Termination Analysis of Term Rewriting by Polynomial Interpretations and Matrix Interpretations

dissertation

by

Friedrich Neurauter

submitted to the Faculty of Mathematics, Computer Science and Physics of the University of Innsbruck

in partial fulfillment of the requirements for the degree of "Doktor der technischen Wissenschaften"

advisor: Univ.-Prof. Dr. Aart Middeldorp

Innsbruck, March 2012

dissertation

Termination Analysis of Term Rewriting by Polynomial Interpretations and Matrix Interpretations

Friedrich Neurauter

March 2012

advisor: Univ.-Prof. Dr. Aart Middeldorp

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht.
Die vorliegende Arbeit wurde bisher in gleicher oder ähnlicher Form noch nicht als Magister-/Master-/Diplomarbeit/Dissertation eingereicht.

Datum	Unterschrift

Abstract

This thesis is concerned with termination and complexity analysis of term rewrite systems. Term rewriting is a formal model of computation based on equational logic. Due to Turing-completeness, all interesting properties of term rewrite systems are undecidable. Nevertheless, many powerful termination techniques have been developed in the course of time. In this thesis, we focus on polynomial interpretations and matrix interpretations, giving the answer to a number of open research questions related to termination and complexity analysis.

The method of polynomial interpretations is one of the oldest termination techniques, but still successfully employed in many automatic termination analyzers. One distinguishes three variants, polynomial interpretations with real, rational and integer coefficients, which raises the question of their mutual relationship with regard to termination proving power. In 2006, a partial answer was given by Lucas who managed to prove that there are term rewrite systems that can be shown terminating by polynomial interpretations with rational coefficients, but cannot be shown terminating using integer polynomials only. He also proved the existence of systems that can only be handled by polynomial interpretations with real (algebraic) coefficients. In this thesis, we extend these results and give the full picture of the relationship, thereby refuting a common yet unproven belief in the term rewriting community about this relationship.

Since their inception in 2006, matrix interpretations have evolved into one of the most important termination techniques. In analogy to polynomial interpretations, there are three variants depending on the domain of the matrix entries, matrix interpretations over the real, rational and natural numbers. We clarify their relationship by showing that matrix interpretations over the reals are more powerful than matrix interpretations over the rationals, which are in turn more powerful than matrix interpretations over the natural numbers. We also show how the choice of the matrix dimension affects termination proving power. Beyond termination analysis, matrix interpretations are the most important technique for obtaining polynomial upper bounds on the (derivational) complexity of term rewrite systems, where the aim is to obtain information about the maximal length of rewrite sequences in relation to the size of their initial term. In particular, triangular matrix interpretations over the natural numbers are known to induce polynomial upper bounds. Recently, this method was improved by an automata-based approach giving a complete characterization of polynomially bounded matrix interpretations over the natural numbers. In this thesis, we present an algebraic approach which subsumes all previous approaches and provides a complete characterization of polynomially bounded matrix interpretations over the real, rational and natural numbers.

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my supervisor Aart Middeldorp for his guidance and continuous support during the process of writing this thesis, always being all ears to my problems and keeping me focused. I truly appreciate his expertise and outstanding ability to ask the right questions at the right time.

I am also grateful to Georg Moser for introducing me into research and supervising my master's thesis, which marked the beginning of my adventures in computer science.

In the matter of funding, I am especially indebted to the University of Innsbruck for giving me the opportunity to obtain a university position as well as for providing me with further financial support during the last year in the form of a "Doktoratsstipendium aus der Nachwuchsförderung 2010".

Furthermore, I would like to thank all the people who have supported me in the course of my studies, including all members, past and present, of the Computational Logic group, namely, Aart, Andreas, Bertram, Chris, Christian, Clemens, Georg, Harald, Martin A., Martin K., Martina, Nao, René, Sarah, Simon and Stefan. It was a pleasure to be a part of this research group.

Special thanks are due to Chris for providing all those neat TeX templates and to my office mates Bertram, Harald and Martin K. for many interesting and fruitful discussions. I am especially indebted to Harald for doing a marvellous job at implementing some of the theoretical results obtained in this thesis and for being my main source of information on questions related to the termination analyzer $\mathsf{T_TT_2}$. Bertram also deserves extra credit for his helpful comments in the early stages of some of the work contained in this thesis.

Last but not least, my utmost appreciation and thanks go to Claudia, my dear companion and partner in life, and my family for their faith and support in a matter so unfamiliar to them.

Contents

Introduction

Since the advent of ubiquitous computing, computer-based systems play an ever-increasing role in virtually all aspects of everyday life. While the failure of certain systems may be tolerable, there are many applications of computers where safety is mission-critical (e.g., in the part of the on-board computer system of a car which controls its anti-skid braking system, in systems controlling nuclear power plants, etc.). Needless to say, proving the correctness of such safety-critical systems is of utmost importance. A crucial task in establishing formal correctness proofs of computer programs is to show that they always yield a result after a finite number of computation steps. This essential property is called *termination* and is well-known to be undecidable in general. So there cannot be a single method capable of analyzing (i.e., proving or disproving) termination for all programs.

In the course of this PhD thesis, termination analysis is studied at the level of *term rewrite systems*. In theoretical computer science, *term rewriting* [4, 70] is a conceptually simple but Turing-complete model of computation whose foundation is equational logic and which is very close to functional programming. Due to Turing-completeness, any Turing machine, as well as any program written in some contemporary programming language, can be simulated by a corresponding term rewrite system. Hence, one can reduce the question of termination of programs to termination of term rewrite systems. What distinguishes term rewriting from equational logic is that equations are used as *directed* reduction rules, where left-hand sides can be replaced by the corresponding right-hand sides, but not vice versa. Thus, a term rewrite system can be viewed as a set of directed equations, called *rewrite rules*, between first-order *terms* built from *variables* and *function symbols*, which models computation as a sequence of *rewrite steps*, starting from some initial term, where each step corresponds to the replacement of an occurrence of the left-hand side of some rewrite rule in a term by its corresponding right-hand side. Immediately, two questions come to mind in this context. First, can this process go on forever or is it guaranteed to terminate after finitely many steps? Second, assuming that all rewrite sequences (or *derivations*) terminate, what can be said about their maximal lengths? Alas, due to Turing-completeness, all interesting properties of term rewrite systems are undecidable. In particular, termination is undecidable (cf. e.g. [4, 70]). Nevertheless, many powerful termination techniques for term rewrite systems have been developed in the course of time, most notably the *dependency pair framework* [3, 25–28, 71], the state-of-the-art framework for termination analysis due to its ability to integrate and combine arbitrary termination techniques in a modular and uniform way [71]. Moreover, Hofbauer and Lautemann observe in [31] that "proving termination with one of these specific techniques in general proves more than just the absence of infinite derivations. It turns out that in many cases such a proof implies an upper bound on the maximal length of derivations", which they consider as a natural measure for the complexity of (terminating) term rewrite systems, formalized in the notion of *derivational complexity*, where the length of a longest derivation is related to the size of its initial term. In the recent past much progress has been made in establishing sufficient and *automatable* criteria for termination and complexity analysis, as is evident in the results of the (annual) international competition for termination and complexity tools.[1]

In this thesis, we put our focus on two specific termination techniques, namely, *polynomial interpretations* (in the first part) and *matrix interpretations* (in the second part), giving the answer to a number of open research questions related to termination and complexity analysis.

Polynomial Interpretations

The method of polynomial interpretations is one of the oldest techniques for proving termination of term rewrite systems, dating back to the (late) seventies of the last century. While originally conceived by Lankford [43] as a means for establishing direct termination proofs, polynomial interpretations are nowadays often used in the context of the dependency pair framework. In the classical approach of Lankford [43], one considers polynomial algebras over the well-founded domain of the natural numbers $\mathbb{N}$ induced by interpreting each function symbol occurring in a given term rewrite system by a polynomial function with integer coefficients that is required to return a natural number whenever all its arguments are from $\mathbb{N}$ and that must be monotone (in all arguments) with respect to the natural order $>_{\mathbb{N}}$ on $\mathbb{N}$. This induces a mapping from terms to natural numbers in the obvious way. Termination can be concluded if for each rewrite rule $\ell \to r$, the polynomial P_ℓ associated with the left-hand side is greater (with respect to $>_{\mathbb{N}}$) than P_r, the corresponding polynomial of the right-hand side, (for all arguments ranging over $\mathbb{N}$) because then any rewrite step between two terms causes a decrease in the associated (natural) numbers. Hence, the well-foundedness of $>_{\mathbb{N}}$ implies the absence of infinite rewrite sequences.

Already back in the seventies an alternative approach using polynomials with real coefficients instead of integers was proposed by Dershowitz [18]. However, due to the fact that the real numbers $\mathbb{R}$ equipped with the natural order $>_{\mathbb{R}}$ are not well-founded, a subterm property is explicitly required to ensure well-foundedness. It was not until 2005 that this limitation was overcome, when Lucas [45] presented a framework for proving (polynomial) termination over the real numbers, where well-foundedness is basically achieved by replacing $>_{\mathbb{R}}$ by a new ordering $>_{\mathbb{R},\delta}$ requiring comparisons between real numbers to not be below a given positive real number δ. The two approaches of [18] and [45] were compared in [47], with the result that the latter is *strictly better* than the former. Therefore, we employ the notion of polynomial interpretations over the real numbers of [45], which also facilitates polynomial interpretations over the rational numbers $\mathbb{Q}$.

[1] http://termcomp.uibk.ac.at

Thus, one distinguishes three variants of polynomial interpretations, polynomial interpretations with real, rational and integer coefficients, which raises the obvious question:

What is their relationship with regard to termination proving power?

Giving a complete answer to this question is the primary aim of the first part of this thesis. Despite the fact that polynomials with real coefficients include polynomials with rational coefficients, which in turn include polynomials with integer coefficients, no clear statement about the relative power of the derived termination techniques appeared in the literature until in 2006 a partial answer was given by Lucas [46] for direct polynomial termination (where all rules $\ell \to r$ must satisfy $P_\ell > P_r$). He proved that there are term rewrite systems that can be shown terminating by polynomial interpretations with rational coefficients, but cannot be shown terminating using polynomials with integer coefficients only. Likewise, he proved that there are term rewrite systems that can be handled by polynomial interpretations with real (algebraic) coefficients, but cannot be handled by polynomial interpretations with rational coefficients. In this thesis, we extend these results and give the full picture of the relationship between the different variants of polynomial interpretations, thereby refuting the common yet unproven belief (expressed in e.g. [10, 47]) in the term rewriting community that polynomial interpretations with real coefficients properly subsume polynomial interpretations with rational coefficients, which in turn properly subsume polynomial interpretations with integer coefficients. In this respect, our main contributions are as follows:

1. First, we show that polynomial interpretations with real or rational coefficients do not properly subsume polynomial interpretations with integer coefficients.

2. Then we show that there are term rewrite systems that can be proved terminating by polynomial interpretations with real and with integer coefficients but not with rational coefficients.

3. Our third result shows that polynomial interpretations with real coefficients subsume polynomial interpretations with rational coefficients.

4. Finally, extending an earlier result of [47], we prove that transcendental real numbers are irrelevant for termination proofs based on polynomial interpretations, a result confirming that automatic termination tools may restrict to the real algebraic numbers $\mathbb{R}_{\mathsf{alg}}$ (without losing power).

Figure A illustrates both our results and the earlier results of [46] (for direct polynomial termination). We also consider the possibility of establishing termination by using polynomial interpretations in an *incremental*[2] fashion, showing that the same relationship applies, and we indicate how to adapt these results to the dependency pair framework, thereby obtaining evidence that the

[2] That is, proving termination by a sequence of polynomial interpretations, each of which removes some rewrite rules until eventually all rewrite rules have been removed.

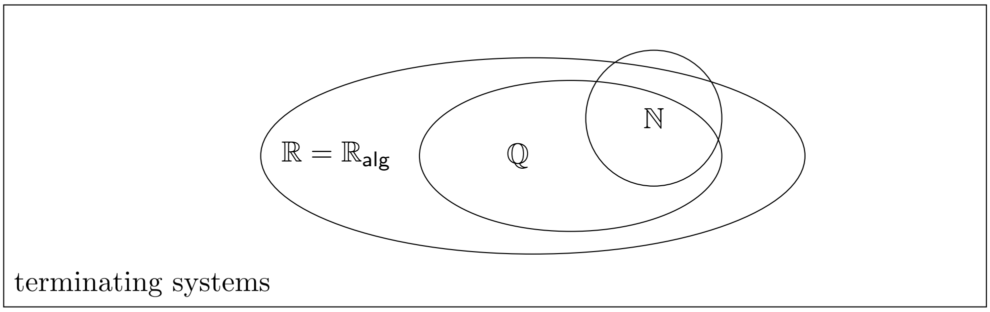

Figure A: Comparison.

relationship is no different in that setting. Furthermore, we study monotonicity criteria for polynomials with respect to the orders $>_{\mathbb{N}}$ and $>_{\mathbb{R},\delta}$ ($>_{\mathbb{Q},\delta}$), based on which we obtain a new result in connection with simple termination showing that polynomial interpretations with real (rational) coefficients (as proposed in [45]) can prove termination of non-simply terminating systems, in contrast to the approaches of Lankford [43] and Dershowitz [18], which are well-known to enforce simple termination. (This result cannot be established using the monotonicity criteria proposed in [45].)

