup_{i=1}^{m} (S_i^I + S_i^F)$. ■

In this chapter, using the results obtained in the previous chapters, we characterize the internal structure of a large class of weak deterministic RVA recognizing sets definable in $\langle \mathbb{R}, +, <, 1 \rangle$. Similar results are already known for NDDs recognizing sets definable in $\langle \mathbb{Z}, +, < \rangle$ [Ler05, Lat05].

Thanks to Theorem 7.1 as well as the discussions of Chapter 4, this characterization leads then to a full documentation of the structure of most of the weak deterministic RVA recognizing sets definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

We will thus characterize the internal structure of automata recognizing sets $S \subseteq \mathbb{R}^n$ definable in $\langle \mathbb{R}, +, <, 1 \rangle$. However, we will limit ourselves to a "good" class of sets, in a sense that will be defined later.

The motivation behind this work is the following : Our aim is to be able to benefit from the advantages of RVA, which mainly reside in their easy algorithmic manipulation and their canonicity, while managing to avoid some of their drawbacks. Precisely, we would like to deeply understand the structure of automata to identify parts of the transition relation of RVA that could be represented implicitly rather than by large sets of states.

7.1 Polyhedra

In this chapter, as mentioned, we focus on the sets $S \subseteq \mathbb{R}^n$ definable in $\langle \mathbb{R}, +, <, 1 \rangle$. As noticed in Section 2.2.2, the theory $\langle \mathbb{R}, +, <, 1 \rangle$ admits the elimination of quantifiers. As a consequence, a set $S \subseteq \mathbb{R}^n$ is definable in $\langle \mathbb{R}, +, <, 1 \rangle$ if and only if it can be expressed as a finite Boolean combination of linear constraints with rational coefficients. Such a set is called a *polyhedron.*

Definition 7.2 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $S \subseteq \mathbb{R}^n$. The set S is a* polyhedron *if and only if there exist $m \in \mathbb{N}$, $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_m} \in \mathbb{Z}^n$ and*

$b_1, b_2, \ldots, b_m \in \mathbb{Z}$ *such that* S *is a finite Boolean combination of the sets*

$$\{\vec{x} \in \mathbb{R}^n \mid \vec{a_i}\vec{x} \leq b_i\},$$

where i *ranges over the set* $\{1, 2, \ldots, m\}$. □

Thanks to existing procedures for eliminate the quantifiers of the formulas defining sets of $\langle \mathbb{R}, +, <, 1 \rangle$, we have the following immediate result :

Theorem 7.3 ([FR75]) *Let* $n \in \mathbb{N}_{>0}$ *be a dimension, and* $S \subseteq \mathbb{R}^n$. *The set* S *is a polyhedron if and only if it is definable in* $\langle \mathbb{R}, +, <, 1 \rangle$.

Remark that polyhedra are in general neither closed nor convex. Actually, the class of polyhedra that we have just defined coincides with the class of *Nef Polyhedra* introduced in [Nef78, Bie95, GV04], which is defined as *the subsets of* $\mathbb{R}^n$ *that can be obtained by a finite sequence of complement and intersection set operations over linear half-spaces* [GV04]. The remaining of this section is adapted from [Nef78, Bie95, GHH$^+$03, GV04] and introduces definitions and basic properties of polyhedra.

Example 7.4 *Figure 7.1 represents a* 2*-dimensional polyhedron* S. *It contains three* vertices *(*$A \in S$ *and* $B, C \notin S$*), three* edges *(*$a, c \subseteq S$ *and* $b \subseteq (\mathbb{R}^2 \setminus S)$*), the* interior $I \subseteq S$ *and the* exterior $E \subseteq (\mathbb{R}^2 \setminus S)$. ◇

The following definition generalizes Definition 6.17 : It now considers subsets of $\mathbb{R}^n$ instead of subsets of $[0, 1]^n$.

Definition 7.5 *Let* $n \in \mathbb{N}_{>0}$ *be a dimension. A set* $S \subseteq \mathbb{R}^n$ *is a* conical set of apex $\vec{v} \in \mathbb{R}^n$ *if and only if for all* $\vec{x} \in \mathbb{R}^n$, $f \in \mathbb{R}_{>0}$, *we have*

$$\vec{x} \in S \Leftrightarrow \vec{v} + f(\vec{x} - \vec{v}) \in S.$$

□

The apexes of conical sets form affine spaces, as proved in the following theorem.

Theorem 7.6 *Let* $n \in \mathbb{N}_{>0}$ *be a dimension, and* $S \subseteq \mathbb{R}^n$ *be a conical set. There exist* $m \in \mathbb{N}$, $\vec{v}, \vec{u_1}, \vec{u_2}, \ldots, \vec{u_m} \in \mathbb{R}^n$ *such that the set of apexes of* S *is the affine space*

$$\{\vec{v} + \lambda_1\vec{u_1} + \lambda_2\vec{u_2} + \ldots + \lambda_m\vec{u_m} \mid \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R}\}.$$

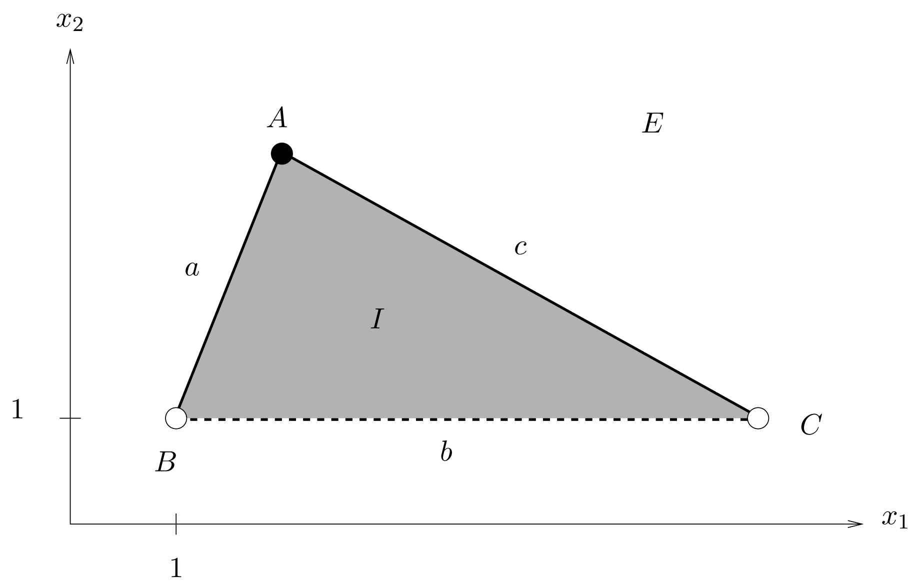

Figure 7.1: Example of a polyhedron in $\mathbb{R}^2$.

Proof It is sufficient to prove that, if $\vec{v_1}, \vec{v_2} \in \mathbb{R}^n$ are two distinct apexes of S, then each point on the line $L = \{\mu\vec{v_1} + (1-\mu)\vec{v_2} \mid \mu \in \mathbb{R}\}$, containing $\vec{v_1}$ and $\vec{v_2}$, is also an apex of S. This can be achieved by showing that S is invariant under any translation parallel to L, i.e., that

$$\vec{x} \in S \Leftrightarrow \vec{x} + \mu(\vec{v_1} - \vec{v_2}) \in S$$

for all $\vec{x} \in \mathbb{R}^n$ and $\mu \in \mathbb{R}$. Without loss of generality, we can restrict ourselves to the values $\mu \in \mathbb{R}_{\geq 0}$, since the apexes $\vec{v_1}$ and $\vec{v_2}$ can be exchanged if needed.

Let $\vec{x} \in \mathbb{R}^n$ be an arbitrary vector, and consider an arbitrary value $\mu \in \mathbb{R}_{\geq 0}$. We define

$$\vec{x'} = \vec{x} + \mu(\vec{v_1} - \vec{v_2})$$

and

$$f = \frac{1}{1+\mu}.$$

Since S is a conical set of apex $\vec{v_1}$, we have

$$\vec{x} \in S \Leftrightarrow \vec{v_1} + f(\vec{x} - \vec{v_1}) \in S. \tag{7.1}$$

Exploiting the conical structure of S with respect to the apex $\vec{v_2}$, we also get

$$\vec{x'} \in S \Leftrightarrow \vec{v_2} + f(\vec{x'} - \vec{v_2}) \in S. \tag{7.2}$$

By replacing $\vec{x'}$ by $\vec{x} + \mu(\vec{v_1} - \vec{v_2})$ and μ by $\dfrac{1-f}{f}$, we deduce

$$\begin{aligned}\vec{v_2} + f(\vec{x'} - \vec{v_2}) &= f\left(\vec{x} + \frac{1-f}{f}(\vec{v_1} - \vec{v_2}) - \vec{v_2}\right) + \vec{v_2} \\ &= \vec{v_1} + f(\vec{x} - \vec{v_1}).\end{aligned}$$

Combining Equations 7.1 and 7.2 finally yields $\vec{x} \in S \Leftrightarrow \vec{x'} \in S$. ■

Pyramids are defined as the sets $S \subseteq \mathbb{R}^n$ that satisfy both Definitions 7.2 and 7.5.

Definition 7.7 *Let $n \in \mathbb{N}_{>0}$ be a dimension. A set $S \subseteq \mathbb{R}^n$ is a* pyramid *if and only if S is a conical polyhedron.* □

Thanks to Theorem 7.6, we know that each pyramid is conical with respect to an affine space of apexes. The *dimension* of a pyramid is defined as the dimension of this affine space.

Definition 7.8 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $S \subseteq \mathbb{R}^n$ be a pyramid. Let $S' \subseteq \mathbb{R}^n$ be the set of vectors $\vec{x} \in \mathbb{R}^n$ such that S is a conical set of apex $\vec{x}$. The* dimension of S, *denoted* $\dim(S)$, *is defined by* $\dim(S')$. □

Example 7.9 *The polyhedron depicted in Figure 7.1 is not a pyramid because it is not a conical set.*

However, the polyhedron of Figure 7.2 is a pyramid. The sets of points $\vec{x} \in \mathbb{R}^n$ such that this polyhedron is conical with respect to $\vec{x}$ only contains the point $(1, 1)$. Hence, its dimension is 0. ◇

When dealing with sets $S \subseteq [0,1]^n$, we need the concept of *bounded pyramid*, whose definition uses the bounded conical sets, introduced in Definition 6.17.

Definition 7.10 *Let $n \in \mathbb{N}_{>0}$ be a dimension. A set $S \subseteq [0,1]^n$ is a* bounded pyramid *if and only if S is simultaneously a polyhedron and a bounded conical set.* □

For bounded pyramids, we can define an adapted notion of dimension.