Matrix Interpretations

The second part of this thesis is dedicated to matrix interpretations, a generalization of linear polynomial interpretations. Since their inception in 2006, matrix interpretations have evolved into one of the most important methods for termination and complexity analysis of term rewrite systems. While originally introduced by Hofbauer and Waldmann as a stand-alone method for establishing termination proofs in the context of string rewriting [32, 33], it was not long until Endrullis *et al.* [20] generalized (one particular instance of) the matrix method to term rewriting and also incorporated it into the dependency pair framework. The basic principle underlying the approach of [20] is the same as that of polynomial interpretations except that terms are mapped to a well-founded domain of vectors of natural numbers (rather than the natural numbers). Function symbols are interpreted by suitable linear mappings represented by square matrices of natural numbers. In [1, 22, 76] the method was lifted to the non-negative rational and real numbers using the same technique that was already used to lift polynomial interpretations from the natural numbers to the rationals and reals. So, in analogy to polynomial interpretations, we have matrix interpretations over the real, rational and natural numbers, which again raises the question of their mutual relationship with regard to termination proving power. Giving a complete answer to this open question is the first major goal in the second part of this thesis. Before stating our results, we mention related work appearing in [23] and [48]. In [23] a relative termination problem is presented that can be handled with matrix interpretations over the rationals but not with matrix interpretations over the natural numbers. However, *relative* termination is essential in this example because the relative component is the key ingredient for precluding matrix interpretations over the natural numbers. As the latter component consists of a single non-terminating rule, the entire example does not readily generalize to (real) termination problems. Besides, there is no evidence in [23] demonstrating the benefit of using irrational numbers in matrix interpretations. In contrast, the author of [48] raises the question whether *rational numbers are somehow unnecessary when dealing with matrix interpretations* based on the observation that matrix interpretations over the rationals can sometimes be simulated with matrix interpretations over the natural numbers. However, our results show that the answer is in the negative. To be precise, we prove that matrix interpretations over the real numbers are more powerful with respect to proving termination than matrix interpretations over the rational numbers, which are in turn more powerful than matrix interpretations over the natural numbers (as a stand-alone termination technique as well as in the setting of the dependency pair framework). Besides, we also show how the choice of the matrix dimension affects termination proving power.

Beyond termination analysis, matrix interpretations are also apt for analyzing the *derivational complexity* of term rewrite systems. In fact, they are the most important technique for obtaining (non-linear) polynomial upper bounds, which are of special interest since they are associated with *feasible* computations. However, in general, the complexity bounds obtained from matrix interpretations are exponential. So in order to obtain polynomial upper bounds, additional conditions must be satisfied. Historically, the first approach appearing in the literature [53] was developed for matrix interpretations over the natural numbers, achieving its goal by restricting the shape of all matrices to upper triangular form. In [73] this method of triangular matrix interpretations was subsumed by an automata-based approach, where matrices are viewed as weighted (word) automata computing a weight function, which is required to be polynomially bounded. The result is a complete characterization (i.e., necessary and sufficient conditions) of polynomially bounded matrix interpretations over the natural numbers. In this thesis, we present an algebraic approach which subsumes all previous approaches and provides a complete characterization of polynomially bounded matrix interpretations over the real, rational and natural numbers.

Personal Contribution and Outline

Most of the results presented in the course of this thesis have already been published in various papers [49, 54–56, 58, 59]. An extended journal version of [54] has been submitted for publication (cf. [57]). Conceptually, the focus of the thesis lies entirely on my personal contributions to these papers. This includes all theoretical results of [54–59] as well as the main results of [49], but excludes the experimental results of [49, 58, 59], which are due to H. Zankl. All results obtained by my co-authors (or others) are indicated as such by referencing the corresponding papers.

The remainder of this thesis is divided into two parts. In Part I, consisting of Chapters 1 – 4, we present our research on polynomial interpretations. After introducing some mathematical preliminaries as well as terminology and notation related to term rewriting in Chapter 1, we introduce all relevant concepts and definitions related to polynomial interpretations in Chapter 2 and show that polynomial interpretations over the rationals and reals do not imply simple termination. In Chapter 3, we study monotonicity criteria for polynomial interpretations, before presenting the main result of Part I in Chapter 4, where we give the full picture of the relationship between the various instances of polynomial interpretations.

In Part II, consisting of Chapters 5 – 7, we present our results on matrix interpretations. We start by introducing matrix interpretations and all the necessary background material in Chapter 5. Then, in Chapter 6, we clarify the relationship between matrix interpretations over the real, rational and natural numbers. Finally, in Chapter 7, we give a complete characterization of matrix interpretations inducing polynomial upper bounds on the derivational complexity of term rewrite systems.

Part I

Polynomial Interpretations

Chapter 1

Preliminaries

In this chapter, we introduce some mathematical preliminaries as well as terminology and notation related to term rewriting. For an in-depth introduction to the latter, we refer to [4, 70]. In Section 1.1, we recall a few basic notions from polynomial algebra, before presenting the relevant background material on term rewriting in Sections 1.2, 1.3 and 1.4.

Throughout this thesis, $\mathbb{N}$ denotes the set $\{0, 1, 2, \dots\}$ of natural numbers, whereas $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ refer to the integer, rational and real numbers, respectively. An *irrational* number is a real number, which is not in $\mathbb{Q}$. A real number is said to be *algebraic* if it is a root of a non-zero polynomial in one indeterminate with integer coefficients, otherwise it is said to be *transcendental*. The set of all real algebraic numbers is denoted by $\mathbb{R}_{\mathsf{alg}}$. Given $D \in \{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$ and some $m \in D$, $>_D$ (resp. $>$ if D is clear from the context) denotes the natural order of the respective domain, $\geqslant_D$ (resp. $\geqslant$) its reflexive closure, and D_m abbreviates $\{x \in D \mid x \geqslant m\}$; for example, $\mathbb{Q}_0$ ($\mathbb{R}_0$) refers to the set of all non-negative rational (real) numbers. We use the following notation for *intervals* of real numbers: we write $[a, b]$ for the *closed interval* $\{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$ and (a, b) for the *open interval* $\{x \in \mathbb{R} \mid a < x < b\}$.

1.1 Polynomials

Let R be a commutative ring (e.g., $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$, $\mathbb{R}$). We denote the associated *polynomial ring* in n *indeterminates* $x_1, \dots, x_n$ by $R[x_1, \dots, x_n]$, the elements of which are finite sums of products of the form $c \cdot x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$, where the *coefficient* c is an element of R and the exponents $i_1, \dots, i_n$ in the *monomial* $x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$ are non-negative integers. As usual, we implicitly assume (without loss of generality) that no two summands have the same (sequence of) exponents. The coefficient associated with the monomial $x_1^0 x_2^0 \cdots x_n^0$ is called the *constant coefficient*. If $c \neq 0$, we call a product $c \cdot x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$ a *term*. An element p of $R[x_1, \dots, x_n]$ is called an *(n-variate) polynomial with coefficients in R*. For example, the polynomial $2x^2 - x + 1$ is an element of $\mathbb{Z}[x]$, the ring of all univariate polynomials with integer coefficients. Polynomials are well-known to be closed under addition, multiplication and composition. The *degree* of a polynomial p, denoted by $\deg(p)$, is the maximum degree of its terms, where the degree of a term is just the sum of the exponents of the monomial associated with it. The degree of the zero polynomial $p = 0$ is defined to be $-\infty$.

In the special case $n = 1$, a polynomial $p \in R[x]$ can be written as $p(x) = \sum_{k=0}^{d} a_k x^k$ for some $d \in \mathbb{N}$. For the largest k such that $a_k \neq 0$, we call $a_k x^k$ the *leading term* of p, a_k its *leading coefficient* and note that $\deg(p) = k$. We say that p is *monic* if its leading coefficient is one. It is said to be *linear*, *quadratic*, *cubic* if its degree is one, two, three.

Every polynomial $p \in R[x_1, \dots, x_n]$ induces a function $p\colon R^n \to R$ as follows: given $a := (a_1, \dots, a_n) \in R^n$, replace x_i by a_i for $i = 1, \dots, n$ in the expression for p. If the result equals zero, then a is said to be a *root* of p (in R). In the remainder of this thesis, we will often identify a polynomial with the function it induces.

For $D \subseteq \mathbb{R}$, we call a function $f\colon D^n \to D$ *positive* (*non-negative*) on $A \subseteq D^n$ if $f(a) \underset{(=)}{>} 0$ for all $a \in A$. In particular, positiveness (non-negativeness) of a polynomial with coefficients in $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ or $\mathbb{R}$ corresponds to positiveness (non-negativeness) of its induced function.

1.2 Term Rewriting

A *signature* is a set of *function symbols*, each of which is equipped with a fixed *arity*. Function symbols of arity zero are also called *constant symbols* (or just *constants*), whereas function symbols of arity one (two) are referred to as *unary* (*binary*) function symbols. For a signature $\mathcal{F}$ and a countably infinite set of *variables* $\mathcal{V}$ disjoint from $\mathcal{F}$, the set of *terms* over $\mathcal{F}$ and $\mathcal{V}$ is denoted by $\mathcal{T}(\mathcal{F}, \mathcal{V})$. The set of variables (function symbols) occurring in a term t is denoted by $\mathcal{V}\mathrm{ar}(t)$ ($\mathcal{F}\mathrm{un}(t)$), and $|t|_a$ denotes the number of occurrences of a symbol $a \in \mathcal{F} \cup \mathcal{V}$ in t. A term is called *linear* if each variable occurs at most once in it. In case $\mathcal{V}\mathrm{ar}(t) = \emptyset$, t is said to be a *ground* term. The set of all ground terms over $\mathcal{F}$ is denoted by $\mathcal{T}(\mathcal{F}, \emptyset)$ or simply $\mathcal{T}(\mathcal{F})$. The *size* $|t|$ of a term t is defined as the number of function symbols and variables occurring in it, whereas its *depth* is defined as follows: $\mathrm{depth}(t) = 0$ if t is a variable or a constant, otherwise, if $t = f(t_1, \dots, t_n)$, $\mathrm{depth}(t) = 1 + \max\{\, \mathrm{depth}(t_i) \mid 1 \leqslant i \leqslant n \,\}$. The *root* of a non-variable term $t = f(t_1, \dots, t_n)$ is defined as $\mathrm{root}(t) = f$.

Let $\Box$ be a fresh constant symbol, called *hole*, not occurring in $\mathcal{F} \cup \mathcal{V}$. A *context* is a term from $\mathcal{T}(\mathcal{F} \cup \{\Box\}, \mathcal{V})$ with exactly one occurrence of $\Box$. If C is a context and t a term, then the expression $C[t]$ denotes the term obtained by replacing the hole in C by t. We say that a term s is a *subterm* of a term t, denoted by $s \trianglelefteq t$, if there exists a context C such that $t = C[s]$; s is called a *proper subterm* if $C \neq \Box$, in which case we write $s \triangleleft t$. A *substitution* is a mapping from $\mathcal{V}$ to $\mathcal{T}(\mathcal{F}, \mathcal{V})$. As usual, $t\sigma$ denotes the result of applying a substitution σ to a term t, that is, replacing each variable x occurring in t by $\sigma(x)$. A binary relation R on terms is *closed under contexts* if $s \mathrel{R} t$ implies $C[s] \mathrel{R} C[t]$ for all terms s, t and contexts C. It is *closed under substitutions* if $s \mathrel{R} t$ implies $s\sigma \mathrel{R} t\sigma$ for all terms s, t and substitutions σ. A *rewrite relation* is a binary relation on terms that is closed under contexts and substitutions.

A *rewrite rule* is a pair of terms (ℓ, r), conveniently written as $\ell \to r$, such that the *left-hand side* ℓ is not a variable and all variables of the *right-hand side* r are contained in ℓ, i.e., $\mathcal{V}\mathrm{ar}(r) \subseteq \mathcal{V}\mathrm{ar}(\ell)$. A rewrite rule $\ell \to r$ is called *left-linear* (*right-linear*) if ℓ (r) is a linear term. It is said to be *linear* if both ℓ and r are linear and *duplicating* if some variable x occurs more often in r than in ℓ, i.e., $|r|_x > |\ell|_x$. A *term rewrite system* (TRS) over $\mathcal{T}(\mathcal{F},\mathcal{V})$ is a finite set of rewrite rules $\ell \to r$ such that $\ell, r \in \mathcal{T}(\mathcal{F},\mathcal{V})$. If all function symbols occurring in a TRS are unary, then we call it a *string rewrite system* (SRS). A TRS is *linear* (*left-linear*, *right-linear*) if all its rewrite rules have the corresponding property, it is *non-duplicating* if it contains no duplicating rule. The *rewrite relation* $\to_\mathcal{R}$ induced by a TRS $\mathcal{R}$ is a binary relation on terms, which is defined as follows: $s \to_\mathcal{R} t$ for two terms s and t if and only if there exist a rewrite rule $\ell \to r \in \mathcal{R}$, a substitution σ and a context C such that $s = C[\ell\sigma]$ and $t = C[r\sigma]$. We call $s \to_\mathcal{R} t$ a *rewrite step* (or simply an $\mathcal{R}$*-step*). As usual, $\to_\mathcal{R}^+$ ($\to_\mathcal{R}^*$) denotes the transitive (and reflexive) closure of $\to_\mathcal{R}$ and $\to_\mathcal{R}^n$ its n-th iterate. For notational convenience, we sometimes drop the subscript $\mathcal{R}$ if it is clear from the context. A term s is called a *normal form* if there is no term t such that $s \to_\mathcal{R} t$. A TRS $\mathcal{R}$ over $\mathcal{T}(\mathcal{F},\mathcal{V})$ is *terminating* if $\to_\mathcal{R}$ is well-founded, that is, if there is no infinite rewrite sequence $t_1 \to_\mathcal{R} t_2 \to_\mathcal{R} t_3 \to_\mathcal{R} t_4 \to_\mathcal{R} \cdots$ of terms in $\mathcal{T}(\mathcal{F},\mathcal{V})$. A TRS $\mathcal{R}$ is *terminating relative to* a TRS $\mathcal{S}$ if every rewrite sequence in $\mathcal{R} \cup \mathcal{S}$ contains only finitely many $\mathcal{R}$-steps.

We illustrate some of the above concepts by means of the following example.

Example 1.1. Let $\mathcal{F} = \{\mathsf{0}, \mathsf{s}, \mathsf{add}\}$ be a finite signature consisting of a binary function symbol add, a unary symbol s and a constant $\mathsf{0}$, and let us consider the TRS $\mathcal{R}$ consisting of the following rewrite rules:

$$\begin{aligned} \mathsf{add}(\mathsf{0}, y) &\to y \\ \mathsf{add}(\mathsf{s}(x), y) &\to \mathsf{s}(\mathsf{add}(x, y)) \end{aligned}$$

This TRS specifies the addition of two natural numbers in unary notation, where the natural number zero is represented by the constant $\mathsf{0}$, whereas the unary symbol s represents the successor function $x \mapsto x + 1$ on the natural numbers. Thus, we can represent the set of all natural numbers $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ by the set of ground terms $\mathcal{T}_\mathbb{N} = \{\mathsf{0}, \mathsf{s}(\mathsf{0}), \mathsf{s}(\mathsf{s}(\mathsf{0})), \mathsf{s}(\mathsf{s}(\mathsf{s}(\mathsf{0}))), \ldots\}$. For two terms t_1 and t_2 representing the numbers n_1 and n_2, respectively, the addition $n_1 + n_2$ is represented by the term $t = \mathsf{add}(t_1, t_2)$. The result of such a computation is obtained by applying the rewrite rules of the TRS $\mathcal{R}$ exhaustively to t; e.g., the following rewrite sequence computes the result of adding two and two:

$$\begin{aligned} \mathsf{add}(\mathsf{s}(\mathsf{s}(\mathsf{0})), \mathsf{s}(\mathsf{s}(\mathsf{0}))) &\to_\mathcal{R} \mathsf{s}(\mathsf{add}(\mathsf{s}(\mathsf{0}), \mathsf{s}(\mathsf{s}(\mathsf{0})))) \to_\mathcal{R} \\ &\qquad \mathsf{s}(\mathsf{s}(\mathsf{add}(\mathsf{0}, \mathsf{s}(\mathsf{s}(\mathsf{0}))))) \to_\mathcal{R} \mathsf{s}(\mathsf{s}(\mathsf{s}(\mathsf{s}(\mathsf{0})))) \end{aligned}$$

Later, in Example 1.8, we shall see that any such computation is guaranteed to come to an end after finitely many steps. In other words, we will show that the TRS $\mathcal{R}$ is terminating.

1.3 Monotone Algebras

Monotone algebras play an important role in the context of termination analysis of TRSs. We use the following terminology (cf. [20, 70]).

Definition 1.2. Let $\mathcal{F}$ be a signature and $\mathcal{V}$ a countably infinite set of variables disjoint from $\mathcal{F}$. An *$\mathcal{F}$-algebra* $\mathcal{A}$ consists of a non-empty *carrier* set A and a collection of *interpretation functions* $f_A\colon A^n \to A$ for each n-ary function symbol $f \in \mathcal{F}$. A *variable assignment* for $\mathcal{A}$ is a mapping from $\mathcal{V}$ to A. The *evaluation* or *interpretation* $[\alpha]_{\mathcal{A}}(t)$ of a term $t \in \mathcal{T}(\mathcal{F}, \mathcal{V})$ with respect to a variable assignment α is inductively defined as follows:

$$[\alpha]_{\mathcal{A}}(t) = \begin{cases} \alpha(t) & \text{if } t \in \mathcal{V} \\ f_A([\alpha]_{\mathcal{A}}(t_1), \ldots, [\alpha]_{\mathcal{A}}(t_n)) & \text{if } t = f(t_1, \ldots, t_n) \end{cases}$$

We shall sometimes abbreviate $[\alpha]_{\mathcal{A}}(t)$ by $[\alpha](t)$ if $\mathcal{A}$ is clear from the context. For $i \in \{1, \ldots, n\}$, an interpretation function $f_A\colon A^n \to A$ is *monotone in its i-th argument* with respect to a binary relation $\sqsupset$ on A if $a_i \sqsupset b$ implies

$$f_A(a_1, \ldots, a_i, \ldots, a_n) \sqsupset f_A(a_1, \ldots, b, \ldots, a_n)$$

for all $a_1, \ldots, a_n, b \in A$. It is said to be *monotone* with respect to $\sqsupset$ if it is monotone in all its arguments.

Definition 1.3. Let $\mathcal{A} = (A, \{f_A\}_{f \in \mathcal{F}})$ and $\mathcal{B} = (B, \{f_B\}_{f \in \mathcal{F}})$ be $\mathcal{F}$-algebras. A *homomorphism* from $\mathcal{A}$ to $\mathcal{B}$ is a mapping $h\colon A \to B$ such that for each n-ary function symbol $f \in \mathcal{F}$, $h(f_A(a_1, \ldots, a_n)) = f_B(h(a_1), \ldots, h(a_n))$ for all $a_1, \ldots, a_n \in A$. An *isomorphism* is a bijective homomorphism.

Definition 1.4. Let $\mathcal{F}$ and $\mathcal{V}$ be as above, and let $(\mathcal{A}, >, \geqslant)$ be an $\mathcal{F}$-algebra $\mathcal{A} = (A, \{f_A\}_{f \in \mathcal{F}})$ together with two binary relations $>$ and $\geqslant$ on A. We lift the latter from A to $\mathcal{T}(\mathcal{F}, \mathcal{V})$ as follows: for $s, t \in \mathcal{T}(\mathcal{F}, \mathcal{V})$, we define

- $s >_{\mathcal{A}} t$ if and only if $[\alpha]_{\mathcal{A}}(s) > [\alpha]_{\mathcal{A}}(t)$ for all assignments α, and
- $s \geqslant_{\mathcal{A}} t$ if and only if $[\alpha]_{\mathcal{A}}(s) \geqslant [\alpha]_{\mathcal{A}}(t)$ for all assignments α.

For a set of rewrite rules $\mathcal{R}$ we say that $(\mathcal{A}, >, \geqslant)$ and $\mathcal{R}$ are *(weakly) compatible* if $\ell >_{\mathcal{A}} r$ ($\ell \geqslant_{\mathcal{A}} r$) for each rewrite rule $\ell \to r \in \mathcal{R}$. We use the following abbreviations: $\mathcal{R} \subseteq {>_{\mathcal{A}}}$ for compatibility and $\mathcal{R} \subseteq {\geqslant_{\mathcal{A}}}$ for weak compatibility. The triple $(\mathcal{A}, >, \geqslant)$ (or just $\mathcal{A}$ if $>$ and $\geqslant$ are clear from the context) is a *weakly (strictly) monotone $\mathcal{F}$-algebra* if $>$ is well-founded, ${>} \cdot {\geqslant} \subseteq {>}$ and for each $f \in \mathcal{F}$, f_A is *weakly (strictly) monotone*, that is, monotone with respect to $\geqslant$ ($>$). It is said to be an *extended monotone $\mathcal{F}$-algebra* if it is both weakly monotone and strictly monotone. Finally, we call $(\mathcal{A}, >, \geqslant)$ a *well-founded monotone $\mathcal{F}$-algebra* if $>$ is a well-founded order (i.e., a transitive and irreflexive relation) on A, $\geqslant$ is its reflexive closure, and each interpretation function is strictly monotone.

It is well-known that well-founded monotone algebras provide a complete characterization of termination (cf. [70, Theorem 6.2.2]).

Theorem 1.5. *A TRS is terminating if and only if there exists a well-founded monotone algebra that is compatible with it.* □

Similarly, extended monotone algebras characterize relative termination according to [20, Theorem 2].

Theorem 1.6. *A TRS $\mathcal{R}$ is terminating relative to a TRS $\mathcal{S}$ if and only if there exists an extended monotone algebra that is compatible with $\mathcal{R}$ and weakly compatible with $\mathcal{S}$.* □

Thus, for $\mathcal{S} = \emptyset$, we conclude that a TRS is terminating if and only if there exists an extended monotone algebra that is compatible with it. In fact, as weak compatibility is not involved in the latter characterization of termination, weak monotonicity of the interpretation functions can be dispensed with altogether. This is obvious from the proof of [20, Theorem 2]. So we obtain the following corollary.

Corollary 1.7. *A TRS is terminating if and only if there exists a strictly monotone algebra that is compatible with it.* □

Note that strictly monotone algebras correspond to well-founded monotone algebras with respect to compatibility of TRSs if $>$ is not only a well-founded relation but also an order. This will be the case throughout this thesis. Moreover, note that any well-founded monotone algebra $(\mathcal{A}, >, \geqslant)$ is also an extended monotone algebra because $\geqslant$ is the reflexive closure of $>$. Therefore, monotonicity with respect to $>$ implies monotonicity with respect to $\geqslant$.

Next we mention another important aspect related to extended monotone algebras, namely, the fact that they facilitate *incremental termination proofs* (cf. [20, Theorem 3]). To this end, let $\mathcal{A}$ be an extended monotone algebra and suppose $\mathcal{R}$ is a TRS such that $\mathcal{R} \subseteq \geqslant_{\mathcal{A}}$ and $\mathcal{S} \subseteq >_{\mathcal{A}}$ for some non-empty subset $\mathcal{S}$ of $\mathcal{R}$. Then, after removing all $\mathcal{S}$-rules from $\mathcal{R}$, termination of $\mathcal{R} \setminus \mathcal{S}$ implies termination of $\mathcal{R}$. Thus, one is free to choose a different extended monotone algebra for the remaining rules $\mathcal{R} \setminus \mathcal{S}$. This process is continued until eventually all rewrite rules have been removed.

Example 1.8. We use the concepts introduced above to show that the TRS $\mathcal{R}$ of Example 1.1 is terminating. For this purpose, let us interpret all functions symbols in an $\mathcal{F}$-algebra over the carrier $\mathbb{N}$, where $\mathcal{F} = \{\mathsf{0}, \mathsf{s}, \mathsf{add}\}$:

$$\mathsf{0}_{\mathbb{N}} = 0 \qquad \mathsf{s}_{\mathbb{N}}(x) = x + 1 \qquad \mathsf{add}_{\mathbb{N}}(x, y) = 2x + y$$

It is easily verified that the pair $\mathcal{A} = (\mathbb{N}, \{f_{\mathbb{N}}\}_{f \in \mathcal{F}})$ is indeed a valid $\mathcal{F}$-algebra according to Definition 1.2. This is due to the fact that all interpretation functions are polynomial functions with non-negative integer coefficients. In addition, all of them are monotone with respect to the natural order $>_{\mathbb{N}}$ on $\mathbb{N}$. In particular, monotonicity of $\mathsf{0}_{\mathbb{N}}$ is vacuously satisfied (as it has no arguments, cf. Definition 1.2). Furthermore, we also have monotonicity of all interpretation functions with respect to $\geqslant_{\mathbb{N}}$. As a consequence, the triple $(\mathcal{A}, >_{\mathbb{N}}, \geqslant_{\mathbb{N}})$ is an extended monotone $\mathcal{F}$-algebra. Together with a variable assignment α mapping (term) variables to natural numbers, it gives rise to the following interpretations of the rules of $\mathcal{R}$:

$$[\alpha]_{\mathcal{A}}(\mathsf{add}(\mathsf{0}, y)) = \alpha(y) \geqslant \alpha(y) = [\alpha]_{\mathcal{A}}(y)$$
$$[\alpha]_{\mathcal{A}}(\mathsf{add}(\mathsf{s}(x), y)) = 2\alpha(x) + \alpha(y) + 2 > 2\alpha(x) + \alpha(y) + 1 = [\alpha]_{\mathcal{A}}(\mathsf{s}(\mathsf{add}(x, y)))$$

These inequalities simplify to $0 \geqslant 0$ and $1 > 0$, so they hold for any variable assignment α. Therefore, the algebra $\mathcal{A}$ is weakly compatible with both rules of $\mathcal{R}$ and (strictly) compatible with the second rule. After removing this rule, termination of the first rule implies termination of the overall system. But termination of the first rule can easily be shown by the following interpretation: $\mathsf{0}_{\mathbb{N}} = 1$ and $\mathsf{add}_{\mathbb{N}}(x, y) = x + y$. Hence, termination of $\mathcal{R}$ follows. Alternatively, a slight modification of the original interpretation to

$$\mathsf{0}_{\mathbb{N}} = 0 \qquad \mathsf{s}_{\mathbb{N}}(x) = x + 1 \qquad \mathsf{add}_{\mathbb{N}}(x, y) = 2x + y + 1$$

results in an extended monotone $\mathcal{F}$-algebra that is compatible with both rules (at the same time), thus establishing termination of $\mathcal{R}$ directly via Corollary 1.7.

Weakly monotone algebras play an important role in the context of termination analysis in the dependency pair framework.

1.4 Dependency Pair Framework

The dependency pair (DP) framework [3, 25–28, 71] is the state-of-the-art framework for termination analysis of TRSs. Being a modular extension of the dependency pair method of Arts and Giesl [3], which was originally regarded as just another termination technique (amongst many others), the main benefit of the DP framework is due to its ability to integrate and combine arbitrary termination techniques in a modular and uniform way [71] (besides facilitating the development of new methods for termination analysis). In what follows, we give a simplified account of the DP framework which is sufficient for our purposes.

Let $\mathcal{R}$ be a TRS over some signature $\mathcal{F}$. The set of *defined symbols* of $\mathcal{R}$ is given by $\mathcal{F}_{\mathcal{D}} = \{\mathrm{root}(\ell) \mid \ell \to r \in \mathcal{R}\}$. For each defined symbol $f \in \mathcal{F}_{\mathcal{D}}$, we introduce a fresh *dependency pair symbol* $f^\sharp$ of the same arity (as f), and for a term $t = f(t_1, \ldots, t_n)$ with $f \in \mathcal{F}_{\mathcal{D}}$, we denote by $t^\sharp$ the result of replacing its root symbol f by $f^\sharp$.