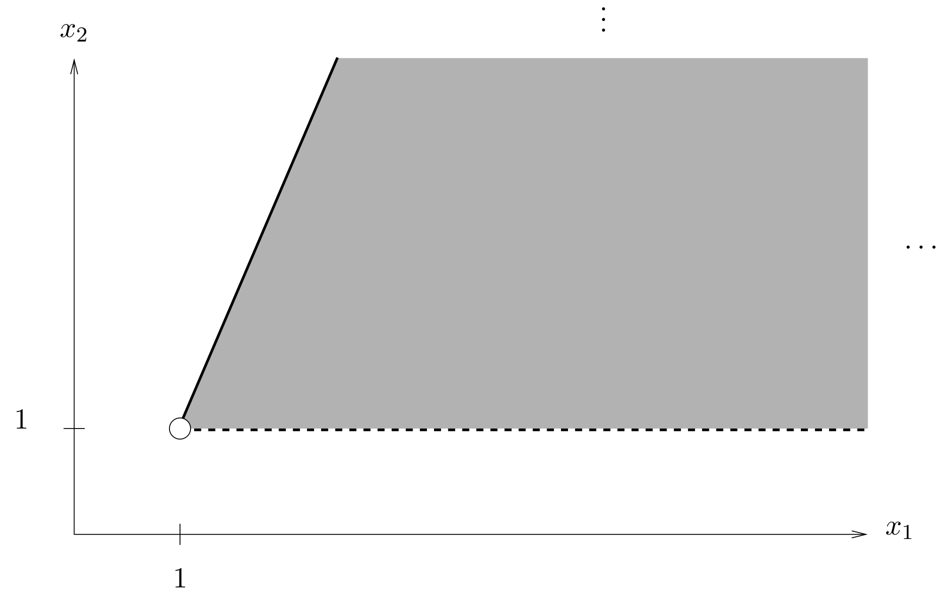

Figure 7.2: Example of a pyramid in $\mathbb{R}^2$.

Definition 7.11 *Let $n \in \mathbb{N}_{>0}$ be a dimension, and $S \subseteq \mathbb{R}^n$ be a bounded pyramid. Let $S' \subseteq \mathbb{R}^n$ be the set of vectors $\vec{x} \in \mathbb{R}^n$ such that S is a bounded conical set of apex $\vec{x}$. The* dimension of S*, denoted* $\dim(S)$*, is defined by* $\dim(\text{aff}(S'))$. □

A property of polyhedra is the following : Given a polyhedron S and a point $\vec{x} \in \mathbb{R}^n$, there exists a neighborhood $N_\varepsilon(\vec{x})$, with $\varepsilon \in \mathbb{R}_{>0}$ and with respect to Euclidean distance, such that, for all neighborhoods $N_{\varepsilon'}(\vec{x})$ included in $N_\varepsilon(x)$, the sets $S \cap N_\varepsilon(x)$ and $S \cap N_{\varepsilon'}(x)$ are equal up to some zoom factor. In the following definition, we use this property to introduce the notion of *local adjoined pyramid*.

Definition 7.12 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $S \subseteq \mathbb{R}^n$ be a polyhedron, and $\vec{x} \in \mathbb{R}^n$. Let $\varepsilon \in \mathbb{R}_{>0}$ be such that, for all $\varepsilon' \in]0, \varepsilon]$, the pyramids*

$$P_S(\vec{x}) = \vec{x} + \{\lambda((S \cap N_{\varepsilon'}(\vec{x})) - \vec{x}) \mid \lambda \in \mathbb{R}_{>0}\}$$

are identical. The set $P_S(\vec{x})$ is called the local adjoined pyramid *of S in $\vec{x}$.* □

Example 7.13 *The possible local adjoined pyramids of the polyhedron of Figure 7.1 are represented in Figure 7.3.* ◇

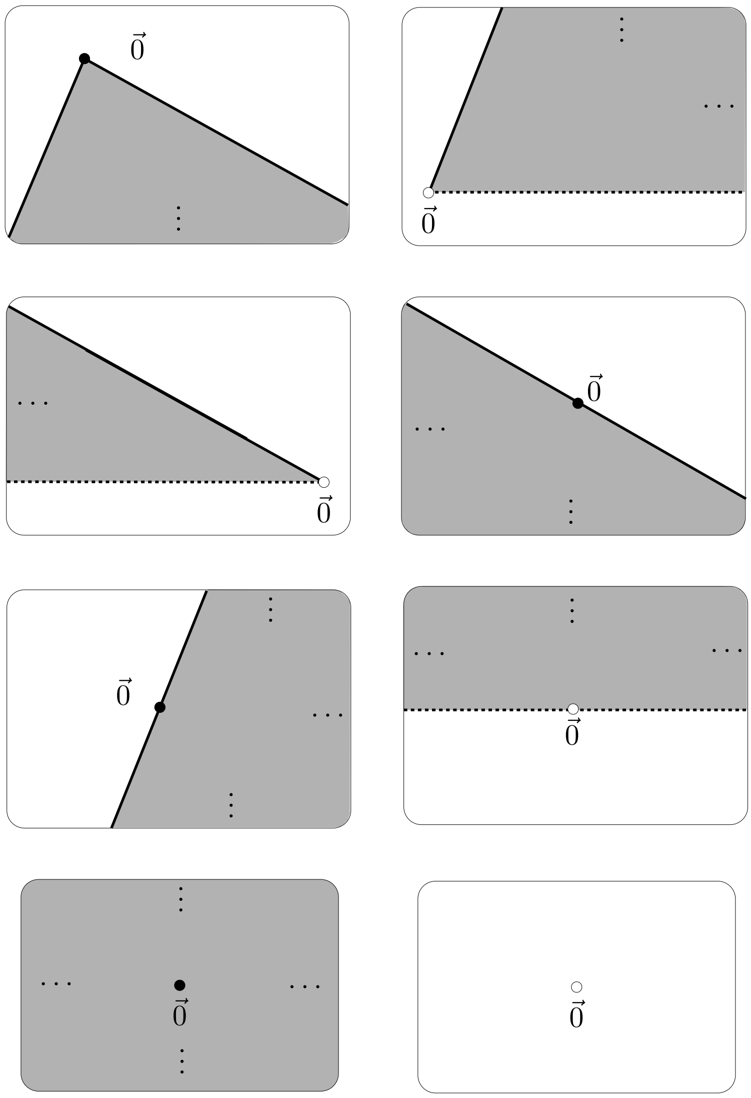

Figure 7.3: Local adjoined pyramids of the polyhedron of Figure 7.1.

When fixing a polyhedron S, the points $\vec{x} \in \mathbb{R}^n$ can be classified into *components* of S : Two points $\vec{x}, \vec{y}$ belong to the same component if and only if their local adjoined pyramids are equal.

Definition 7.14 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $S \subseteq \mathbb{R}^n$ be a polyhedron, and $\vec{x}, \vec{y} \in \mathbb{R}^n$. We have $\vec{x} \sim_S \vec{y}$ if and only if $P_S(\vec{x}) = P_S(\vec{y})$. The equivalence classes of the relation $\sim_S$ define the* components *of S.* □

Example 7.15 *In the polyhedron represented in Figure 7.1, there are eight components : The three vertices A, B, C, the three edges a, b, c, the interior I and the exterior E.* ◇

The components of a polyhedron S satisfy some properties [Nef78, Bie95, GV04]. In particular, it has been proved that they are finite in number. The *dimension* of a component is defined in the following way.

Definition 7.16 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $S \subseteq \mathbb{R}^n$ be a polyhedron, and $C \subseteq \mathbb{R}^n$ be a component of S. If $\vec{x}$ is an element of C, then the* dimension *of C is the dimension of the affine space forming the set of points $\vec{x'} \in \mathbb{R}^n$ such that $P_S(\vec{x})$ is conical with respect to $\vec{x'}$.* □

Given a polyhedron, its components are linked by an *incidence relation*, as formalized in the next definition.

Definition 7.17 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $S \subseteq \mathbb{R}^n$ be a polyhedron, and $C_1, C_2 \subseteq \mathbb{R}^n$ be distinct components of S. The component C_1 is* incident to *the component C_2, which is denoted $C_1 \prec C_2$, if and only if C_1 is a subset of the closure of C_2, i.e., a subset of the union of C_2 and the set of its limit points.* □

Example 7.18 *In the polyhedron of Figure 7.1, we have $A \prec c \prec I$ and $A \not\prec C$.* ◇

To conclude this section, the following lemma gives a link between the components of a polyhedron and the set of linear inequations from which it is defined.

Lemma 7.19 *Let* $n \in \mathbb{N}_{>0}$ *be a dimension, and* $S \subseteq \mathbb{R}^n$ *be a polyhedron. Suppose that there exist* $m \in \mathbb{N}$, $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_m} \in \mathbb{Z}^n$ *and* $b_1, b_2, \ldots, b_m \in \mathbb{Z}$ *such that* S *is a finite Boolean combination of the sets* $\{\vec{x} \in \mathbb{R}^n \mid \vec{a_i}\vec{x} \leq b_i\}$, *where* i *ranges over the set* $\{1, 2, \ldots, m\}$. *Let* $\vec{x} \in \mathbb{R}^n$ *be an arbitrary vector,* $S' = \vec{x} + P_S(\vec{x})$, *and* $S'' \subseteq \mathbb{R}^n$ *be the affine space of vectors* $\vec{x}'$ *such that* S' *is conical with respect to* $\vec{x}'$. *There exists* $I \subseteq \{1, 2, \ldots, m\}$ *such that*

$$S'' \subseteq \bigcap_{i \in I} \{\vec{x} \in \mathbb{R}^n \mid \vec{a_i}\vec{x} = b_i\}$$

and $\dim(S') = \dim(S'') = \dim(\bigcap_{i \in I} \{\vec{x} \in \mathbb{R}^n \mid \vec{a_i}\vec{x} = b_i\})$.

Proof If $\dim(S') = n$, then $S' = S'' = \mathbb{R}^n$, which corresponds to the intersection of zero sets.

Otherwise, a non-empty subset $I \subseteq \{1, 2, \ldots, m\}$ satisfying the expected conditions necessarily exists since S is a finite Boolean combination of the sets $\{\vec{x} \in \mathbb{R}^n \mid \vec{a_i}\vec{x} \leq b_i\}$. ■

7.2 Suitable polyhedra

In Chapter 2, and in particular in Theorem 2.48, we mentioned that, given a base $r \in \mathbb{N}_{>1}$, there sometimes exist distinct r-encodings, but sharing the same integer part and encoding the same real number; such encodings are called *dual encodings*. Dual encodings are useful in theory since, when dealing with languages containing *all* the encodings of the numbers they represent, operations on the represented sets can generally directly be implemented by the corresponding well-known algorithms on infinite-word automata, as observed in Section 2.5.3.

However, as a drawback behind the use of dual encodings, the structures of automata are sometimes more complex than necessary.