Definition 1.9. For $\mathcal{F}$ and $\mathcal{R}$ as above, the set of *dependency pairs* of $\mathcal{R}$ is given by $\mathsf{DP}(\mathcal{R}) = \{\ell^\sharp \to u^\sharp \mid \ell \to r \in \mathcal{R}, u \trianglelefteq r, u \ntriangleleft \ell, \mathrm{root}(u) \in \mathcal{F}_{\mathcal{D}}\}$.

In the DP framework, the problem of establishing termination of a given TRS is modularized by splitting it into several subproblems called DP problems, which can then be treated separately using different termination techniques. A *DP problem* is a pair $(\mathcal{P}, \mathcal{S})$, where $\mathcal{P}$ and $\mathcal{S}$ are finite sets of rewrite rules such that the root symbols of the rules in $\mathcal{P}$ neither occur in $\mathcal{S}$ nor in proper subterms of the left- and right-hand sides of the rules in $\mathcal{P}$. In the sequel, we sometimes write $(\mathcal{P}, \_)$ (resp. $(\_, \mathcal{S})$) to indicate that we are only interested in the first (resp. second) component of a DP problem. A DP problem $(\mathcal{P}, \mathcal{S})$ is *finite* if there is no infinite rewrite sequence $s_1 \to_{\mathcal{P}} t_1 \to_{\mathcal{S}}^* s_2 \to_{\mathcal{P}} t_2 \to_{\mathcal{S}}^* \cdots$ such that all terms t_i $(i = 1, 2, \ldots)$ are terminating with respect to $\to_{\mathcal{S}}$. Such an infinite sequence is said to be *minimal*. Every DP problem of the form $(\emptyset, \_)$ is finite. The following theorem captures the main result underlying the dependency pair approach.

Theorem 1.10. *A TRS $\mathcal{R}$ is terminating if and only if the DP problem* $(\mathsf{DP}(\mathcal{R}), \mathcal{R})$ *is finite.* $\square$

The pair $(\mathsf{DP}(\mathcal{R}), \mathcal{R})$ is referred to as the *initial DP problem* of $\mathcal{R}$. So according to Theorem 1.10, proving termination amounts to proving finiteness of a DP problem. The latter is achieved by means of *DP processors*, functions taking a DP problem as input and returning a set of DP problems as output. In order to use a DP processor Φ for proving termination, it must be *sound*, that is, if all DP problems in $\Phi(\mathcal{P}, \mathcal{S})$ are finite, then the original problem $(\mathcal{P}, \mathcal{S})$ is finite. Of course, the intention is to transform $(\mathcal{P}, \mathcal{S})$ into a set of "simpler" problems, simpler in the sense that proving finiteness of the latter is easier than proving finiteness of $(\mathcal{P}, \mathcal{S})$. Therefore, one is especially interested in processors that decrease $\mathcal{P}$ and/or $\mathcal{S}$.

The general procedure for establishing termination of a TRS $\mathcal{R}$ in the DP framework is to (try to) prove finiteness of its initial DP problem by recursively applying sound DP processors, thereby creating a tree whose nodes are DP problems and whose root is the initial DP problem. The children of a node $(\mathcal{P}, \mathcal{S})$ to which some processor Φ is applied are the single DP problems in $\Phi(\mathcal{P}, \mathcal{S})$. If at some point no new DP problems are generated, that is, if the final processors return empty sets of DP problems, then all leaves in the tree are finite DP problems (typically but not necessarily of the form $(\emptyset, \_)$). This implies finiteness of the initial DP problem, hence termination of $\mathcal{R}$, as finiteness propagates right up from the leaves of the tree to its root due to soundness of all processors.

In the context of this thesis, we only consider DP processors based on reduction pairs. Given a DP problem $(\mathcal{P}, \mathcal{S})$, the aim of such a processor is to return a simplified version of its input by removing rules from the $\mathcal{P}$ component. Formally, a *reduction pair* $(>, \gtrsim)$ consists of a well-founded order $>$ and a reflexive and transitive relation $\gtrsim$ (on terms) such that

1. $>$ is closed under substitutions,
2. $\gtrsim$ is closed under contexts and substitutions, and
3. $> \cdot \gtrsim \subseteq >$ or $\gtrsim \cdot > \subseteq >$.

Theorem 1.11. *For any reduction pair $(>, \gtrsim)$, the processor that maps a DP problem $(\mathcal{P}, \mathcal{S})$ to*

- $\{(\mathcal{P} \setminus \mathcal{P}', \mathcal{S})\}$ *if* $\mathcal{P}' \subseteq >$ *and* $(\mathcal{P} \setminus \mathcal{P}') \cup \mathcal{S} \subseteq \gtrsim$ *for some* $\mathcal{P}' \subseteq \mathcal{P}$
- $\{(\mathcal{P}, \mathcal{S})\}$ *otherwise*

is sound. $\square$

Such processors are called *reduction pair processors*. We say that a reduction pair processor Φ *succeeds* on a DP problem $(\mathcal{P}, \mathcal{S})$ if $\Phi(\mathcal{P}, \mathcal{S}) \neq (\mathcal{P}, \mathcal{S})$, otherwise it *fails*. It is well-known (and easily verified) that $(>_{\mathcal{A}}, \geqslant_{\mathcal{A}})$ is a reduction pair for every weakly monotone algebra $(\mathcal{A}, >, \geqslant)$, where $>$ is not only a well-founded relation but also an order and $\geqslant$ is reflexive and transitive. (The latter requirements on $>$ and $\geqslant$ are satisfied by all weakly monotone algebras considered in this thesis. However, these properties are not essential for obtaining a sound processor [20].) We say that a weakly monotone algebra $(\mathcal{A}, >, \geqslant)$ *succeeds* (*fails*) on a DP problem if the reduction pair processor based on $(>_{\mathcal{A}}, \geqslant_{\mathcal{A}})$ succeeds (fails) on it.

The number of constraints that need to be satisfied in Theorem 1.11 can often be reduced based on the observation that a reduction pair does not necessarily have to satisfy $\ell \gtrsim r$ for all rules $\ell \to r$ in $\mathcal{S}$ but just for a subset of rules known as the usable rules [3,26,28,71]. Formally, the set of *usable rules* of a DP problem $(\mathcal{P}, \mathcal{S})$ is given by $\mathsf{U}(\mathcal{P}, \mathcal{S}) = \{\ell \to r \in \mathcal{S} \mid \mathrm{root}(\ell) \in \mathsf{US}(t) \text{ for some } s \to t \in \mathcal{P}\}$, where $\mathsf{US}(t) = \emptyset$ if t is a variable and $\mathsf{US}(t) = \mathsf{US}(f(t_1, \ldots, t_n))$ is the least set such that $f \in \mathsf{US}(t)$, $\mathsf{US}(t_i) \subseteq \mathsf{US}(t)$ for all $i \in \{1, \ldots, n\}$, and for all rules $\ell \to r \in \mathcal{S}$, $\mathrm{root}(\ell) \in \mathsf{US}(t)$ implies $\mathcal{F}\mathrm{un}(r) \subseteq \mathsf{US}(t)$. The integration of usable rules into the reduction pair processor of Theorem 1.11 allows to weaken the requirement $\mathcal{S} \subseteq \gtrsim$ to $\mathsf{U}(\mathcal{P}, \mathcal{S}) \subseteq \gtrsim$, provided that $\mathcal{C}_\varepsilon$*-compatibility* is guaranteed, which means that for a fresh function symbol c the conditions $\mathsf{c}(x, y) \gtrsim x$ and $\mathsf{c}(x, y) \gtrsim y$ must hold (cf. [26,28]).

Theorem 1.12. *For any $\mathcal{C}_\varepsilon$-compatible reduction pair $(>, \gtrsim)$, the processor that maps a DP problem $(\mathcal{P}, \mathcal{S})$ to*

- $\{(\mathcal{P} \setminus \mathcal{P}', \mathcal{S})\}$ *if* $\mathcal{P}' \subseteq >$ *and* $(\mathcal{P} \setminus \mathcal{P}') \cup \mathsf{U}(\mathcal{P}, \mathcal{S}) \subseteq \gtrsim$ *for some* $\mathcal{P}' \subseteq \mathcal{P}$
- $\{(\mathcal{P}, \mathcal{S})\}$ *otherwise*

is sound. □

The result of Theorem 1.12 can be strengthened by coupling usable rules with *argument filters* [26] and/or by improving the computation of usable rules using unification instead of just looking at the root symbols (cf. e.g. [71]). Argument filters are implicit in reduction pairs based on weakly monotone algebras induced by polynomial interpretations and matrix interpretations. Because of this, and due to the fact that all reduction pairs considered in the sequel fall into this category, we skip a detailed discussion of argument filters here.

Chapter 2

Polynomial Interpretations

In this chapter, we introduce all relevant concepts and definitions related to polynomial interpretations and recall some well-known basic results concerning decidability, total termination and simple termination. In particular, we formally define the following variants of polynomial interpretations: polynomial interpretations over the real, rational and natural numbers. We also present a new result in connection with simple termination showing that, unlike polynomial interpretations over the natural numbers, polynomial interpretations over the rationals and reals do not enforce simple termination.

The idea of using polynomial interpretations for proving termination of TRSs dates back to the (late) seventies of the last century. In the classical approach of Lankford [43], one considers polynomials with integer coefficients inducing polynomial algebras over the well-founded domain of the natural numbers. To be precise, each n-ary function symbol f is interpreted by a polynomial in n indeterminates with integer coefficients, which induces a mapping from terms to integer numbers in the obvious way. In order to conclude termination of a given TRS, three conditions have to be satisfied. First, every polynomial must be *well-defined*, that is, it must induce a well-defined polynomial function $f_{\mathbb{N}} \colon \mathbb{N}^n \to \mathbb{N}$ over the natural numbers. In addition, all interpretation functions are required to be monotone with respect to the natural order $>_{\mathbb{N}}$ on $\mathbb{N}$. Finally, one has to show compatibility of the interpretation with the given TRS, i.e., for each rewrite rule $\ell \to r$, the polynomial P_ℓ associated with the left-hand side must be greater (with respect to $>_{\mathbb{N}}$) than P_r, the corresponding polynomial of the right-hand side, for all values of the indeterminates (ranging over $\mathbb{N}$).

Using the terminology of Chapter 1, polynomial interpretations à la Lankford are just special well-founded monotone algebras (resp. extended monotone algebras) over the carrier $\mathbb{N}$, where all interpretation functions are given by polynomials with integer coefficients, and proving termination of a TRS $\mathcal{R}$ amounts to finding an interpretation that is compatible with (all rules of) $\mathcal{R}$. However, a thorough examination of [43] reveals that Lankford was already aware of the possibility of establishing termination by using polynomial interpretations in an *incremental* way (essentially using the approach outlined in Section 1.3 for extended monotone algebras).

Apart from using polynomial interpretations as a stand-alone termination method (as described above), they are nowadays often employed in the context of the DP framework, which has the great advantage that the induced algebras only need to be weakly monotone rather than strictly monotone (cf. Section 1.3).

Already back in the seventies an alternative approach using polynomials with real coefficients instead of integers was proposed by Dershowitz [18]. However, due to the fact that the real numbers $\mathbb{R}$ equipped with the natural order $>_{\mathbb{R}}$ are not well-founded, a subterm property is explicitly required to ensure well-foundedness, i.e., each interpretation function $f_{\mathbb{R}} \in \mathbb{R}[x_1, \ldots, x_n]$ must additionally satisfy $f_{\mathbb{R}}(\ldots, x_i, \ldots) >_{\mathbb{R}} x_i$ for all $i \in \{1, \ldots, n\}$ and all $x_1, \ldots, x_n \in A \subseteq \mathbb{R}$, where $A \neq \emptyset$ is the intended carrier of the induced algebra. It was not until 2005 that this limitation was overcome, when Lucas [45] presented a framework for proving (polynomial) termination over the real numbers, where well-foundedness is basically achieved by replacing $>_{\mathbb{R}}$ by a new ordering $>_{\mathbb{R},\delta}$ requiring comparisons between real numbers to not be below a given positive real number δ. The two approaches of [18] and [45] were compared in [47], with the result that the latter is *strictly better* than the former. Therefore, we employ the notion of polynomial interpretations over the real numbers of [45], which also facilitates polynomial interpretations over the rational numbers without further ado.

2.1 Definitions

For polynomial interpretations, the notion of well-definedness of a (polynomial) function is important.

Definition 2.1. Let f be a function from D^n to D with $D \subseteq \mathbb{R}$ and $n \geqslant 0$. We say that f is *well-defined* over $A \subseteq D$ if $f(x_1, \ldots, x_n) \in A$ for all $x_1, \ldots, x_n \in A$; in particular, $f \in A$ if $n = 0$.

In other words, well-definedness over $A \subseteq D$ of a function f from D^n to D means that the restriction of f to A^n is a function from A^n to A. In particular, well-definedness of f over D_0 is equivalent to non-negativeness of f on D_0^n. Further, note that well-definedness of a polynomial (with coefficients in $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ or $\mathbb{R}$) corresponds to well-definedness of the function it induces.

2.1.1 Polynomial Interpretations over the Natural Numbers

Definition 2.2. A *polynomial interpretation over* $\mathbb{N}$ for a signature $\mathcal{F}$ consists of a polynomial $f_{\mathbb{N}} \in \mathbb{Z}[x_1, \ldots, x_n]$ for each n-ary function symbol $f \in \mathcal{F}$ such that $f_{\mathbb{N}}$ is well-defined over $\mathbb{N}$.

Due to well-definedness, each of the polynomials $f_{\mathbb{N}}$ induces a function from $\mathbb{N}^n$ to $\mathbb{N}$. Hence, the pair $\mathcal{N} = (\mathbb{N}, \{f_{\mathbb{N}}\}_{f \in \mathcal{F}})$ constitutes an $\mathcal{F}$-algebra over the carrier $\mathbb{N}$. In the sequel, we often identify a polynomial interpretation with its associated $\mathcal{F}$-algebra. In conjunction with the natural order $>_{\mathbb{N}}$ on $\mathbb{N}$ and its reflexive closure $\geqslant_{\mathbb{N}}$, we obtain an algebra $(\mathcal{N}, >_{\mathbb{N}}, \geqslant_{\mathbb{N}})$, which is easily seen to be a weakly monotone $\mathcal{F}$-algebra if for each $f \in \mathcal{F}$, $f_{\mathbb{N}}$ is weakly monotone. It is strictly monotone if each interpretation function $f_{\mathbb{N}}$ is strictly monotone, in which case it is even an extended monotone algebra because strict monotonicity of $f_{\mathbb{N}}$ implies weak monotonicity, that is, monotonicity with respect to $>_{\mathbb{N}}$ implies monotonicity with respect to $\geqslant_{\mathbb{N}}$.

Definition 2.3. A polynomial interpretation over $\mathbb{N}$ is said to be *weakly (strictly) monotone* if the algebra $(\mathcal{N}, >_{\mathbb{N}}, \geqslant_{\mathbb{N}})$ is weakly (strictly) monotone. Similarly, we say that a polynomial interpretation over $\mathbb{N}$ is *(weakly) compatible* with a set of rewrite rules $\mathcal{R}$ if the algebra $(\mathcal{N}, >_{\mathbb{N}}, \geqslant_{\mathbb{N}})$ is (weakly) compatible with $\mathcal{R}$. Finally, a TRS $\mathcal{R}$ is *polynomially terminating over* $\mathbb{N}$ if there exists a polynomial interpretation over $\mathbb{N}$ that is both compatible with $\mathcal{R}$ and strictly monotone.

Note that the notion of polynomial termination over $\mathbb{N}$ as defined above does indeed make sense because any TRS that is polynomially terminating over $\mathbb{N}$ is indeed terminating due to Corollary 1.7. Moreover, note that the coefficients of the polynomials occurring in a polynomial interpretation over $\mathbb{N}$ are not restricted to $\mathbb{N}$. Some of them may be negative, even if strict monotonicity of all interpretation functions is required.

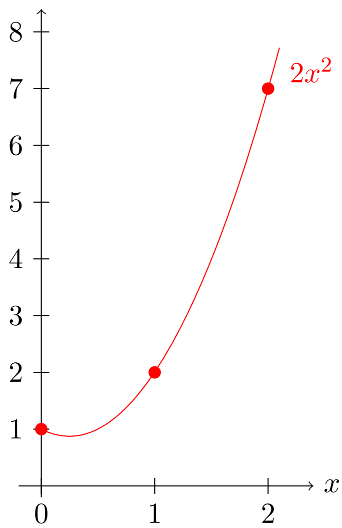

Figure 2.1: The polynomial function $x \mapsto 2x^2 - x + 1$.

Example 2.4. The univariate integer polynomial $p(x) = 2x^2 - x + 1 \in \mathbb{Z}[x]$ induces the polynomial function $f_{\mathbb{Z}} \colon \mathbb{Z} \to \mathbb{Z}$, $x \mapsto p(x)$, which is easily seen to be well-defined over $\mathbb{N}$ (cf. Figure 2.1). So its restriction $f_{\mathbb{N}} \colon \mathbb{N} \to \mathbb{N}$, $x \mapsto p(x)$ to the natural numbers is indeed permissible in a polynomial interpretation over $\mathbb{N}$, even in a strictly monotone one because $f_{\mathbb{N}}$ is strictly monotone (i.e., monotone with respect to $>_{\mathbb{N}}$). However, viewing $p(x)$ as a function of a non-negative real variable, monotonicity (with respect to the natural order on the real numbers) does not hold.

Next we present an example that illustrates the use of polynomial interpretations for proving a given TRS polynomially terminating and/or *incrementally* polynomially terminating. The formal definition of incremental polynomial termination will be given in the next subsection.

Example 2.5. Consider the TRS $\mathcal{R}$ consisting of the following rewrite rules:

$$\begin{aligned} \mathsf{f}(\mathsf{g}(x)) &\to \mathsf{g}(\mathsf{g}(\mathsf{f}(x))) \\ \mathsf{g}(\mathsf{s}(x)) &\to \mathsf{s}(\mathsf{s}(\mathsf{g}(x))) \end{aligned}$$

We claim that $\mathcal{R}$ is polynomially terminating over $\mathbb{N}$ by the following polynomial interpretation: $\mathsf{f}_{\mathbb{N}}(x) = x^2$, $\mathsf{g}_{\mathbb{N}}(x) = 3x + 5$ and $\mathsf{s}_{\mathbb{N}}(x) = x + 1$. Indeed, all these functions are polynomial functions that are well-defined over $\mathbb{N}$ as all coefficients are non-negative integers. In addition, all of them are monotone with respect to $>_{\mathbb{N}}$ because if the value of x is increased (by at least one), then the value of the functions also increases (by at least one). Finally, for compatibility with the rules of $\mathcal{R}$, the two inequalities

$$\begin{aligned} 9x^2 + 30x + 25 &>_{\mathbb{N}} 9x^2 + 20 \\ 3x + 8 &>_{\mathbb{N}} 3x + 7 \end{aligned}$$

must hold for all $x \in \mathbb{N}$, which is obviously true. Hence, $\mathcal{R}$ is polynomially terminating over $\mathbb{N}$. Besides, one can show that no interpretation using only linear polynomials of the form $ax + b$ can be used to establish polynomial termination of $\mathcal{R}$. However, the corresponding argument crucially relies on the fact that the interpretation must be (strictly) compatible with all rewrite rules. Indeed, using the incremental approach outlined in Section 1.3, which allows for weak compatibility with some (but not all) rules, termination of $\mathcal{R}$ can be established using linear polynomials. To this end, let us first consider the interpretation $\mathsf{f}_{\mathbb{N}}(x) = 3x$, $\mathsf{g}_{\mathbb{N}}(x) = x + 1$ and $\mathsf{s}_{\mathbb{N}}(x) = x$, which is easily seen to be a strictly monotone polynomial interpretation over $\mathbb{N}$ that is compatible with the first rule of $\mathcal{R}$ and weakly compatible with the second rule. Thus, after removing the first rule, termination of the second rule implies termination of $\mathcal{R}$. But termination of the second rule can easily be shown by the following interpretation: $\mathsf{g}_{\mathbb{N}}(x) = 3x$ and $\mathsf{s}_{\mathbb{N}}(x) = x + 1$. Hence, termination of $\mathcal{R}$ follows.

2.1.2 Polynomial Interpretations over the Rationals and Reals

Now if one wants to extend the notion of polynomial interpretations to the rational, real or real algebraic numbers, the main problem one is confronted with is the non-well-foundedness of these domains with respect to their natural order (even for non-negative numbers). In [30], and later in [45], this problem is overcome as follows. Let $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, fix some positive number $\delta \in D$, and define a new order $>_{D,\delta}$ on D as a replacement for the natural order $>_D$:

$$x >_{D,\delta} y \quad :\Longleftrightarrow \quad x - y \geqslant_D \delta \quad \text{for all } x, y \in D$$

Thus, $>_{D,\delta}$ is well-founded on subsets of D that are bounded from below. Therefore, any set D_m ($m \in D$) could in principle be used as the carrier for polynomial interpretations over D. However, it is well-known that one may *restrict* to D_0 without loss of generality, that is, without losing any power with respect to proving termination (cf. [16, 45, 70]). We shall further elaborate on this aspect at the end of this section.

Remark 2.6. Obviously, one can also define $>_{D,\delta}$ for $D = \mathbb{Z}$. Then, for $\delta = 1$, the order $>_{\mathbb{Z},1}$ corresponds to the natural order on the integers.

Definition 2.7. Let $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$. A *polynomial interpretation over D* for a signature $\mathcal{F}$ consists of a polynomial $f_D \in D[x_1, \ldots, x_n]$ for each n-ary function symbol $f \in \mathcal{F}$ and some positive number $\delta \in D$ such that f_D is well-defined over D_0.

As for polynomial interpretations over $\mathbb{N}$, the pair $\mathcal{D} = (D_0, \{f_D\}_{f \in \mathcal{F}})$ constitutes an $\mathcal{F}$-algebra over the carrier D_0 due to the well-definedness of all interpretation functions. Together with $>_{D_0,\delta}$ and $\geqslant_{D_0}$, the restrictions of $>_{D,\delta}$ and $\geqslant_D$ to D_0, we obtain an algebra $(\mathcal{D}, >_{D_0,\delta}, \geqslant_{D_0})$, where $>_{D_0,\delta}$ is well-founded (on D_0) and $>_{D_0,\delta} \cdot \geqslant_{D_0} \subseteq >_{D_0,\delta}$. Hence, if for each $f \in \mathcal{F}$, f_D is weakly (strictly) monotone, that is, monotone with respect to $\geqslant_{D_0}$ ($>_{D_0,\delta}$), then $(\mathcal{D}, >_{D_0,\delta}, \geqslant_{D_0})$ is a weakly (strictly) monotone $\mathcal{F}$-algebra. However, unlike for polynomial interpretations over $\mathbb{N}$, strict monotonicity of $(\mathcal{D}, >_{D_0,\delta}, \geqslant_{D_0})$ does not entail weak monotonicity as it can very well be the case that an interpretation function is monotone with respect to $>_{D_0,\delta}$ but not with respect to $\geqslant_{D_0}$. We shall see an example of such a function in Section 3.3 (cf. Lemma 3.19).

Definition 2.8. Let $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$. A polynomial interpretation over D is said to be *weakly (strictly) monotone* if the algebra $(\mathcal{D}, >_{D_0,\delta}, \geqslant_{D_0})$ is weakly (strictly) monotone. Similarly, we say that a polynomial interpretation over D is *(weakly) compatible* with a set of rewrite rules $\mathcal{R}$ if the algebra $(\mathcal{D}, >_{D_0,\delta}, \geqslant_{D_0})$ is (weakly) compatible with $\mathcal{R}$. Finally, a TRS $\mathcal{R}$ is *polynomially terminating over D* if there exists a polynomial interpretation over D that is both compatible with $\mathcal{R}$ and strictly monotone.

Again, the notion of polynomial termination over $\mathbb{Q}$ ($\mathbb{R}_{\mathsf{alg}}$, $\mathbb{R}$) as defined above is sensible due to Corollary 1.7. Next we formally define the notion of incremental polynomial termination, which is essentially based on the incremental termination approach for extended monotone algebras given in Section 1.3.

Definition 2.9. For $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$ and $n \in \mathbb{N}$, $n \geqslant 1$, a TRS $\mathcal{R}$ is said to be *polynomially terminating over D in n steps* if

- $n = 1$ and $\mathcal{R}$ is polynomially terminating over D, or if
- $n > 1$ and there exists a polynomial interpretation $\mathcal{P}$ over D and a non-empty subset $\mathcal{S} \subsetneq \mathcal{R}$ such that
  1. $\mathcal{P}$ is weakly and strictly monotone,
  2. $\mathcal{R} \subseteq \geqslant_{\mathcal{P}}$ and $\mathcal{S} \subseteq >_{\mathcal{P}}$, and
  3. $\mathcal{R} \setminus \mathcal{S}$ is polynomially terminating over D in $n - 1$ steps.

Furthermore, we call a TRS $\mathcal{R}$ *incrementally polynomially terminating over D* (or *polynomially terminating over D^**) if there exists some $n \in \mathbb{N}$, $n \geqslant 1$, such that $\mathcal{R}$ is polynomially terminating over D in n steps.

Note that the interpretation $\mathcal{P}$ in Definition 2.9 is an extended monotone algebra that establishes relative termination of $\mathcal{S}$ with respect to $\mathcal{R}$ according to Theorem 1.6. So every infinite rewrite sequence in $\mathcal{R}$ contains only finitely many $\mathcal{S}$-steps. That is, after a finite number of steps we are left with an infinite rewrite sequence in $\mathcal{R} \setminus \mathcal{S}$. Thus, termination of $\mathcal{R} \setminus \mathcal{S}$ implies termination of $\mathcal{R}$ (cf. also [20, Theorem 3]). This observation, together with an easy induction on the number of steps n, by which termination of $\mathcal{R} \setminus \mathcal{S}$ in item (3) follows, yields the following soundness result. In particular, note that the notions of polynomial termination over D and incremental polynomial termination over D coincide for singleton TRSs.

Lemma 2.10. *Let $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, and let $\mathcal{R}$ be a TRS. If $\mathcal{R}$ is incrementally polynomially terminating over D, then it is terminating.* □

In principle, one could take any set $\mathbb{N}_m$ (or even $\mathbb{Z}_m$) instead of $\mathbb{N}$ as the carrier for polynomial interpretations with integer coefficients. However, it is well-known [16, 70] that this does not have any effect on the power of the interpretations (with respect to proving termination). Thus, one may *restrict* to the carrier $\mathbb{N}$ without loss of generality. Likewise, for polynomial interpretations with coefficients from $\mathbb{Q}$ ($\mathbb{R}_{\mathsf{alg}}$, $\mathbb{R}$), one does not lose any power by choosing the carrier $\mathbb{Q}_0$ ($\mathbb{R}_{\mathsf{alg},0}$, $\mathbb{R}_0$), as was already observed in [45]. In fact, this is a direct consequence of the following more general result.