Example 7.20 *A weak deterministic RVA recognizing the set* $\{\vec{x} \in \mathbb{R}^2 \mid \vec{x}[1] - \vec{x}[2] = 0\}$ *in base* 2 *is represented in Figure 7.4. In this automaton, all the encodings labeled by paths ending cycling in States* 6 *and* 7 *have dual encodings ending cycling in State* 5. ◇

In this section, we define a class of *r-suitable polyhedra*, with respect to a base $r \in \mathbb{N}_{>1}$. The RVA representing those polyhedra in base r have in general a more simple transition graph than general polyhedra; for instance,

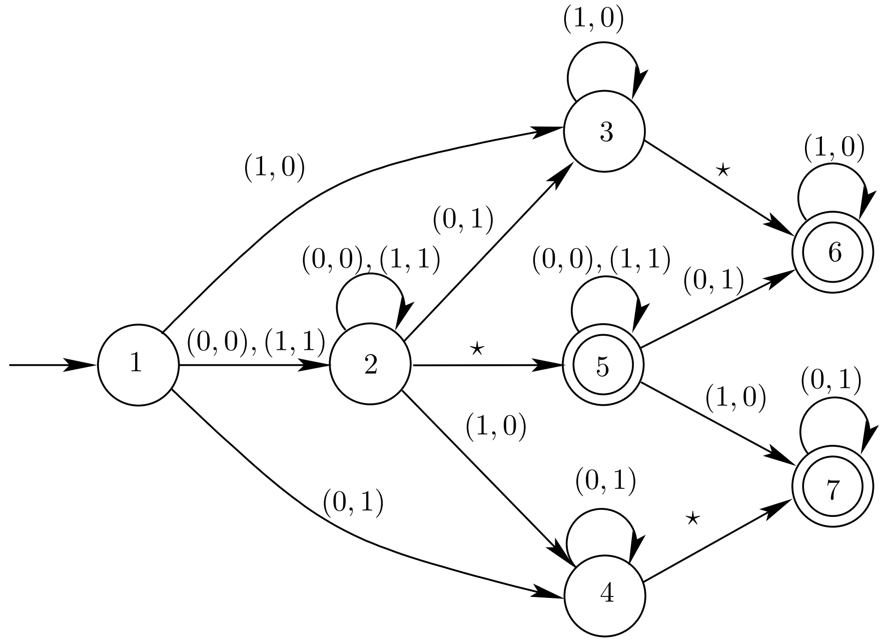

Figure 7.4: RVA representing the set $\{\vec{x} \in \mathbb{R}^2 \mid \vec{x}[1] - \vec{x}[2] = 0\}$.

we will see that the polyhedron introduced in Example 7.20 is not 2-suitable. The internal structure of RVA recognizing r-suitable polyhedra will then be documented in the next section.

Definition 7.21 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq \mathbb{R}^n$ be a polyhedron. The polyhedron S is* r-suitable *if and only if*

- *there exist $m \in \mathbb{N}$, $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_m} \in \mathbb{Z}^n$ and $b_1, b_2, \ldots, b_m \in \mathbb{Z}$;*
- *S is a finite Boolean combination of the sets $\{\vec{x} \in \mathbb{R}^n \mid \vec{a_i}\vec{x} \leq b_i\}$, where i ranges over the set $\{1, 2, \ldots, m\}$;*
- *for each subset $I \subseteq \{1, 2, \ldots, m\}$, let S_I be the set*

$$S_I = \bigcap_{i \in I} \{\vec{x} \in \mathbb{R}^n \mid \vec{a_i}\vec{x} = b_i\},$$

and let d_I be the dimension of the affine space S_I, i.e., $d_I = \dim(S_I)$. For all $I \subseteq \{1, 2, \ldots, m\}$, $\vec{\imath} \in \mathbb{Z}^n$ and $k \in \mathbb{Z}$, if

$$S_{I,\vec{\imath},k} = S_I \cap \left(\frac{\vec{\imath} + [0,1]^n}{r^k} \right) \neq \emptyset,$$

then $\dim(\mathrm{aff}(S_{I,\vec{\imath},k})) = d_I$.

□

Intuitively speaking, a polyhedron is r-suitable if, for all weak deterministic RVA r-recognizing the sets S_I introduced in Definition 7.21, the states q of their fractional parts are such that the dimensions of the affine hulls of the sets $S(q)$ are equal to the dimension of the affine space S_I.

For instance, we give in the following example a polyhedron that is not 2-suitable.

Example 7.22 *Let S be the two-dimensional polyhedron*

$$S = \{\vec{x} \in \mathbb{R}^2 \mid \vec{x}[1] - \vec{x}[2] \leq 0\}.$$

It is itself a set of the form $\{\vec{x} \in \mathbb{R}^2 \mid \vec{a}\vec{x} \leq b\}$. Let $S' = \{\vec{x} \in \mathbb{R}^2 \mid \vec{x}[1] - \vec{x}[2] = 0\}$, and $S'' = S' \cap \left(\dfrac{(1,0) + [0,1]^n}{2}\right)$. We have $S'' = \left\{\left(\frac{1}{2}, \frac{1}{2}\right)\right\} \neq \emptyset$, but $\dim(\mathrm{aff}(S'')) = 0$ whereas $\dim(S') = 1$. Hence, S is not 2-suitable. Figure 7.4, introduced in Example 7.20, depicts the minimal weak deterministic RVA recognizing S' in base 2. Let q be the state 7 of this automaton. We have $S(q) = \{(0,1)\}$, and $\dim(\mathrm{aff}(S(q))) = 0$. ◇

The following example provides a 2-suitable polyhedron.

Example 7.23 *Let S be the two dimensional polyhedron*

$$S = \{\vec{x} \in \mathbb{R}^2 \mid 3.\vec{x}[1] - 3.\vec{y}[1] = 1\}.$$

S is a 2-suitable polyhedron. The minimal weak deterministic RVA recognizing S is base 2 is depicted in Figure 7.5. ◇

7.3 Internal structure of RVA recognizing the r-suitable polyhedra

In this section, we focus on the r-suitable polyhedra, with $r \in \mathbb{N}_{>1}$, and we give insight into the structure of the fractional parts of weak deterministic RVA recognizing them. Additionally, we show that the polyhedra that are not r-suitable have in general a more complex structure.

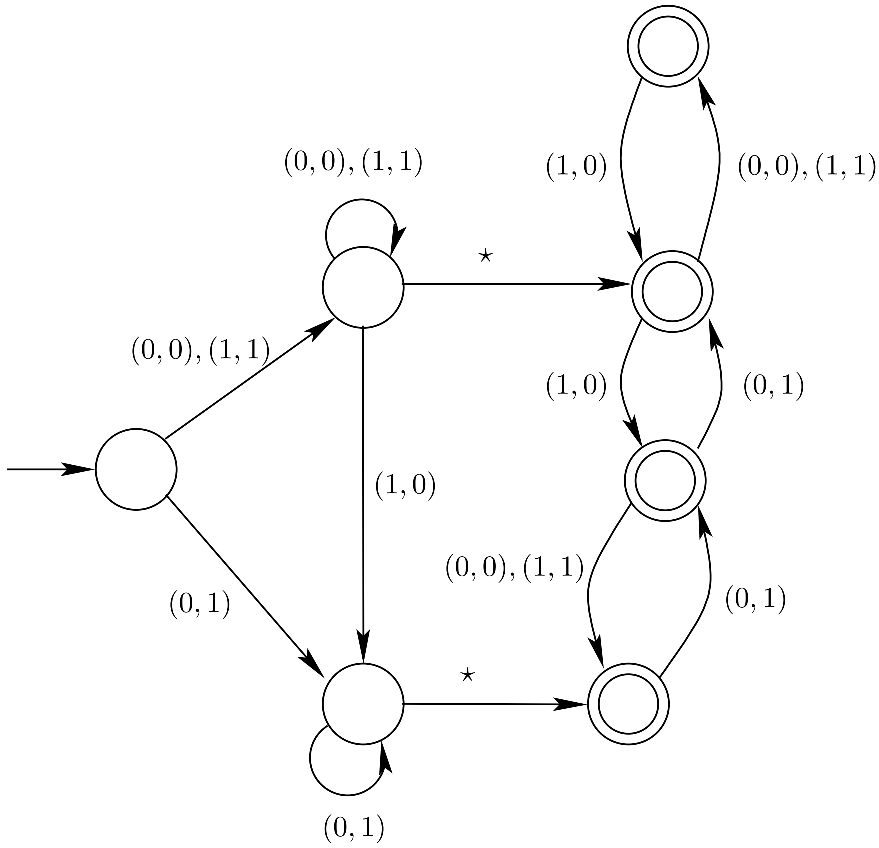

Figure 7.5: RVA representing the set $\{\vec{x} \in \mathbb{R}^2 \mid 3.\vec{x}[1] - 3.\vec{x}[2] = 1\}$.

In the following theorem, we consider the states q belonging to non-trivial strongly connected components of the fractional part of RVA recognizing polyhedra. We establish that each set $S(q)$ is a bounded pyramid. The theorem holds even for polyhedra that are not r-suitable, and uses the results of Chapter 6.

Theorem 7.24 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq \mathbb{R}^n$ be a polyhedron recognized by a weak deterministic RVA $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ in base r. Let $q \in Q$ be a state such that there exist $w_1 \in \{0, r-1\}^n(\{0, 1, \ldots, r-1\}^n)^*$ and $w_2 \in (\{0, 1, \ldots, r-1\}^n)^*$ such that $\delta^*(q_0, w_1 \star w_2)$ is defined and equal to q. If q belongs to a non-trivial strongly connected component of $\mathcal{A}$, then $S(q)$ is a bounded pyramid.*

Proof If $S(q) = \emptyset$, then the result is immediate.

Otherwise, it suffices to establish that $S(q)$ is simultaneously a polyhedron and a bounded conical set.

Since S is a polyhedron, it is definable in $\langle \mathbb{R}, +, <, 1 \rangle$. By Theorem 3.31, S is weakly s-recognizable, for every base $s \in \mathbb{N}_{>1}$. In particular, there exists a base $s \in \mathbb{N}_{>1}$ such that r and s are multiplicatively independent and S is weakly s-recognizable. Applying Lemma 6.5, we deduce that $S(q)$ is simultaneously weakly r- and weakly s-recognizable. It then follows from Theorem 6.26 that $S(q)$ is definable in $\langle \mathbb{R}, +, <, 1 \rangle$. Hence, it is a polyhedron.

Since q belongs to a non-trivial strongly connected component of $\mathcal{A}$, there exists $w \in (\{0, 1, \ldots, r-1\}^n)^+$ such that $\delta^*(q, w) = q$. Hence, by Lemma 6.8, $S(q)$ is $r^{|w|}$-product-stable with respect to the pivot $\langle \vec{0} \star w^\omega \rangle_{r,n}$. By Theorems 6.24 and 6.26, $S(q)$ is a bounded conical set. ∎

We now move to r-suitable polyhedra. Let $\mathcal{A}$ be a weak deterministic RVA recognizing such a polyhedron. In the next theorem, we consider states q, q' that belong to non-trivial strongly connected components of the fractional part of $\mathcal{A}$ and such that q' is reachable from q. By Theorem 7.24, the sets $S(q)$ and $S(q')$ are bounded pyramids. Here, we prove $\dim(S(q)) \leq \dim(S(q'))$.