Let $\mathcal{F}$ be a signature, $D \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, and let $(\mathcal{B}, >_{D_m,\delta}, \geqslant_{D_m})$ be an $\mathcal{F}$-algebra over the carrier D_m with $m \in D$ (whose interpretation functions are not necessarily polynomials). Then one can always define an $\mathcal{F}$-algebra $(\mathcal{A}, >_{D_0,\delta}, \geqslant_{D_0})$ over the carrier D_0 such that $\mathcal{A}$ is isomorphic to $\mathcal{B}$ and induces the same (ordering) relations $\geqslant_{\mathcal{A}} = \geqslant_{\mathcal{B}}$ and $>_{\mathcal{A}} = >_{\mathcal{B}}$ on terms. In particular, for $D = \mathbb{Z}$ and $\delta = 1$, one can always define an isomorphic $\mathcal{F}$-algebra $(\mathcal{A}, >_{\mathbb{N}}, \geqslant_{\mathbb{N}})$ over the carrier $\mathbb{N}$ because $D_0 = \mathbb{N}$, $\geqslant_{D_0} = \geqslant_{\mathbb{N}}$ and $>_{D_0,\delta} = >_{\mathbb{N}}$ (cf. Remark 2.6). In order to prove this, let us consider the translation

$$\varphi \colon D_0 \to D_m,\ x \mapsto x + m$$

which is obviously bijective. In addition, it satisfies the following condition:

$$\forall\, x, y \in D_0 \quad x \geqslant_{D_0} y \iff \varphi(x) \geqslant_{D_m} \varphi(y) \tag{2.1}$$

(Here, $\geqslant_{D_m}$ denotes the restriction of $\geqslant_D$ to D_m.) This is due to the fact that $\geqslant_D$ is closed under addition and the observation that $x, y \in D_0$ if and only if $\varphi(x), \varphi(y) \in D_m$. Hence, φ is an order-isomorphism from $(D_0, \geqslant_{D_0})$ to $(D_m, \geqslant_{D_m})$, and therefore also from $(D_0, >_{D_0,\delta})$ to $(D_m, >_{D_m,\delta})$ as the orders $>_{D_0,\delta}$ and $>_{D_m,\delta}$ are defined in terms of $\geqslant_{D_0}$ and $\geqslant_{D_m}$, respectively. Moreover, φ naturally extends to a bijection $\widetilde{\varphi}$ from D_0^n to D_m^n, based on which we establish the following diagram:

$$\begin{tikzcd}
D_0^n \arrow[r, "\widetilde{\varphi}"] \arrow[d, "f_{D_0}"'] & D_m^n \arrow[d, "f_{D_m}"] \\
D_0 \arrow[r, "\varphi"] & D_m
\end{tikzcd}$$

Figure 2.2: A commutative diagram.

This diagram is commutative if $\varphi \circ f_{D_0} = f_{D_m} \circ \widetilde{\varphi}$, that is, if

$$f_{D_m} = \varphi \circ f_{D_0} \circ \widetilde{\varphi}^{-1} \text{ or equivalently } f_{D_0} = \varphi^{-1} \circ f_{D_m} \circ \widetilde{\varphi} \tag{2.2}$$

Thus, given f_{D_0}, defining $f_{D_m} = \varphi \circ f_{D_0} \circ \widetilde{\varphi}^{-1}$ makes the diagram commutative, and vice versa. Now let $\mathcal{A} = (D_0, \{f_{D_0}\}_{f\in\mathcal{F}})$ and $\mathcal{B} = (D_m, \{f_{D_m}\}_{f\in\mathcal{F}})$ be two $\mathcal{F}$-algebras. Then commutation of the diagram amounts to an isomorphism between the two algebras.

Lemma 2.11. *Consider the $\mathcal{F}$-algebras $(\mathcal{A}, >_{D_0,\delta}, \geqslant_{D_0})$ and $(\mathcal{B}, >_{D_m,\delta}, \geqslant_{D_m})$ and assume that for each n-ary function symbol $f \in \mathcal{F}$, the diagram in Figure 2.2 commutes. Then the following statements hold:*

1. *$\mathcal{A}$ and $\mathcal{B}$ are isomorphic as $\mathcal{F}$-algebras by the isomorphism φ,*
2. *$\varphi \circ [\alpha]_{\mathcal{A}}(\cdot) = [\varphi \circ \alpha]_{\mathcal{B}}(\cdot)$ for any variable assignment α, and*
3. *$\geqslant_{\mathcal{A}} = \geqslant_{\mathcal{B}}$ and $>_{\mathcal{A}} = >_{\mathcal{B}}$.*

Proof. Concerning the first item, we observe that the translation φ is a bijective mapping from the carrier of $\mathcal{A}$ to the carrier of $\mathcal{B}$, which is a homomorphism from $\mathcal{A}$ to $\mathcal{B}$ due to the assumption that the diagram in Figure 2.2 commutes, which means that for each n-ary function symbol $f \in \mathcal{F}$, $\varphi \circ f_{D_0} = f_{D_m} \circ \widetilde{\varphi}$.

Concerning the second item, we remark that $[\alpha]_{\mathcal{A}}(t) = \varphi^{-1}([\varphi\circ\alpha]_{\mathcal{B}}(t))$ follows by a straightforward induction on t using the commutation property (2.2).

The third item follows from the second item and the order-isomorphism property (2.1). The equality $>_{\mathcal{A}} = >_{\mathcal{B}}$ can be shown as follows:

$$
\begin{aligned}
s >_{\mathcal{A}} t \quad &\iff \quad \forall\alpha \;\; [\alpha]_{\mathcal{A}}(s) >_{D_0,\delta} [\alpha]_{\mathcal{A}}(t) \\
&\iff \quad \forall\alpha \;\; [\alpha]_{\mathcal{A}}(s) \geqslant_{D_0} [\alpha]_{\mathcal{A}}(t) + \delta \\
&\iff \quad \forall\alpha \;\; \varphi([\alpha]_{\mathcal{A}}(s)) \geqslant_{D_m} \varphi([\alpha]_{\mathcal{A}}(t) + \delta) \\
&\iff \quad \forall\alpha \;\; \varphi([\alpha]_{\mathcal{A}}(s)) \geqslant_{D_m} \varphi([\alpha]_{\mathcal{A}}(t)) + \delta \\
&\iff \quad \forall\alpha \;\; [\varphi\circ\alpha]_{\mathcal{B}}(s) \geqslant_{D_m} [\varphi\circ\alpha]_{\mathcal{B}}(t) + \delta \\
&\iff \quad \forall\alpha \;\; [\varphi\circ\alpha]_{\mathcal{B}}(s) >_{D_m,\delta} [\varphi\circ\alpha]_{\mathcal{B}}(t) \\
&\iff \quad \forall\beta \;\; [\beta]_{\mathcal{B}}(s) >_{D_m,\delta} [\beta]_{\mathcal{B}}(t) \\
&\iff \quad s >_{\mathcal{B}} t
\end{aligned}
$$

Here, α (β) is a variable assignment for $\mathcal{A}$ ($\mathcal{B}$), and we use the fact that there is a one-to-one correspondence between assignments for $\mathcal{A}$ and assignments for $\mathcal{B}$ mapping any assignment α for $\mathcal{A}$ to an assignment $\beta = \varphi \circ \alpha$ for $\mathcal{B}$. Finally, we remark that the equality $\geqslant_{\mathcal{A}} = \geqslant_{\mathcal{B}}$ follows by analogous reasoning after replacing $>_{D_0,\delta}$ by $\geqslant_{D_0}$. ☐

Thus, if $\mathcal{B} = (D_m, \{f_{D_m}\}_{f\in\mathcal{F}})$ is a polynomial interpretation over the carrier D_m, then we can always define a polynomial interpretation $\mathcal{A} = (D_0, \{f_{D_0}\}_{f\in\mathcal{F}})$ over the carrier D_0 by letting $f_{D_0} = \varphi^{-1} \circ f_{D_m} \circ \widetilde{\varphi}$ such that $\mathcal{A}$ is isomorphic to $\mathcal{B}$ and induces the same (ordering) relations $\geqslant_{\mathcal{A}} = \geqslant_{\mathcal{B}}$ and $>_{\mathcal{A}} = >_{\mathcal{B}}$ on terms. Note that f_{D_0} is a polynomial whenever f_{D_m} is a polynomial because polynomials are closed under translation.

Example 2.12. Consider the TRS $\mathcal{R}$ of Example 2.5. As shown previously, this system is polynomially terminating over $\mathbb{N}$ by the following interpretation:

$\mathsf{f}_{\mathbb{N}}(x) = x^2$, $\mathsf{g}_{\mathbb{N}}(x) = 3x+5$ and $\mathsf{s}_{\mathbb{N}}(x) = x+1$. As a consequence of Lemma 2.11, termination of $\mathcal{R}$ can also be established via the following polynomial interpretation over the carrier $\mathbb{N}_1$:

$$\mathsf{f}_{\mathbb{N}_1}(x) = \mathsf{f}_{\mathbb{N}}(x-1)+1 = x^2-2x+2 \qquad \mathsf{g}_{\mathbb{N}_1}(x) = \mathsf{g}_{\mathbb{N}}(x-1)+1 = 3x+3$$
$$\mathsf{s}_{\mathbb{N}_1}(x) = \mathsf{s}_{\mathbb{N}}(x-1)+1 = x+1$$

To this end, we first observe that all interpretation functions are indeed well-defined over $\mathbb{N}_1$. This is obvious for $\mathsf{g}_{\mathbb{N}_1}$ and $\mathsf{s}_{\mathbb{N}_1}$, but it also holds for $\mathsf{f}_{\mathbb{N}_1}$ because $\mathsf{f}_{\mathbb{N}_1}(x) = (x-1)^2+1$ has a global minimum of value one. In addition, all functions are monotone with respect to $>_{\mathbb{N}_1}$. In particular, monotonicity of $\mathsf{f}_{\mathbb{N}_1}$ can be shown as follows. If $x >_{\mathbb{N}_1} y$ (i.e., $x >_{\mathbb{N}} y$ for $x, y \in \mathbb{N}_1$), then $x-1 >_{\mathbb{N}} y-1$ and both operands are non-negative, which implies $(x-1)^2 >_{\mathbb{N}} (y-1)^2$ and $\mathsf{f}_{\mathbb{N}_1}(x) = (x-1)^2+1 >_{\mathbb{N}} (y-1)^2+1 = \mathsf{f}_{\mathbb{N}_1}(y)$. Hence, $\mathsf{f}_{\mathbb{N}_1}(x) >_{\mathbb{N}_1} \mathsf{f}_{\mathbb{N}_1}(y)$ due to well-definedness of $\mathsf{f}_{\mathbb{N}_1}$. Finally, compatibility with the rules of $\mathcal{R}$ amounts to the satisfaction of the two inequalities $30x >_{\mathbb{N}_1} 25$ and $3x+6 >_{\mathbb{N}_1} 3x+5$ for all $x \in \mathbb{N}_1$, both of which obviously hold. This implies termination due to Corollary 1.7. In the same way, termination of $\mathcal{R}$ can also be established via a polynomial interpretation over a carrier containing negative numbers, like $\mathbb{Z}_{-1}$, for example:

$$\mathsf{f}_{\mathbb{Z}_{-1}}(x) = \mathsf{f}_{\mathbb{N}}(x+1)-1 = x^2+2x \qquad \mathsf{g}_{\mathbb{Z}_{-1}}(x) = \mathsf{g}_{\mathbb{N}}(x+1)-1 = 3x+7$$
$$\mathsf{s}_{\mathbb{Z}_{-1}}(x) = \mathsf{s}_{\mathbb{N}}(x+1)-1 = x+1$$

We conclude this section with the formal definition of linear polynomial interpretations.

Definition 2.13. Let $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$. A *polynomial interpretation over* D is said to be *linear* if all its interpretation functions are linear polynomials, that is, $f_D(x_1, \dots, x_n) = a_n x_n + \cdots + a_1 x_1 + a_0$ for each n-ary function symbol f.

In the remainder of this thesis, we will sometimes use the term "polynomial interpretations with integer coefficients" as a synonym for polynomial interpretations over $\mathbb{N}$. Likewise, the term "polynomial interpretations with rational (real, real algebraic) coefficients" refers to polynomial interpretations over $\mathbb{Q}$ ($\mathbb{R}$, $\mathbb{R}_{\mathsf{alg}}$).

2.2 Basic Facts

In this section, we present some well-known facts about polynomial interpretations as well as a new result in connection with simple termination.

We start with some easy observations which will be used freely in the sequel. Let $\mathcal{P}$ be a polynomial interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$ and $\ell \to r$ a rewrite rule in the variables $x_1, \dots, x_m$. One easily verifies (e.g. by induction on ℓ) that the interpretation $[\alpha]_{\mathcal{P}}(\ell)$ with respect to a variable assignment α can be written as $P_\ell(\alpha(x_1), \dots, \alpha(x_m))$ for some m-variate polynomial P_ℓ (which does not depend on α). Likewise, one can associate a polynomial P_r with r such that $[\alpha]_{\mathcal{P}}(r) = P_r(\alpha(x_1), \dots, \alpha(x_m))$. Denoting the difference $P_\ell - P_r$

by $P_{\ell,r}$, we thus conclude that compatibility of $\mathcal{P}$ with the rule $\ell \to r$ holds if and only if $P_{\ell,r}(\alpha(x_1), \ldots, \alpha(x_m)) > 0$ for all assignments α, or equivalently, if $P_{\ell,r}(x_1, \ldots, x_m) > 0$ for all $x_1, \ldots, x_m \in D_0$. Similarly, $\mathcal{P}$ is weakly compatible with $\ell \to r$ if and only if $P_{\ell,r}(x_1, \ldots, x_m) \geqslant 0$ for all $x_1, \ldots, x_m \in D_0$.

We also note that compatibility with a set of rewrite rules implies weak compatibility with that set of rules. This is due to the fact that $>_{\mathbb{N}} \subseteq \geqslant_{\mathbb{N}}$ and $>_{D_0,\delta} \subseteq \geqslant_{D_0}$ for all $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$.