Theorem 7.25 *Let $n \in \mathbb{N}_{>0}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, and $S \subseteq \mathbb{R}^n$ be a r-suitable polyhedron recognized by a weak deterministic RVA $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ in base r. Let $q, q' \in Q$ be states that belong to non-trivial strongly connected components of $\mathcal{A}$ and such that there exist $w_1 \in \{0, r-1\}^n(\{0, 1, \ldots, r-1\}^n)^*$ and $w_2, w_3 \in (\{0, 1, \ldots, r-1\}^n)^*$ such that $\delta^*(q_0, w_1 \star w_2) = q$ and $\delta^*(q, w_3) = q'$. We have $\dim(S(q)) \leq \dim(S(q'))$.*

Proof Let $m \in \mathbb{N}$, $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_m} \in \mathbb{Z}^n$ and $b_1, b_2, \ldots, b_m \in \mathbb{Z}$ provided by Definition 7.21 applied to S. These values exist since S is a r-suitable polyhedron.

By Theorem 7.24, $S(q)$ and $S(q')$ are both bounded pyramids.

By the deterministic nature of $\mathcal{A}$, we have the following equalities :

$$S \cap \left(\langle w_1 \star w_2 \vec{0}^\omega \rangle_{r,n} + \frac{[0,1]^n}{r^{|w_2|}} \right) = \langle w_1 \star w_2 \vec{0}^\omega \rangle_{r,n} + \frac{S(q)}{r^{|w_2|}}$$

and

$$S \cap \left(\langle w_1 \star w_2 w_3 \vec{0}^\omega \rangle_{r,n} + \frac{[0,1]^n}{r^{|w_2|+|w_3|}} \right) = \langle w_1 \star w_2 w_3 \vec{0}^\omega \rangle_{r,n} + \frac{S(q')}{r^{|w_2|+|w_3|}}.$$

Let $\vec{x} \in [0,1]^n$ such that $S(q')$ is a bounded conical set of apex $\vec{x}$, and let

$$\vec{x'} = \langle w_1 \star w_2 w_3 \vec{0}^\omega \rangle_{r,n} + \frac{\vec{x}}{r^{|w_2|+|w_3|}}.$$

We have

$$\vec{x'} \in \langle w_1 \star w_2 w_3 \vec{0}^\omega \rangle_{r,n} + \frac{[0,1]^n}{r^{|w_2|+|w_3|}} = \frac{\langle w_1 w_2 w_3 \rangle_{r,n} + [0,1]^n}{r^{|w_2|+|w_3|}}$$

and thus also

$$\vec{x'} \in \langle w_1 \star w_2 \vec{0}^\omega \rangle_{r,n} + \frac{[0,1]^n}{r^{|w_2|}} = \frac{\langle w_1 w_2 \rangle_{r,n} + [0,1]^n}{r^{|w_2|}}.$$

Let $S' = \vec{x'} + P_S(\vec{x'})$, and $S'' \subseteq \mathbb{R}^n$ be the affine space of vectors $\vec{x''}$ such that S' is a conical set of apex $\vec{x''}$. By Lemma 7.19, there exists $I \subseteq \{1, 2, \ldots, m\}$ such that

$$S'' \subseteq S_I$$

and $\dim(S') = \dim(S'') = \dim(S_I)$, where

$$S_i = \bigcap_{i \in I} \{\vec{x} \in \mathbb{R}^n \mid \vec{a_i}\vec{x} = b_i\}.$$

If $S(q)$ is a bounded conical set of apex $\langle \vec{0} \star w_3 \vec{0}^\omega \rangle_{r,n} + \frac{\vec{x}}{r^{|w_3|}}$, then, since $S' = \vec{x'} + P_S(\vec{x})$ and by definition of the local adjoined pyramid $P_S(\vec{x})$, we have $\vec{x'} \in S''$. Hence, by Definition 7.21, we get

$$\begin{aligned} & \dim\left(\operatorname{aff}\left(S_I \cap \frac{\langle w_1 w_2 \rangle_{r,n} + [0,1]^n}{r^{|w_2|}}\right)\right) \\ = \; & \dim\left(\operatorname{aff}\left(S_I \cap \frac{\langle w_1 w_2 w_3 \rangle_{r,n} + [0,1]^n}{r^{|w_2|+|w_3|}}\right)\right) \\ = \; & \dim(S_I). \end{aligned}$$

In this case, it follows that $\dim(S(q)) = \dim(S(q')) = \dim(S_I)$.

Otherwise, the set $S(q)$ is conical, but $\langle \vec{0} \star w_3 \vec{0}^\omega \rangle_{r,n} + \frac{\vec{x}}{r^{|w_3|}}$ does not belong to its set of apexes. Let $\vec{x''}$ be an apex of $S(q)$. The component of S to which $\vec{x'}$ belongs is necessarily incident to the component of $\langle w_1 \star w_2 \vec{0}^\omega \rangle_{r,n} + \frac{\vec{x}}{r^{|w_2|}}$. Hence, we have $\dim(P_S(\vec{x'})) > \dim \left(P_S \left(\langle w_1 \star w_2 \vec{0}^\omega \rangle_{r,n} + \frac{\vec{x}}{r^{|w_2|}} \right) \right)$. By Definition 7.21, we thus have

$$\begin{aligned} & \dim \left(\text{aff} \left(S_I \cap \frac{\langle w_1 w_2 \rangle_{r,n} + [0,1]^n}{r^{|w_2|}} \right) \right) \\ < \; & \dim \left(\text{aff} \left(S_I \cap \frac{\langle w_1 w_2 w_3 \rangle_{r,n} + [0,1]^n}{r^{|w_2|+|w_3|}} \right) \right) \\ = \; & \dim(S_I). \end{aligned}$$

■

Theorems 7.24 and 7.25 provide a relatively precise description of the structure of the fractional parts of weak deterministic RVA recognizing r-suitable polyhedra. States q belonging to non-trivial strongly connected components have the following property : $S(q)$ is a bounded pyramid. Moreover, in the graph of the RVA, if a state q' is reachable from a state q, and if q and q' are both in strongly connected components, then $\dim(S(q)) \leq \dim(S(q'))$. Additionally, the readings of the encodings of the apexes of $S(q)$ end in a strongly connected component whose states q'' are such that $\dim(S(q'')) = \dim(S(q))$.

When dealing with non-necessarily r-suitable polyhedra, this property generally does not hold and weak deterministic RVA recognizing such polyhedra can have a more complicated structure.

Example 7.26 *Consider the RVA introduced in Example 7.20 and depicted in Figure 7.4. This RVA recognizes the set* $S = \{\vec{x} \in \mathbb{R}^2 \mid \vec{x}[1] - \vec{x}[2] \leq 0\}$. *Let* q *and* q' *be respectively States* 5 *and* 6 *in Figure 7.4. We have* $\dim(S(q)) = 1$ *and* $\dim(S(q')) = 0$. ◇

7.4 Towards a characterization of the r-suitable polyhedra

The results of the previous section consist in a documentation of the internal structure of automata recognizing r-suitable polyhedra, which where defined in Section 7.2. However, this definition is quite heavy, and it could be useful to have a more algebraic characterization of those polyhedra. That is the subject of this section. Precisely, we will establish a characterization of the polyhedra having the property to be *r-suitable $(n-1)$-planes*, i.e., of the r-suitable sets of the form $\{\vec{x} \in \mathbb{R}^n \mid \vec{a}\vec{x} = b\}$, with $\vec{a} \in \mathbb{Z}^n$, $\vec{a} \neq \vec{0}$ and $b \in \mathbb{Z}$. Additionally, by a counter-example, we show that this characterization cannot be simply extended to general r-suitable polyhedra.

Let $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ be a weak deterministic RVA recognizing a r-suitable $(n-1)$-plane. For each state q of the fractional part of $\mathcal{A}$, the set $S(q)$ is a conical set since we have

$$S \cap \left(\langle w_1 \star w_2 \vec{0}^\omega \rangle_{r,n} + \frac{[0,1]^n}{r^{|w_2|}} \right) = \langle w_1 \star w_2 \vec{0}^\omega \rangle_{r,n} + \frac{S(q)}{r^{|w_2|}},$$

where $w_1 \star w_2 \in \{0, r-1\}^n (\{0,1,\ldots,r-1\}^n)^* \star (\{0,1,\ldots,n-1\}^n)^*$ is a word that leads to q from the initial state q_0 of $\mathcal{A}$, i.e., $\delta^*(q_0, w_1 \star w_2) = q$. Moreover, by definition of the r-suitable polyhedra, if $S(q) \neq \emptyset$, then $\dim(S(q)) = n - 1$.

Let $S = \{\vec{x} \in \mathbb{R}^n \mid \vec{a}\vec{x} = b\}$, with $\vec{a} \in \mathbb{Z}^n$ and $b \in \mathbb{Z}$. If $\vec{a}[j] = 0$ for some $j \in \{1, 2, \ldots, n\}$, then the possible values of $\vec{x}[j]$ are arbitrary. In the following lemma, we use this property to establish that, in this case, the set S is r-suitable if and only if the set

$$S' = \{\vec{x} \in \mathbb{R}^{n-1} \mid \vec{a'}\vec{x} = b\},$$

where $\vec{a'} = (\vec{a}[1], \vec{a}[2], \ldots, \vec{a}[j-1], \vec{a}[j+1], \ldots, \vec{a}[n])$, is r-suitable as well.

Lemma 7.27 *Let $n \in \mathbb{N}_{>1}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, $\vec{a} \in \mathbb{Z}^n$ and $b \in \mathbb{Z}$ such that $\vec{a}[j] = 0$ for a given $j \in \{1, 2, \ldots, n\}$. The set*

$$S = \{\vec{x} \in \mathbb{R}^n \mid \vec{a}\vec{x} = b\}$$

is r-suitable if and only if the set

$$S' = \{\vec{x} \in \mathbb{R}^{n-1} \mid \vec{a'}\vec{x} = b\},$$

where $\vec{a'} = (\vec{a}[1], \vec{a}[2], \ldots, \vec{a}[j-1], \vec{a}[j+1], \ldots, \vec{a}[n])$, is r-suitable.

Proof If $\vec{a} = 0$, then we either have $S = \mathbb{R}^n$ or $S = \emptyset$. In this case, the set S' is respectively equal to $\mathbb{R}^{n-1}$ or to $\emptyset$. The lemma thus trivially holds.

Otherwise, the set S is r-suitable if, for all $\vec{\imath} \in \mathbb{Z}^n$ and $k \in \mathbb{Z}$, $S \cap \left(\frac{\vec{\imath} + [0,1]^n}{r^k} \right) \neq \emptyset$ implies $\dim \left(S \cap \left(\frac{\vec{\imath} + [0,1]^n}{r^k} \right) \right) = n - 1$. It suffices to prove that this condition is equivalent to the fact that $S' \cap \left(\frac{\vec{i'} + [0,1]^n}{r^k} \right) \neq \emptyset$ implies $\dim \left(S' \cap \left(\frac{\vec{i'} + [0,1]^n}{r^k} \right) \right) = n - 2$, where

$$\vec{i'} = (\vec{\imath}[1], \vec{\imath}[2], \ldots, \vec{\imath}[j-1], \vec{\imath}[j+1], \ldots, \vec{\imath}[n]).$$

This is actually the case, since

$$\vec{x} \in S' \cap \left(\frac{\vec{i'} + [0,1]^n}{r^k} \right)$$
$$\Leftrightarrow$$
$$(\vec{x}[1], \vec{x}[2], \ldots, \vec{x}[j-1], \alpha, \vec{x}[j+1], \ldots, \vec{x}[n]) \in S \cap \left(\frac{\vec{\imath} + [0,1]^n}{r^k} \right),$$

for each $\alpha \in \mathbb{R}$. ■

Note that if $n = 1$, then a $(n-1)$-plane can only be a singleton $\{x\}$, with $x \in \mathbb{Q}$, which is necessarily r-suitable.

Thanks to Lemma 7.27, the problem of characterizing the r-suitable $(n-1)$-planes is reduced to the same problem for sets that can be expressed as $\{\vec{x} \in \mathbb{R}^n \mid \vec{a}\vec{x} = b\}$, with $\vec{a}[i] \neq 0$ for all $i \in \{1, 2, \ldots, n\}$.

We first establish that such $(n-1)$-planes, with $n \in \mathbb{N}_{>1}$, are not r-suitable if and only if there exist $\vec{v} \in \mathbb{Z}^n$ and $k \in \mathbb{N}$ such that $\vec{a}\left(\frac{\vec{v}}{r^k}\right) = b$. We establish the two parts of this condition separately.

Lemma 7.28 *Let $n \in \mathbb{N}_{>1}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, $\vec{a} \in (\mathbb{Z}_{\neq 0})^n$, $b \in \mathbb{Z}$ and $S = \{\vec{x} \in \mathbb{R}^n \mid \vec{a}\vec{x} = b\}$. If S is not r-suitable, then there exist $\vec{v} \in \mathbb{Z}^n$ and $k \in \mathbb{Z}$ such that $\vec{a}\left(\frac{\vec{v}}{r^k}\right) = b$.*

Proof Since S is not r-suitable, there exist $\vec{\imath} \in \mathbb{Z}^n$ and $k \in \mathbb{Z}$ such that $S \cap S' \neq \emptyset$ and $\dim(\text{aff}(S \cap S')) < n - 1$, where

$$S' = \frac{\vec{\imath} + [0,1]^n}{r^k}.$$

The set $S \cap S'$ does not contain an element of the interior of the n-cube S' since, otherwise, $\text{aff}(S \cap S')$ would be of dimension $n - 1$. Hence, the intersection is on the faces of S'. It follows that, since S is a $(n-1)$-plane, $S \cap S'$ contains a vertex of S', i.e., there exists $\vec{\jmath} \in \{0,1\}^n$ such that

$$\frac{\vec{\imath} + \vec{\jmath}}{r^k} \in S \cap S'.$$

The conclusion follows since $\vec{v}$ can be defined by $\vec{v} = \vec{\imath} + \vec{\jmath}$. ∎

Lemma 7.29 *Let $n \in \mathbb{N}_{>1}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, $\vec{a} \in (\mathbb{Z}_{\neq 0})^n$, $b \in \mathbb{Z}$ and $S = \{\vec{x} \in \mathbb{R}^n \mid \vec{a}\vec{x} = b\}$. If there exist $\vec{v} \in \mathbb{Z}^n$ and $k \in \mathbb{Z}$ such that $\vec{a}\left(\frac{\vec{v}}{r^k}\right) = b$, then S is not r-suitable.*

Proof By contradiction, suppose that S is r-suitable.

For each $i \in \{1, 2, \ldots, n\}$, let S_i be the set

$$S_i = \frac{\left(\vec{v} - (\underbrace{0, \ldots, 0}_{i-1}, 1, \underbrace{0, \ldots, 0}_{n-i})\right) + [0,1]^n}{r^k}.$$

We have $\frac{\vec{v}}{r^k} \in S_i$ for each $i \in \{1, 2, \ldots, n\}$. Moreover, by definition, this vector belongs to S. Since S is r-suitable, we have $\dim(\mathrm{aff}(S \cap S_i)) = n-1$ for each $i \in \{1, 2, \ldots, n\}$. Since $\vec{a}[i] \neq 0$ for all i, the intersections $S \cap S_i$ have elements in all the interiors of the n-cubes S_i. Hence, since S is an affine space, we have $S = \mathbb{R}^n$. This is a contradiction with the fact that $\vec{a} \in (\mathbb{Z}_{\neq 0})^n$. Hence, S is not r-suitable. ∎

Combining Lemmas 7.28 and 7.29, we immediately get the following corollary :

Corollary 7.30 *Let $n \in \mathbb{N}_{>1}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, $\vec{a} \in (\mathbb{Z}_{\neq 0})^n$, $b \in \mathbb{Z}$ and $S = \{\vec{x} \in \mathbb{R}^n \mid \vec{a}\vec{x} = b\}$. There exist $\vec{v} \in \mathbb{Z}^n$ and $k \in \mathbb{Z}$ such that $\vec{a}\left(\frac{\vec{v}}{r^k}\right) = b$ if and only if S is not r-suitable.*

Corollary 7.30 will now be used to get an algebraic characterization of the r-suitable $(n-1)$-planes.

We need the following classical result of number theory.

Theorem 7.31 (Bézout's identity) *Let $n \in \mathbb{N}_{>0}$ and $a_1, a_2, \ldots, a_n \in \mathbb{Z}_{\neq 0}$. If $\gcd(a_1, a_2, \ldots, a_n) = 1$, then there exist $u_1, u_2, \ldots, u_n \in \mathbb{Z}$ such that*

$$\sum_{i=1}^{n} u_i a_i = 1.$$

We also need the following lemma :

Lemma 7.32 *Let $c \in \mathbb{Z}$, $n \in \mathbb{N}_{>0}$ and $a_1, a_2, \ldots, a_n \in \mathbb{Z}_{\neq 0}$. If there exist $u_1, u_2, \ldots, u_n \in \mathbb{Z}$ such that*

$$\sum_{i=1}^{n} u_i a_i = c,$$

then there exists $m \in \mathbb{Z}$ such that $c = m. \gcd(a_1, a_2, \ldots, a_n)$.

Proof By definition, the number $\gcd(a_1, a_2, \ldots, a_n)$ divides a_i for each $i \in \{1, 2, \ldots, n\}$. It thus divides each linear combination

$$\sum_{i=1}^{n} \lambda_i a_i,$$

with $\lambda_i \in \mathbb{Z}$ for all $i \in \{1, 2, \ldots, n\}$. In particular, $\gcd(a_1, a_2, \ldots, a_n)$ divides c. ■

We are now able to establish the main result of this section.

Theorem 7.33 *Let $n \in \mathbb{N}_{>1}$ be a dimension, $r \in \mathbb{N}_{>1}$ be a base, $\vec{a} \in (\mathbb{Z}_{\neq 0})^n$, $b \in \mathbb{Z}$ and $S = \{\vec{x} \in \mathbb{R}^n \mid \vec{a}\vec{x} = b\}$. Let $d \in \mathbb{N}_{>0}$ be the denominator of the reduced fraction of*

$$\frac{b}{\gcd(\vec{a}[1], \vec{a}[2], \ldots, \vec{a}[n])}.$$

The set S is not r-suitable if and only if d is a product of prime factors of r.

Proof Suppose that S is not r-suitable. By Lemma 7.28, there exist $\vec{v} \in \mathbb{Z}^n$ and $k \in \mathbb{Z}$ such that $\vec{a}\left(\dfrac{\vec{v}}{r^k}\right) = b$. We thus have

$$br^k = \vec{a}\vec{v}.$$

It then follows from Lemma 7.32 that there exists $m \in \mathbb{Z}$ such that

$$br^k = m.\gcd(\vec{a}[1], \vec{a}[2], \ldots, \vec{a}[n]).$$

Hence, we have

$$\frac{b}{\gcd(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n)} = \frac{m}{r^k}.$$

The denominator d of the reduced form of this fraction is necessarily a product of prime factors of r.

Reciprocally, suppose that d is a product of prime factors of r, and let $\vec{a'} \in \mathbb{Z}^n$ be the vector satisfying

$$\vec{a'} = \frac{\vec{a}}{\gcd(\vec{a}[1], \vec{a}[2], \ldots, \vec{a}[n])}.$$

We have $\gcd(\vec{a'}[1], \vec{a'}[2], \ldots, \vec{a'}[n])$. By Theorem 7.31, there exists $\vec{u} \in \mathbb{Z}^n$ such that $\vec{u}\vec{a'} = 1$. Multiplying by b, we get

$$b(\vec{u}\vec{a'}) = \left(\frac{b\vec{u}}{\gcd(\vec{a}[1], \vec{a}[2], \ldots, \vec{a}[n])}\right)\vec{a} = b.$$

It follows that the vector $\frac{b}{\gcd(\vec{a}[1], \vec{a}[2], \ldots, \vec{a}[n])}\vec{u}$ belongs to S. By hypothesis, the denominator d of the reduced form of $\frac{b}{\gcd(\vec{a}[1], \vec{a}[2], \ldots, \vec{a}[n])}$ is a product $r_1 r_2 \ldots r_p$ of prime factors $r_1, r_2, \ldots, r_p$ of r, with $p \in \mathbb{N}$. Multiplying both the numerator and denominator by

$$\prod_{i=1}^{p} \frac{r}{r_i},$$

we deduce that there exist $\vec{v} \in \mathbb{Z}^n$ and $k \in \mathbb{Z}$ such that $\frac{\vec{v}}{r^k} \in S$. Hence, by Lemma 7.29, S is not r-suitable. ■

Together, Lemma 7.27 and Theorems 7.33 provide an algebraic characterization of the r-suitable sets of the form $\{\vec{x} \in \mathbb{R}^n \mid \vec{a}\vec{x} = b\}$, with $\vec{a} \in \mathbb{Z}^n$ and $b \in \mathbb{Z}$.

Moreover, we know, by Definition 7.21, that the r-suitable polyhedra S are strongly related to the intersections of sets $\{\vec{x} \in \mathbb{R}^n \mid \vec{a_i}\vec{x} = b_i\}$, where S is a finite Boolean combination of the sets $\{\vec{x} \in \mathbb{R}^n \mid \vec{a_i}\vec{x} \leq b_i\}$. We could thus imagine that it could be easily possible to obtain a characterization of such sets S by considering individually the sets $\{\vec{x} \in \mathbb{R}^n \mid \vec{a_i}\vec{x} = b_i\}$. However, it is not the case, as shown in the following example.