2.2.1 Decidability Issues

Next we recall some decidability results related to polynomial interpretations. Suppose $\mathcal{R}$ is a finite set of rewrite rules over some signature $\mathcal{F}$ and $\{f_{\mathbb{Z}}\}_{f \in \mathcal{F}}$ a collection of polynomials with integer coefficients. In order to verify whether these polynomials constitute a valid polynomial interpretation over $\mathbb{N}$ that is compatible with $\mathcal{R}$, we have to show

1. well-definedness (over $\mathbb{N}$) of all interpretation functions, and
2. for each rule $\ell \to r \in \mathcal{R}$, $[\alpha]_{\mathcal{N}}(\ell) >_{\mathbb{N}} [\alpha]_{\mathcal{N}}(r)$ for all assignments α.

Clearly, if the first condition holds, then $\mathcal{N} = (\mathbb{N}, \{f_{\mathbb{Z}}\}_{f \in \mathcal{F}})$ is indeed a valid polynomial interpretation over $\mathbb{N}$, which is compatible with $\mathcal{R}$ if the second condition is satisfied as well. Using the observations made at the beginning of this section, we conclude that the family of interpretations $\{f_{\mathbb{Z}}\}_{f \in \mathcal{F}}$ is a valid polynomial interpretation over $\mathbb{N}$ that is (weakly) compatible with $\mathcal{R}$ if the following conditions hold:

1. for each n-ary $f \in \mathcal{F}$, $f_{\mathbb{Z}}(x_1, \ldots, x_n) \geqslant 0$ for all $x_1, \ldots, x_n \in \mathbb{N}$, and
2. for each rule $\ell \to r \in \mathcal{R}$, $P_{\ell,r}(x_1, \ldots, x_m) \underset{(\geq)}{} 0$ for all $x_1, \ldots, x_m \in \mathbb{N}$.

Here, the variables $x_1, \ldots, x_m$ correspond to the ones occurring in $\ell \to r$. Besides, in connection with termination, the interpretation is required to be weakly monotone (if used in the DP framework) or strictly monotone (if used as a stand-alone termination method). The corresponding conditions for strict (weak) monotonicity are as follows: for each n-ary function symbol $f \in \mathcal{F}$, $f_{\mathbb{Z}}(x_1, \ldots, x_i, \ldots, x_n) \underset{(\geq)}{} f_{\mathbb{Z}}(x_1, \ldots, y, \ldots, x_n)$ for all $x_1, \ldots, x_n, y \in \mathbb{N}$ with $x_i \underset{(\geq)}{} y$ and all $i \in \{1, \ldots, n\}$.

Using the fact that $x >_{\mathbb{N}} y$ if and only if $x \geqslant_{\mathbb{N}} y + 1$, all of the conditions mentioned above, most notably the ones arising from compatibility with rewrite rules, can be phrased as instances of the problem of checking *positiveness of a polynomial* $p \in \mathbb{Z}[x_1, \ldots, x_n]$ on $\mathbb{N}^n$. However, this problem is known to be undecidable for $n > 1$ (by reduction from Hilbert's 10-th problem, cf. [70, Proposition 6.2.11]), as was already observed by Lankford [43]. Nevertheless, the following partial method for testing positiveness (and non-negativeness), which is due to [34], has proven to work sufficiently well in practice, that is, in the context of automated termination proofs (cf. [16, 34] for details). We shall refer to it as the *absolute positiveness* approach and adapt it as follows.

Definition 2.14. A polynomial is said to be *absolutely positive* (*absolutely non-negative*) if its constant coefficient is positive (non-negative) and all other coefficients are non-negative.

Lemma 2.15. *If a polynomial $p \in \mathbb{Z}[x_1, \dots, x_n]$ is absolutely positive (absolutely non-negative), then $p(x_1, \dots, x_n) \underset{(\geq)}{>} 0$ for all $x_1, \dots, x_n \in \mathbb{N}$.* $\square$

Another way of obtaining a decidable sufficient condition for positiveness of a polynomial $p \in \mathbb{Z}[x_1, \dots, x_n]$ on $\mathbb{N}^n$, which was already mentioned in [43], is to evaluate p in $\mathbb{R}$ rather than in $\mathbb{N}$ as positiveness on $\mathbb{N}^n$ follows from positiveness on $\mathbb{R}_0^n$. And the latter is decidable by Tarski's celebrated decision procedure for the first-order theory of real closed fields using the method of quantifier elimination [6, 14, 69], which implies the decidability of the first-order theory of the real numbers ($\mathbb{R}$ being a real closed field).

More importantly, from Tarski's result we immediately obtain the decidability of the problem considered at the beginning of this section for polynomials with real coefficients. That is to say that it is decidable whether a given collection of polynomials with real (algebraic) coefficients constitutes a valid polynomial interpretation over $\mathbb{R}$ ($\mathbb{R}_{\mathsf{alg}}$) that is compatible (resp. weakly compatible) with a given set of rewrite rules, and weak (resp. strict) monotonicity of the interpretation is decidable, too. (Note that both $\mathbb{R}$ and $\mathbb{R}_{\mathsf{alg}}$ are real closed fields.) While this result is of great theoretical importance, it is less relevant in practice due to the high computational complexity of the decision procedures for the theory of real closed fields (cf. [6], for example). In fact, as of yet, most automatic termination tools prefer the absolute positiveness approach also for polynomials with real (algebraic) coefficients. This is sound because Lemma 2.15 extends to this case.

Lemma 2.16. *If a polynomial $p \in \mathbb{R}[x_1, \dots, x_n]$ is absolutely positive (absolutely non-negative), then $p(x_1, \dots, x_n) \underset{(\geq)}{>} 0$ for all $x_1, \dots, x_n \in \mathbb{R}_0$.* $\square$

As far as polynomial interpretations over $\mathbb{Q}$ are concerned, we remark that the corresponding literature (e.g. [4, 45, 70]) does not mention any decidability results. It is known, though, that the full first-order theory of the rationals is undecidable [63, 65]. Despite this negative result, it turns out that decidability does indeed hold for polynomial interpretations over $\mathbb{Q}$. However, we shall postpone the proof of this fact until we have all the necessary ingredients for it at hand in Chapter 3.

2.2.2 Total Termination and Simple Termination

Next we relate the notion of polynomial termination to two other well-known notions of termination, namely, total termination and simple termination. Following [70], we define total termination as follows.

Definition 2.17. A TRS over a signature $\mathcal{F}$ is called *totally terminating* if there exists a compatible well-founded monotone $\mathcal{F}$-algebra $(\mathcal{A}, >, \geqslant)$ such that $>$ is a total order on its carrier.

For simple termination, we employ the following definition.

Definition 2.18. A TRS $\mathcal{R}$ over a signature $\mathcal{F}$ is called *simply terminating* if $\mathcal{R} \cup \mathcal{E}\mathrm{mb}(\mathcal{F})$ is terminating, where the TRS $\mathcal{E}\mathrm{mb}(\mathcal{F})$ consists of all the rules

$$f(x_1, \ldots, x_n) \to x_i$$

with $f \in \mathcal{F}$ an n-ary function symbol, $n \geqslant 1$, and $i \in \{1, \ldots, n\}$. The latter rules are called *embedding rules*.

Hence, compatibility of a strictly monotone $\mathcal{F}$-algebra $(\mathcal{A}, >, \geqslant)$, where $>$ is a total order, with some TRS $\mathcal{R}$ implies total termination of $\mathcal{R}$. Thus, polynomial termination over $\mathbb{N}$ implies total termination as $>_{\mathbb{N}}$ is a total order on $\mathbb{N}$. In contrast, for $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, the order $>_{D_0,\delta}$ is not total (independently of the value of δ). So if a TRS is polynomially terminating over $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ or $\mathbb{R}$, total termination cannot be concluded.

Polynomial termination over $\mathbb{N}$ also implies simple termination. This is due to the fact that total termination implies simple termination according to [70, Propositions 6.3.8 and 6.3.18]. (But note that the latter statement does not hold any more for infinite signatures if one employs the alternative definition of simple termination given in [50].)

Example 2.19. Consider the TRS $\mathcal{R}_0 = \{\mathsf{f}(\mathsf{a}) \to \mathsf{f}(\mathsf{g}(\mathsf{a}))\}$ over the finite signature $\mathcal{F} = \{\mathsf{a}, \mathsf{f}, \mathsf{g}\}$. It is easy to see that this system is not simply terminating because the TRS $\mathcal{R}_0 \cup \mathcal{E}\mathrm{mb}(\mathcal{F}) = \mathcal{R}_0 \cup \{\mathsf{f}(x) \to x, \mathsf{g}(x) \to x\}$ admits the infinite rewrite sequence

$$\mathsf{f}(\mathsf{a}) \to_{\mathcal{R}_0} \mathsf{f}(\mathsf{g}(\mathsf{a})) \to_{\mathcal{E}\mathrm{mb}(\mathcal{F})} \mathsf{f}(\mathsf{a}) \to_{\mathcal{R}_0} \mathsf{f}(\mathsf{g}(\mathsf{a})) \to \cdots$$

Hence, the TRS $\mathcal{R}_0$ cannot be polynomially terminating over $\mathbb{N}$.

Concerning simple termination in connection with polynomial interpretations over $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ and $\mathbb{R}$, we establish the following new result showing that polynomial termination over these domains does not imply simple termination, not even for finite signatures.

Lemma 2.20. *The TRS $\mathcal{R}_0$ of Example 2.19 is polynomially terminating over $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ and $\mathbb{R}$.*

Proof. The following interpretation establishes polynomial termination over $\mathbb{Q}$:

$$\delta = 1 \qquad \mathsf{a}_{\mathbb{Q}} = \tfrac{1}{2} \qquad \mathsf{f}_{\mathbb{Q}}(x) = 4x \qquad \mathsf{g}_{\mathbb{Q}}(x) = x^2$$

To this end, we first note that the compatibility constraint associated with the single rewrite rule gives rise to the inequality $2 >_{\mathbb{Q}_0,1} 1$, which holds by definition of $>_{\mathbb{Q}_0,1}$. Next we observe that all interpretation functions are well-defined over $\mathbb{Q}_0$ as all coefficients are non-negative. So it remains to show monotonicity with respect to $>_{\mathbb{Q}_0,1}$. Assume $x >_{\mathbb{Q}_0,1} y$ for $x, y \in \mathbb{Q}_0$. By definition of $>_{\mathbb{Q}_0,1}$, we have $x \geqslant_{\mathbb{Q}_0} y + 1$, and therefore $4x \geqslant_{\mathbb{Q}_0} 4y + 4 \geqslant_{\mathbb{Q}_0} 4y + 1$, which implies $\mathsf{f}_{\mathbb{Q}}(x) \geqslant_{\mathbb{Q}_0} \mathsf{f}_{\mathbb{Q}}(y) + 1$. Hence, $\mathsf{f}_{\mathbb{Q}}(x) >_{\mathbb{Q}_0,1} \mathsf{f}_{\mathbb{Q}}(y)$ due to well-definedness of $\mathsf{f}_{\mathbb{Q}}$. As to monotonicity of $\mathsf{g}_{\mathbb{Q}}$, we observe that $x >_{\mathbb{Q}_0,1} y$ implies both $x - y \geqslant_{\mathbb{Q}_0} 1$ and $x + y \geqslant_{\mathbb{Q}_0} 1$. So we obtain $(x + y)(x - y) = x^2 - y^2 \geqslant_{\mathbb{Q}_0} 1$, and therefore $\mathsf{g}_{\mathbb{Q}}(x) \geqslant_{\mathbb{Q}_0} \mathsf{g}_{\mathbb{Q}}(y) + 1$. Hence, $\mathsf{g}_{\mathbb{Q}}(x) >_{\mathbb{Q}_0,1} \mathsf{g}_{\mathbb{Q}}(y)$ due to well-definedness of $\mathsf{g}_{\mathbb{Q}}$. This shows polynomial termination over $\mathbb{Q}$. Polynomial termination over $\mathbb{R}_{\mathsf{alg}}$ and $\mathbb{R}$ follows by the same interpretation. □

Corollary 2.21. *For $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, the following statements hold:*

1. *Polynomial termination over D does not imply simple termination.*
2. *Polynomial termination over D does not imply total termination.*

Proof. The first claim follows from Lemma 2.20 because the TRS of Example 2.19 is not simply terminating. The second claim follows from the first one and the fact that total termination implies simple termination in case of finite signatures. □

2.2.3 Algebraic and Transcendental Numbers

According to Definition 2.7, the coefficients of the polynomials occurring in a polynomial interpretation over $\mathbb{R}$ can be arbitrary real numbers, including transcendental numbers, like π, for example. However, according to [47], it suffices to restrict to polynomials with real algebraic coefficients as interpretations of function symbols (without loss of generality). In this thesis, we extend this result by showing that polynomial interpretations over $\mathbb{R}$ are in fact equivalent to polynomial interpretations over $\mathbb{R}_{\mathsf{alg}}$ with respect to proving termination of TRSs (in the context of the DP framework as well as if used as a stand-alone termination method). That is, we show that it is also sufficient to restrict to the non-negative real algebraic numbers (instead of the entire set of non-negative real numbers) as the carrier for polynomial interpretations with real algebraic coefficients (cf. Section 4.2 of Chapter 4).

Chapter 3

Monotonicity Criteria

In order to prove termination by means of a polynomial interpretation, three conditions must be satisfied. First, all polynomials occurring in the interpretation must be well-defined over the respective carrier. Secondly, the interpretation is required to be weakly monotone (if applied in the context of the DP framework) or strictly monotone (if used as a stand-alone termination method). In other words, all interpretation functions have to be weakly monotone in the former case, whereas strict monotonicity is required in the latter case. Last but not least, given a termination problem in the form of a set of rewrite rules, the interpretation has to be compatible with some subset of its rules and weakly compatible with the remaining rules. Ideally, the process of proving termination should be fully automated. One does not want to search for termination proofs by hand, but one would rather have them generated automatically. Interestingly, as far as polynomial interpretations are concerned, automation was an issue right from the beginning [43]. Unlike then, nowadays they are considered to be well-suited for automation. Concerning polynomial interpretations over the natural numbers, the corresponding details can be found in [16], whereas the automation of polynomial interpretations over the rational and real numbers is described in [45]. In both approaches, termination tools are concerned with parametric polynomials whose coefficients (i.e., the parameters) are initially unknown and have to be instantiated suitably such that the resulting concrete polynomials satisfy the conditions mentioned above.