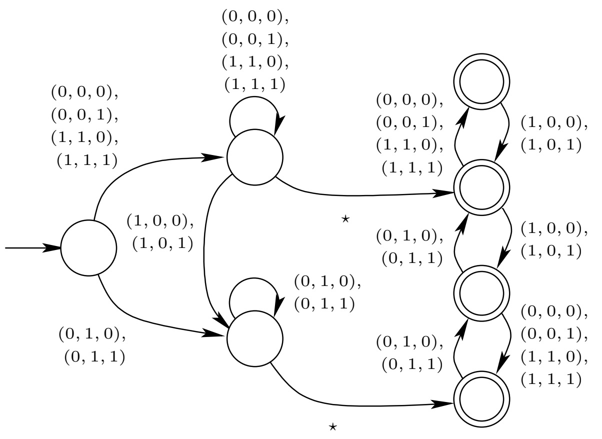

Figure 7.6: RVA representing the set $\{\vec{x} \in \mathbb{R}^3 \mid 3.\vec{x}[1] - 3.\vec{x}[2] = 1\}$.

Example 7.34 *By Theorem 7.33, the sets*

$$S = \{\vec{x} \in \mathbb{R}^3 \mid 3.\vec{x}[1] - 3.\vec{x}[2] = 1\}$$

and

$$S' = \{\vec{x} \in \mathbb{R}^2 \mid 3.\vec{x}[1] + 3.\vec{x}[3] = 1\}$$

are both 2*-suitable. RVA recognizing* S *and* S' *in base* 2 *are respectively represented in Figures 7.6 and 7.7. However, their intersection* $S \cap S'$ *is not* 2*-suitable. A RVA recognizing* $S \cap S'$ *is given in Figure 7.8. In this RVA, we can observe that for any state* $q \in \{10, 11, 12, 13\}$, $\dim(S(q)) = 0$. ◇

7.5 Summary

In this chapter, we began, in Section 7.1, by introducing polyhedra and giving related general results.

Next, in Section 7.2, we defined, with respect to a base $r \in \mathbb{N}_{>1}$, the notion of *r-suitable* polyhedra. These polyhedra form a large class of polyhedra with associated RVA that have a more simple structure than RVA recognizing general polyhedra. Indeed, RVA recognizing not r-suitable polyhedra can contain redundant structures that arise from the necessity of handling all encodings of the vectors that have to be recognized.

Using the results of the previous chapters, the internal structure of RVA recognizing the r-suitable polyhedra was documented in Section 7.3. Let $\mathcal{A}$

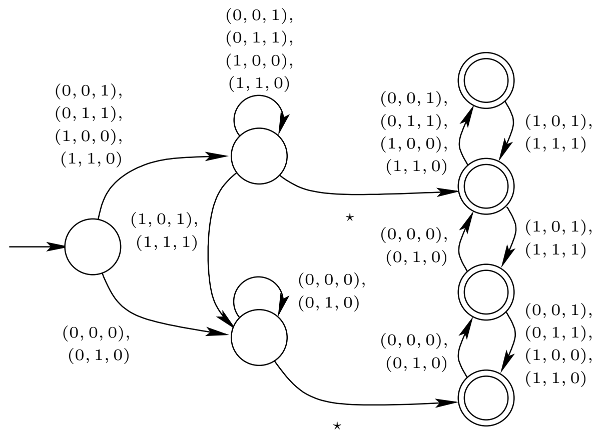

Figure 7.7: RVA representing the set $\{\vec{x} \in \mathbb{R}^2 \mid 3.\vec{x}[1] + 3.\vec{x}[3] = 1\}$.

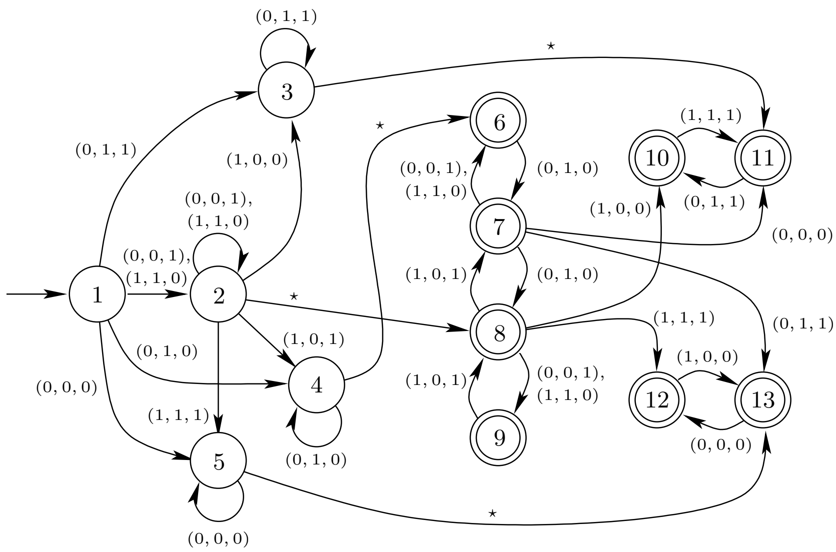

Figure 7.8: RVA representing the set $\{\vec{x} \in \mathbb{R}^2 \mid 3.\vec{x}[1] - 3.\vec{x}[2] = 1 \wedge 3.\vec{x}[1] + 3.\vec{x}[3] = 1\}$.

be such a RVA. Precisely, we obtained two main results. First, we established that the sets of vectors whose encodings are accepted from non-trivial strongly connected components of the fractional part of $\mathcal{A}$ are bounded pyramids. Second, we proved that for any pair q, q' of states of the fractional part of $\mathcal{A}$, if q' is reachable from q, and if S and S' are respectively the set of apexes of the bounded pyramids accepted from q and q', then the dimension of S is smaller or equal than the dimension of S'.

Finally, we gave the first steps of an algebraic characterization of the r-suitable polyhedra. We focused on the sets S of the form $S = \{\vec{x} \in \mathbb{R}^n \mid \vec{a}\vec{x} = b\}$, where $\vec{a} \in \mathbb{Z}^n$ and $b \in \mathbb{Z}$. We showed that we can, without loss of generality, only consider the vectors $\vec{a} \in (\mathbb{Z}_{\neq 0})^n$ and we remarked that if the dimension n is equal to 1, then those sets are all r-suitable. We established the following characterization : A set $S = \{\vec{x} \in \mathbb{R}^n \mid \vec{a}\vec{x} = b\}$, where $n \in \mathbb{N}_{>1}$, $\vec{a} \in (\mathbb{Z}_{\neq 0})^n$ and $b \in \mathbb{Z}$, is r-suitable if and only if the denominator of the reduced fraction of

$$\frac{b}{\gcd(\vec{a}[1], \vec{a}[2], \ldots, \vec{a}[n])}$$

has at least one prime factor not belonging to the set of prime factors of the base $r \in \mathbb{N}_{>1}$. We also noticed, by a counter-example, that this characterization does not trivially generalize to general r-suitable polyhedra.

Chapter 8

Conclusion

8.1 Summary

The main contribution of this thesis is a characterization of the subsets of $\mathbb{R}^n$ that can be recognized by finite automata in multiple integer bases. When a set of vectors is recognized in two sufficiently different bases, we have established that this set is necessarily definable in the first-order additive theory of real and integer variables $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

In the case of weak deterministic automata, used in actual implementations of symbolic representation systems [LASH, FAST, LIRA], we proved that the condition on the bases is the multiplicative independence. Since we established that recognizability in multiplicatively dependent bases is equivalent to recognizability in only one of them, we have thus obtained a complete characterization of the sets of real vectors recognized in multiple bases, similar to the one known for the domain $\mathbb{Z}^n$ [Cob69, Sem77].

For general automata, i.e., Muller or (possibly non-deterministic) Büchi ones, it turned out that the multiplicative independence of the bases is not a strong enough condition; we indeed provided, as a counter-example, a set simultaneously recognizable in two multiplicatively independent bases, but not definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$. In this case, we have demonstrated that the condition having to be satisfied by the bases to force definability of the represented sets in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$, is that the bases must have different sets of prime factors.

Several original developments have been achieved for proving those results. In particular, we established expressiveness properties of weak and general automata recognizing sets of vectors. We also provided reductions of those problems, from the domain $\mathbb{R}^n$ to the domain $[0,1]^n$. Another contribution is the introduction of the notions of product- and sum-stability, aimed at connecting structural properties of automata with arithmetical properties of the sets they represent.

As a corollary of our results, it is now established that for every subset S of $\mathbb{R}^n$ that is recognizable in all bases $r \in \mathbb{N}_{>1}$, the set S is necessarily definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$, which implies that it can be recognized by a weak deterministic automaton. In addition to the practical advantages of these automata, this brings a theoretical justification to their use for representing sets of integer and real vectors : If recognizability by automata has to be achieved regardless of the representation base, then the representable sets are exactly those that can be recognized by weak deterministic automata in every base.

The proofs of those main results lead in particular to two additional contributions.

The first one is a characterization of the sets definable in the theory $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$: They are exactly the sets that can be expressed as finite unions of sets having the form $S^I + S^F$, where $S^I \subseteq \mathbb{Z}^n$ is definable in $\langle \mathbb{Z}, +, < \rangle$, i.e, in Presburger arithmetic, and where $S^F \subseteq [0,1]^n$ is definable in $\langle \mathbb{R}, +, <, 1 \rangle$, i.e., the additive theory of reals without multiplication.

The other additional contribution is a precise insight into the structure of weak deterministic automata recognizing a large class of sets definable in the theory $\langle \mathbb{R}, +, <, 1 \rangle$. In the paper [BBD10], the knowledge of this structure has been exploited for developing a more efficient symbolic representation system suited for the sets definable in $\langle \mathbb{R}, +, <, 1 \rangle$. In particular, this new representation system avoids the unnecessarily large size of automata representing linear constraints with large coefficients; indeed, intuitively, it replaces some components of automata of large size by implicit algebraic descriptions, while keeping a canonical representation.

8.2 Relations with other work

This thesis is centered around automata recognizing sets of real vectors that are encoded positionally in an integer base. Other ways to encode numbers into words have been studied for a long time. For non-negative integer numbers, we can refer to [BM89, Fro92, Sha94, Lor95, BH97, Hol98, Fro01]. The recognizability properties of encodings obtained using such numeration systems have inspired a lot of research. A general class of systems for which the set $\mathbb{N}$ is recognizable is the class of *abstract numeration systems* [Rig01, Cha09]. For real numbers, other generalized numeration systems also exist, e.g., the β-numeration [Rén57, Par60, BS05]. The main difference with our approach is that our motivations are related to practical applications : We are looking for representation systems that can be used in practice for actual applications, in algorithmic and expressiveness perspectives, rather than more theoretical systems.

We have provided extensions of Cobham's and Semenov's theorems to automata recognizing sets of real vectors. It is worth mentioning that these famous theorems have been generalized in many other ways and in various contexts, e.g., [Vil92, Fab94, Bès97, PB97, Han98, Bès00, Dur02, AB08, Dur08]. In the same way as we did in Chapter 5, Kronecker's approximation theorem is often used in work related to Cobham's theorem. However, due to the particular nature of the problems we considered, most of the ideas of this thesis are original.

Another extension of Cobham's theorem can be found in the article [AB11], and is strongly related with the results of Chapter 5. This article does not consider any automata-based representation of sets, but rather deals with the notion of *r-self-similarity*, which can be connected with properties about the states of our RVA. It gives a link between the sets that are both r- and s-self-similar, where r and s are multiplicatively independent, and the form of the set that is represented. However, it only focuses on the one-dimensional sets that are compact; by contrast, our results do not have these limitations.

Automata recognizing sets of vectors of numbers encoded positionally in an integer base have inspired several results that can be related to this thesis. Procedures for synthesizing a formula from a NDD, thus representing a set of vectors of integers, exist [Ler05, Lat05]. Other work were interested to give bounds on the size of automata representing sets definable in $\langle \mathbb{Z}, +, < \rangle$ and $\langle \mathbb{R}, +, <, 1 \rangle$ [Kla08, Kla10], and are of real interest for analyzing the complexity of the relying decision procedures. We did not consider those problems in this thesis. Let us note that some procedures for performing complex operations on arithmetic automata are available. For instance, the paper [CLW09] provides an algorithm for computing a RVA recognizing the convex hull of a set represented by a NDD.

Work on the expressiveness of RVA appears in other publications. RVA were introduced in the paper [BBR97], and, one year later, their expressiveness, with respect to a given base r, was characterized in [BRW98]. The paper [JS01] also established expressiveness properties with respect to a base, but in a more geometrical perspective. In this work, closure properties of the sets that can be recognized are demonstrated, as well as results expressing that some classes of sets cannot be recognized, whatever the representation base is. Later, the article [BJW05] established that the sets definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$ are all recognizable by weak deterministic RVA, regardless of the encoding base $r \in \mathbb{N}_{>1}$. Part of the results of this thesis can be seen as the proof of the reciprocal property : Thanks to our results, we know that the sets that are recognizable in all bases, either by weak deterministic automata, or by general ones, are exactly the sets that are definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$.

The result giving a characterization of the theory $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$ in terms of $\langle \mathbb{Z}, +, < \rangle$ and $\langle \mathbb{R}, +, <, 1 \rangle$, obtained in the beginning of Chapter 7, is related to the paper [Wei99], in which a similar characterization is given, but holding only for one-dimensional sets.

Work aimed at improving the efficiency of arithmetic decision procedures based on RVA was developed in [EK08]. In this article, the authors identify redundant structures in RVA, introduce the notion of *don't care words* permitting to define smaller automata representing sets of $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$, and give adapted algorithms for manipulating such automata.

Finally, as mentioned in the previous section, our results have prompted, in [BBD10], the development of a new representation system, based on RVA, and able to represent the sets definable in $\langle \mathbb{R}, +, <, 1 \rangle$. A similar representation system has been developed in [GHH^{+}03, HK05], but is limited to the handling of three-dimensional sets. Our results give underlying theoretical bases of both representation systems.

8.3 Perspectives

As explained in the introduction of this thesis, our motivation of studying the properties of RVA rely in the development of tools aimed at verifying computer systems. In this perspective, research subjects concerning finite automata representing arithmetic sets are still abundant.

For exploiting the information contained in RVA, it is sometimes useful to be able to convert this information into other types of representations. A potential subject consists in the development of algorithms able to synthesize formulas representing the sets recognized by RVA, in the same way as the procedures provided in [Ler05, Lat05] for NDDs.

It should be possible to extend the results of our Chapter 7 in several ways. For instance, a promising direction is to try to get a characterization of the number of states of automata recognizing the sets of $\langle \mathbb{R}, +, <, 1 \rangle$, in a more precise way than the results of [Kla10], as well as an even precise documentation of the structure of such automata. Also in Chapter 7, we introduced the notion of *r-suitable sets* and we established the first steps of an algebraic characterization of these sets. This result could be extended to get a full characterization of the r-suitable sets.

In Chapter 5 and 6, we connected the structure of parts of RVA with arithmetical properties to deduce characteristics of the represented sets, for instance with the help of the notion of *product-stable sets*. A great challenge would be to try to apply similar techniques to automata recognizing integer vectors, that could perhaps lead to simpler proofs of Cobham's and Semenov's theorems.

We mentioned that a representation system based on RVA, but representing the sets in a more concise way, was developed in [BBD10]. It is called *Implicit Real Vector Automata (IRVA)*. Future research is needed to get a really efficient representation. For instance, in our potential applications, it is important to be able to compute projections of the represented sets; a research direction consists thus in the development of algorithms for projecting and determinizing sets represented by IRVA. Improving the conciseness of IRVA is also challenging. Finally, in some applications, it is useful to manage periodic sets. A long term work could then be to extend IRVA to represent sets definable in $\langle \mathbb{R}, \mathbb{Z}, +, < \rangle$, which can involve naturally periodicities, instead of sets of $\langle \mathbb{R}, +, <, 1 \rangle$.

Bibliography

[AB08] B. Adamczewski and J. Bell. Function fields in positive characteristic: Expansions and Cobham's theorem. *Journal of Algebra*, 319(6):2337–2350, 2008.

[AB11] B. Adamczewski and J. Bell. An analogue of Cobham's theorem for fractals. *Transactions of the American Mathematical Society*, To appear, 2011.

[BB07] B. Boigelot and J. Brusten. A generalization of Cobham's theorem to automata over real numbers. In *Proceedings of the 34th International Colloquium on Automata, Languages and Programming*, volume 4596 of *Lecture Notes in Computer Science*, pages 813–824. Springer, July 2007.

[BB09] B. Boigelot and J. Brusten. A generalization of Cobham's theorem to automata over real numbers. *Theoretical Computer Science*, 410(18):1694–1703, 2009.

[BBB08] B. Boigelot, J. Brusten, and V. Bruyère. On the sets of real numbers recognized by finite automata in multiple bases. In *Proceedings of the 35th International Colloquium on Automata, Languages and Programming*, volume 5126 of *Lecture Notes in Computer Science*, pages 112–123. Springer, July 2008.

[BBB10] B. Boigelot, J. Brusten, and V. Bruyère. On the sets of real numbers recognized by finite automata in multiple bases. *Logical Methods in Computer Science*, 6(1):1–17, 2010.

[BBD10] B. Boigelot, J. Brusten, and J.-F. Degbomont. Implicit Real Vector Automata. In *Proceedings of the 12th International Workshop on Verification of Infinite-State Systems*, volume 39 of *Electronic Proceedings in Theoretical Computer Science*, pages 63–76, Singapore, September 2010.

[BBL09] B. Boigelot, J. Brusten, and J. Leroux. A generalization of Semenov's theorem to automata over real numbers. In *Proceedings of the 22nd International Conference on Automated Deduction*, volume 5663 of *Lecture Notes in Artificial Intelligence*, pages 469–484. Springer, July 2009.

[BBR97] B. Boigelot, L. Bronne, and S. Rassart. An improved reachability analysis method for strongly linear hybrid systems. In *Proceedings of the 9th International Conference on Computer-Aided Verification*, volume 1254 of *Lecture Notes in Computer Science*, pages 167–177. Springer-Verlag, 1997.

[Bès97] A. Bès. Undecidable extensions of Büchi arithmetic and Cobham-Semenov theorem. *Journal of Symbolic Logic*, 62(4):1280–1296, 1997.

[Bès00] A. Bès. An extension of the Cobham-Semenov theorem. *Journal of Symbolic Logic*, 65(1):201–211, 2000.

[BFL08] F. Bouchy, A. Finkel, and J. Leroux. Decomposition of decidable first-order logics over integers and reals. In *Proceedings of the 15th International Symposium on Temporal Representation and Reasoning*, pages 147–155. IEEE Computer Society Press, June 2008.

[BH97] V. Bruyère and G. Hansel. Bertrand numeration systems and recognizability. *Theoretical computer science*, 181:17–43, 1997.

[BHMV94] V. Bruyère, G. Hansel, C. Michaux, and R. Villemaire. Logic and p-recognizable sets of integers. *Bulletin of the Belgian Mathematical Society*, 1(2):191–238, March 1994.

[BHZ08] R. Bagnara, P. M. Hill, and E. Zaffanella. The Parma Polyhedra Library: Toward a complete set of numerical abstractions for the analysis and verification of hardware and software systems. *Science of Computer Programming*, 72(1–2):3–21, 2008.

[Bie95] H. Bieri. Nef polyhedra: A brief introduction. *Computing Suppl.*, 10:43–60, 1995.

[BJW05] B. Boigelot, S. Jodogne, and P. Wolper. An effective decision procedure for linear arithmetic over the integers and reals. *ACM Transactions on Computational Logic*, 6(3):614–633, 2005.

[BL04] B. Boigelot and L. Latour. Counting the solutions of Presburger equations without enumerating them. *Theoretical Computer Science*, 313:17–29, 2004.

[BM89] A. Bertrand-Mathis. Comment écrire les nombres entiers dans une base que n'est pas entière. *Acta Math. Hung.*, 54:237–241, 1989.

[Boi98] B. Boigelot. *Symbolic Methods for Exploring Infinite State Spaces.* PhD thesis, Université de Liège, 1998.

[Bou71] N. Bourbaki. *Topologie générale : chapitres 1 à 4.* Hermann, 1971.

[Bru85] V. Bruyère. *Entiers et automates finis.* Graduate thesis, Université de Mons, 1985.

[Bru06] J. Brusten. *Etude des propriétés des RVA.* Graduate thesis, Université de Liège, May 2006.

[BRW98] B. Boigelot, S. Rassart, and P. Wolper. On the expressiveness of real and integer arithmetic automata. In *Proc. 25th ICALP*, volume 1443 of *Lecture Notes in Computer Science*, pages 152–163, Aalborg, July 1998. Springer.

[BS05] V. Berthé and A. Siegel. Tilings associated with beta-numeration and substitutions. *Integers*, 5(3):A2, 46 pp. (electronic), 2005.

[Büc60] J. R. Büchi. Weak second order arithmetic and finite automata. *Z. Math. Logik Grundlag. Math.*, 6:66–92, 1960.

[Büc62] J. R. Büchi. On a decision method in restricted second order arithmetic. In *Proc. International Congress on Logic, Methodoloy and Philosophy of Science*, pages 1–12, Stanford, 1962. Stanford University Press.

[CGP00] E.M. Clark, O. Grumberg, and D.A. Peled. *Model Checking.* The MIT Press, Cambridge, MA, 2000.

[CH78] P. Cousot and N. Halbwachs. Automatic discovery of linear restraints among variables of a program. In *Conference Record of the 5th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pages 84–97, Tucson, Arizona, 1978. ACM Press, New York, NY.

[Cha09] E. Charlier. *Abstract Numeration Systems : Recognizability, Decidability, Multidimensional S-Automatic Words, and Real Numbers.* PhD thesis, Université de Liège, 2009.

[CLW09] F. Cantin, A. Legay, and P. Wolper. Computing convex hulls by automata iteration. *International Journal of Foundations of Computer Science*, 20(4):647–667, 2009.

[Cob69] A. Cobham. On the base-dependence of sets of numbers recognizable by finite automata. *Mathematical Systems Theory*, 3:186–192, 1969.

[Dur02] F. Durand. A theorem of Cobham for non-primitive substitutions. *Acta Arithmetica*, 104(3):225–241, 2002.

[Dur08] F. Durand. Cobham-Semenov theorem and $\mathbb{N}^d$-subshifts. *Theoretical Computer Science*, 391:20–38, February 2008.

[Eil74] S. Eilenberg. *Automata, languages, and machines (volume A)*. Academic Press, San Diego, CA, 1974.

[EK08] J. Eisinger and F. Klaedtke. Don't care words with an application to the automata-based approach for real addition. *Formal Methods in System Design*, 33(1–3):85–115, 2008.

[Fab94] S. Fabre. Une généralisation du théorème de Cobham. *Acta Arithmetica*, 67:197–208, 1994.

[FAST] Fast Acceleration of Symbolic Transition systems (FAST). Available at : `http://www.lsv.ens-cachan.fr/fast/`.

[FR75] J. Ferrante and C. Rackoff. A decision procedure for the first order theory of real addition with order. *SIAM J. on Computing*, 4(1):69–76, March 1975.

[Fro92] C. Frougny. Representations of numbers and finite automata. *Math. Systems Theory*, 25:37–60, 1992.

[Fro01] C. Frougny. *Algebraic Combinatorics on words*, chapter Numeration Systems, pages 203–236. Cambridge University Press, 2001.

[GHH$^+$03] M. Granados, P. Hachenberger, S. Hert, L. Kettner, K. Mehlhorn, and M. Seel. Boolean operations on 3D selective Nef complexes – data structure, algorithms, and implementation. In *Proc. 11th ESA*, volume 2832 of *Lecture Notes in Computer Science*, pages 654–666. Springer, 2003.

[GS66] S. Ginsburg and E.H. Spanier. Semigroups, Presburger formulas and languages. *Pacific J. Math.*, 16:285–296, 1966.

[GV04] M.A. Granados Velásquez. *Fast point location, ray shooting and intersection test for 3D Nef polyhedra.* Bachelor thesis, Eafit University, 2004.

[Han98] G. Hansel. Systèmes de numération indépendants et syndéticité. *Theoretical Computer Science*, 204(1-2):119–130, 1998.

[HK05] P. Hachenberger and L. Kettner. Boolean operations on 3D selective Nef complexes: Optimized implementation and experiments. In *Proc. ACM Symposium on Solid and Physical Modeling*, pages 163–174. ACM, 2005.

[HMU01] J.E. Hopcroft, R. Motwani, and J.D. Ullman. *Introduction to Automata Theory, Languages, and Computation, Second Edition*. Addison-Wesley, 2001.

[Hol88] G.J. Holzmann. An improved protocol reachability analysis technique. *Software, Practice & Experience*, 18(2):137–161, February 1988.

[Hol98] M. Hollander. Greedy numeration systems and regularity. *Theory Comput. Syst.*, 31:111–133, 1998.

[Hop71] J. E. Hopcroft. An $n \log n$ algorithm for minimizing the states in a finite automaton. In Z. Kohavi, editor, *The Theory of Machines and Computation*, pages 189–196. Academic Press, New York, 1971.

[HW85] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. Oxford University Press, 5th edition, 1985.

[JS01] H. Jürgensen and L. Staiger. Finite automata encoding geometric figures. In O. Boldt and H. Jürgensen, editors, *Automata Implementation*, volume 2214 of *Lecture Notes in Computer Science*, pages 101–108. Springer, 2001.

[Kla08] F. Klaedtke. Bounds on the automata size for Presburger arithmetic. *ACM Transactions on Computational Logic*, 9(2), 2008.

[Kla10] F. Klaedtke. Ehrenfeucht-Fraïssé goes automatic for real addition. *Information and Computation*, 208(11):1283–1295, 2010.

[Kle56] S. C. Kleene. Representation of events in nerve nets and finite automata. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, Annals of Mathematics Studies, pages 3–41. Princeton University Press, 1956.

[KV97] O. Kupferman and M.Y. Vardi. Weak alternating automata are not that weak. In *5th Israeli Symposium on Theory of Computing and Systems*, pages 147–158. IEEE Computer Society Press, 1997.

[Lan69] L. H. Landweber. Decision problems for ω-automata. *Mathematical Systems Theory*, 3(4):376–384, 1969.

[LASH] The Liège Automata-based Symbolic Handler (LASH). Available at : `http://www.montefiore.ulg.ac.be/~boigelot/research/lash/`.

[Lat05] L. Latour. *Presburger arithmetic: from automata to formulas*. PhD thesis, Université de Liège, 2005.

[Ler05] J. Leroux. A polynomial time Presburger criterion and synthesis for number decision diagrams. In *Proceedings of the 20th IEEE Symposium on Logic in Computer Science*, pages 147–156, Chicago, 2005. IEEE Computer Society.

[LIRA] Linear Integer/Real Arithmetic solver (LIRA). Available at : `http://lira.gforge.avacs.org/`.

[Lor95] N. Loraud. β-shift, systèmes de numération et automates. *J. Théorie Nombres Bordeaux*, 7:473–498, 1995.

[LV92] H. Le Verge. A note on Chernikova's algorithm. Technical report, IRISA, Rennes, 1992.

[Löd01] C. Löding. Efficient minimization of deterministic weak ω-automata. *Information Processing Letters*, 79(3):105–109, 2001.

[McM93] K.L. McMillan. *Symbolic Model Checking*. Kluwer Academic Publishers, 1993.

[McN66] R. McNaughton. Testing and generating infinite sequences by a finite automaton. *Information and Control*, 9(5):521–530, 1966.

[MH84] S. Miyano and T. Hayashi. Alternating finite automata on ω-words. *Theoretical Computer Science*, 32:321–330, 1984.

[MS97] O. Maler and L. Staiger. On syntactic congruences for ω-languages. *Theoretical Computer Science*, 183:93–112, 1997.

[Mun75] J. R. Munkres. *Topology: A First Course*. Prentice-Hall, Englewood Cliffs, New Jersey, 1975.

[Nef78] W. Nef. *Beiträge Zur Theorie Der Polyeder Mit Anwendungen in Der Computergraphik (Contributions to the Theory of Polyhedra, with Applications in Computer Graphics)*. Verlag H. Lang & Cie. AG, Bern, 1978.

[Par60] W. Parry. On the β-expansions of real numbers. *Acta Math. Acad. Sci. Hungar.*, 11:401–416, 1960.

[PB97] F. Point and V. Bruyère. On the Cobham-Semenov theorem. *Theory Comput. Syst.*, 30(2):197–220, 1997.

[Per90] D. Perrin. Finite automata. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science, Volume B: Formal Models and Semantics*, pages 1–57. Elsevier and MIT Press, 1990.

[PP04] D. Perrin and J.E. Pin. *Infinite words*, volume 141 of *Pure and Applied Mathematics*. Elsevier, 2004.

[Pre29] M. Presburger. Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt. In *Comptes Rendus du Premier Congrès des Mathématiciens des Pays Slaves*, pages 92–101, Warsaw, 1929.

[Pug92] W. Pugh. The Omega Test : A fast and practical integer programming algorithm for dependence analysis. *Communications of the ACM*, 8:4–13, 1992.

[Rén57] A. Rényi. Representations for real numbers and their ergodic properties. *Acta Math. Acad. Sci. Hungar*, 8:477–493, 1957.

[Rig01] M. Rigo. *Abstract numeration systems on a regular language and recognizability.* PhD thesis, Université de Liège, 2001.

[RS59] M. O. Rabin and D. Scott. Finite automata and their decision problems. *IBM journal of research and developement*, 3(2):114–125, 1959.

[Saf88] S. Safra. On the complexity of ω-automata. In *Proc. 29th Symposium on Foundations of Computer Science*, pages 319–327. IEEE Computer Society, October 1988.

[Sem77] A.L. Semenov. Presburgerness of predicates regular in two number systems. *Siberian Mathematical Journal*, 18:289–299, 1977.

[Sha94] J. O. Shallit. Numeration systems, linear recurrences, and regular sets. *Inform. Comput.*, 113:331–347, 1994.

[Sta83] L. Staiger. Finite-state ω-languages. *Journal of Computer and System Sciences*, 27(3):434–448, 1983.

[SW74] L. Staiger and K. Wagner. Automatentheoretische und automatenfreie Charakterisierungen topologischer Klassen regulärer Folgenmengen. *Elektronische Informationsverarbeitung und Kybernetik*, 10(7):379–392, 1974.

[Var07] M. Vardi. The Büchi complementation saga. In *Proc. 24th. STACS*, volume 4393 of *Lecture Notes in Computer Science*, pages 12–22, Aachen, February 2007. Springer.

[Vil92] R. Villemaire. Joining k- and l-recognizable sets of natural numbers. In *Proceedings of the 9th Annual Symposium on Theoretical Aspects of Computer Science*, volume 577 of *Lecture Notes in Computer Science*, pages 83–94, Cachan, February 1992. Springer.

[WB95] P. Wolper and B. Boigelot. An automata-theoretic approach to Presburger arithmetic constraints. In *Proceedings of the 2nd Static Analysis Symposium*, volume 983 of *Lecture Notes in Computer Science*, pages 21–32, Glasgow, September 1995. Springer.

[WB00] P. Wolper and B. Boigelot. On the construction of automata from linear arithmetic constraints. In *Proceedings of the 6th International Conference on Tools and Algorithms for Construction and analysis of Systems*, volume 1785 of *Lecture Notes in Computer Science*, pages 1–19, Berlin, March 2000. Springer.

[Wei99] V. Weispfenning. Mixed real-integer linear quantifier elimination. In *Proc. ACM SIGSAM ISSAC*, pages 129–136, Vancouver, July 1999. ACM Press.

[Wes78] C.H. West. Generalized technique for communication protocol validation. *IBM Journal of Research and Development*, 22:394–404, 1978.

[Wil93] T. Wilke. Locally threshold testable languages of infinite words. In *Proceedings of the 10th Annual Symposium on Theoretical Aspects of Computer Science*, volume 665 of *Lecture Notes in Computer Science*, pages 607–616, Würzburg, 1993. Springer.