In the first part of this chapter, we consider polynomial interpretations over the natural numbers, putting our focus on monotonicity and well-definedness of parametric polynomials of low degree (e.g., linear, quadratic, etc.). The aim is to provide exact (that is, necessary and sufficient) constraints in terms of the abstract coefficients of the parametric polynomials such that monotonicity and well-definedness of the resulting concrete polynomials are guaranteed for every instantiation of the coefficients satisfying the constraints. In particular, our approach subsumes the one proposed in [16], which is currently used in many termination tools. In contrast to the latter, which is essentially based on the absolute positiveness criterion and therefore does not allow negative coefficients, negative numbers in certain coefficients can be handled without further ado in our approach.

In the second part of this chapter, we investigate monotonicity and well-definedness for polynomial interpretations over the rational and real numbers. Most notably, we provide complete characterizations of weak and strict monotonicity, the latter subsuming the approach proposed in [45]. We also present a

TRS, for which polynomial termination can only be established by our approach. In addition, we investigate the relationship between weak and strict monotonicity, showing that, unlike for polynomial interpretations over the natural numbers, strict monotonicity does not imply weak monotonicity. Moreover, we give necessary and sufficient criteria for monotonicity and well-definedness of parametric polynomials of low degree (also allowing negative numbers in certain coefficients).

Finally, in the third part of this chapter, we show that there are indeed TRSs that can be proved terminating by a polynomial interpretation with negative coefficients, but cannot be proved terminating by a polynomial interpretation where the coefficients of all polynomials are non-negative. Concerning previous work [21, 28] on negative coefficients in polynomial interpretations, we note that these approaches ensure well-definedness and (weak) monotonicity by extending polynomials with the "max"-operation. However, all our interpretation functions are (real) polynomials and our results do also apply to strict monotonicity. Hence, we do not consider these approaches in the sequel.

Last but not least, let us also mention that some of the criteria presented in this chapter have a strong impact on the theoretical part of this thesis, enabling us to derive (new) results, which could not be derived otherwise.

The chapter is organized as follows. In Section 3.1, we introduce some mathematical preliminaries. Then, in Section 3.2, we present our results on monotonicity and well-definedness of polynomial interpretations over the natural numbers, whereas Section 3.3 is dedicated to polynomial interpretations over the rational and real numbers. Negative coefficients are treated in Section 3.4, before we conclude in Section 3.5.

The results presented in Section 3.2 were originally published in the conference paper [58]. The material contained in Sections 3.3 and 3.4 is new.

3.1 Preliminaries

A sequence of real numbers $(x_k)_{k\in\mathbb{N}}$ *converges* to the *limit* x if for every real number $\varepsilon > 0$ there exists a natural number N such that the absolute distance $|x_k - x|$ is less than ε for all $k > N$; we denote this by $\lim_{k\to\infty} x_k = x$. As convergence in $\mathbb{R}^n$ is equivalent to componentwise convergence, we use the same notation also for limits of converging sequences of n-tuples of real numbers $(\vec{x}_k \in \mathbb{R}^n)_{k\in\mathbb{N}}$. A real function $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$, is *continuous* at $\xi \in D$ if for every converging sequence $(\vec{x}_k)_{k\in\mathbb{N}}$ in D with $\lim_{k\to\infty} \vec{x}_k = \xi$ it holds that $\lim_{k\to\infty} f(\vec{x}_k) = f(\xi)$. The function f is *continuous on* D if it is continuous at all $\xi \in D$. Polynomial functions are well-known to be continuous on $\mathbb{R}^n$. Finally, as $\mathbb{Q}$ is dense in $\mathbb{R}$, every real number is a rational number or the limit of a converging sequence of rational numbers.

A *quadratic equation* is a polynomial equation of the form $ax^2 + bx + c = 0$, where x is an indeterminate, and the coefficients a, b and c are real constants with $a \neq 0$. The values of x satisfying such an equation are called *roots* of the equation. By the fundamental theorem of algebra, every quadratic equation has exactly two (not necessarily distinct) roots. They are given by the *quadratic* *formula*:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

In this formula, the expression $D := b^2 - 4ac$ underneath the square root sign is of central importance because it determines the nature of the roots; it is therefore called the *discriminant* of a quadratic equation. Since all coefficients are assumed to be real numbers, one of the following three cases applies:

1. If D is positive, there are two distinct roots, both of which are real numbers.
2. If D is zero, there is exactly one real root, called a *double root*.
3. If D is negative, there are no real roots. Both roots are complex numbers.

Lemma 3.1. *Let $f(x) = ax^2 + bx + c$ with $a, b, c \in \mathbb{R}$. Then $f(x) \geqslant 0$ holds for all $x \in \mathbb{R}_0$ if and only if $a \geqslant 0 \wedge c \geqslant 0 \wedge (b \geqslant 0 \vee 4ac - b^2 \geqslant 0)$.*

Proof. First, we observe that $c \geqslant 0$ is necessary as $f(0) = c$. Moreover, $a \geqslant 0$ is also necessary because otherwise the leading coefficient of the quadratic polynomial $ax^2 + bx + c$, which determines the behaviour for large values of x, is negative. Next we consider the case $a = 0$ explicitly. Then the condition given in the lemma simplifies to $c \geqslant 0 \wedge (b \geqslant 0 \vee b^2 \leqslant 0)$, or equivalently, $c \geqslant 0 \wedge b \geqslant 0$, which is obviously necessary and sufficient for $f(x) = bx + c$ to be non-negative whenever x is non-negative. On the other hand, if $a > 0$, we proceed by case distinction on b. If $b \geqslant 0$, then all three coefficients are non-negative, so the claim holds. In case b is negative, we write

$$f(x) = a\left(\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + \frac{c}{a}\right) = a\left(\left(x + \frac{b}{2a}\right)^2 - \frac{D}{4a^2}\right)$$

where $D := b^2 - 4ac$ is the discriminant of the quadratic equation $ax^2 + bx + c = 0$. From this representation we infer that f has a global minimum at $x_{min} = -b/(2a)$ of value $f(x_{min}) = -D/(4a)$. As b is negative, x_{min} is positive, such that $f(x) \geqslant 0$ holds for all $x \in \mathbb{R}_0$ if and only if $f(x_{min}) = -D/(4a) \geqslant 0$, or equivalently, $4ac - b^2 \geqslant 0$. □

We end this section with the mean value theorem of differential calculus.

Theorem 3.2 ([5, Theorem 3.1])**.** *If a function $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}$, is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), then there exists $\xi \in (a, b)$ such that*

$$f'(\xi) = \frac{\partial f}{\partial x}(\xi) = \frac{f(b) - f(a)}{b - a}$$

□

3.2 Polynomial Interpretations over the Natural Numbers

According to the definition given in the previous chapter, a polynomial interpretation over $\mathbb{N}$ for a signature $\mathcal{F}$ is essentially an $\mathcal{F}$-algebra $\mathcal{N} = (\mathbb{N}, \{f_{\mathbb{N}}\}_{f \in \mathcal{F}})$

over the carrier $\mathbb{N}$ that associates each function symbol with a polynomial with integer coefficients. Depending on whether all interpretation functions are weakly or strictly monotone, the triple $(\mathcal{N}, >_{\mathbb{N}}, \geqslant_{\mathbb{N}})$ is either a weakly monotone $\mathcal{F}$-algebra or a strictly monotone $\mathcal{F}$-algebra (or both). As some form of monotonicity is always required in the context of termination analysis, we conclude that an n-ary polynomial function $f_{\mathbb{N}}$ used in a polynomial interpretation over $\mathbb{N}$ is an element of the polynomial ring $\mathbb{Z}[x_1, \ldots, x_n]$ that must satisfy:

1. well-definedness over $\mathbb{N}$: $f_{\mathbb{N}}(x_1, \ldots, x_n) \geqslant 0$ for all $x_1, \ldots, x_n \in \mathbb{N}$, and
2. strict (weak) monotonicity: $f_{\mathbb{N}}(x_1, \ldots, x_i, \ldots, x_n) \mathrel{{}_{(}\!\geq\!{}_{)}} f_{\mathbb{N}}(x_1, \ldots, y, \ldots, x_n)$ for all $x_1, \ldots, x_n, y \in \mathbb{N}$ with $x_i \mathrel{{}_{(}\!\geq\!{}_{)}} y$ and all $i \in \{1, \ldots, n\}$.

Alas, as mentioned in Section 2.2, both of these properties are in general undecidable.

Termination tools face the following problem. They deal with parametric polynomials (i.e., polynomials whose coefficients are unknowns; e.g., ax^2+bx+c), and their task is to find suitable integer numbers for the unknown coefficients such that the resulting concrete polynomials satisfy both of the above properties. The solution proposed in [16] is to restrict the search space for the unknown coefficients to the non-negative integers (i.e., applying the absolute positiveness criterion) because then well-definedness and weak monotonicity are obtained for free. To obtain strict monotonicity in the i-th argument of a polynomial function $f_{\mathbb{N}}(\ldots, x_i, \ldots)$, at least one of the terms $(c_k x_i^k)_{k>0}$ must have a positive coefficient $c_k > 0$ (or equivalently, $c_k \geqslant 1$).

Obviously, this approach is easy to implement and works quite well in practice. However, it is not optimal in the sense that it excludes certain polynomials, like $2x^2 - x + 1$ (cf. Figure 2.1), which might be useful to prove termination of certain TRSs. So how can we do better? To this end, let us observe that in general termination tools only use restricted forms of polynomials to interpret function symbols. There are restrictions concerning the degree of the polynomials (linear, quadratic, etc.) and sometimes also restrictions that disallow certain kinds of monomials. Now the idea is as follows. Despite the fact that well-definedness and monotonicity are undecidable in general, it might be the case, and indeed it actually is the case, that they are decidable for the restricted forms of polynomials used in practice.

Remark 3.3. Checking (weak) compatibility of a polynomial interpretation over $\mathbb{N}$ with a rewrite rule $\ell \to r$ means showing that the rule gives rise to a (weak) decrease; i.e., $P_\ell - P_r \mathrel{{}_{(}\!\geq\!{}_{)}} 0$. In $\mathbb{N}$, both cases reduce to checking non-negativeness of integer polynomials because $x > y$ holds if and only if $x \geqslant y + 1$ holds. Since well-definedness over $\mathbb{N}$ of an integer polynomial is equivalent to non-negativeness on $\mathbb{N}$, any method that ensures non-negativeness of parametric polynomials can also be used for checking compatibility. However, we remark that the method presented in this chapter is not ideally suited for this purpose as it also enforces monotonicity (weak or strict), which is irrelevant for compatibility.

In the sequel, we analyze parametric polynomials whose only restriction is a bound on the degree. We will first treat linear parametric polynomials. While this does not yield new results or insights, it is instructive to demonstrate our approach in a simple setting. This is followed by an analysis of quadratic and finally also cubic parametric polynomials, both of which yield new results. The following lemmas will be helpful in this analysis. The first one gives a more succinct characterization of monotonicity (which is well-known), whereas the second one relates monotonicity and well-definedness. In these lemmata, as well as in the entire section, well-definedness means well-definedness over $\mathbb{N}$, and weak (strict) monotonicity refers to monotonicity with respect to $\geqslant_{\mathbb{N}}$ ($>_{\mathbb{N}}$). In particular, monotonicity is meant with respect to all arguments (cf. Definitions 1.2 and 1.4).

Lemma 3.4. *A (not necessarily polynomial) function $f_{\mathbb{N}} \colon \mathbb{N}^n \to \mathbb{N}$ is strictly (weakly) monotone if and only if*

$$f_{\mathbb{N}}(x_1, \ldots, x_i + 1, \ldots, x_n) \underset{(=)}{>} f_{\mathbb{N}}(x_1, \ldots, x_i, \ldots, x_n)$$

for all $x_1, \ldots, x_n \in \mathbb{N}$ and all $i \in \{1, \ldots, n\}$. □

Lemma 3.5. *Let $f \colon \mathbb{Z}^n \to \mathbb{Z}$ be the polynomial function associated with a polynomial in $\mathbb{Z}[x_1, \ldots, x_n]$, and let $f_{\mathbb{N}} \colon \mathbb{N}^n \to \mathbb{Z}$ denote its restriction to $\mathbb{N}^n$. Then $f_{\mathbb{N}}$ is strictly (weakly) monotone and well-defined if and only if*

1. *$f_{\mathbb{N}}(0, \ldots, 0) \geqslant 0$, and*
2. *$f_{\mathbb{N}}(x_1, \ldots, x_i + 1, \ldots, x_n) \underset{(=)}{>} f_{\mathbb{N}}(x_1, \ldots, x_i, \ldots, x_n)$*

for all $x_1, \ldots, x_n \in \mathbb{N}$ and all $i \in \{1, \ldots, n\}$.

Proof. Obvious. □

3.2.1 Linear Parametric Polynomials

In this section, we consider the generic linear parametric polynomial function $f_{\mathbb{N}}(x_1, \ldots, x_n) = a_n x_n + a_{n-1} x_{n-1} + \cdots + a_1 x_1 + a_0$ and derive constraints on the coefficients a_i that guarantee monotonicity and well-definedness.

Lemma 3.6. *The function $f_{\mathbb{N}}(x_1, \ldots, x_n) = a_n x_n + \cdots + a_1 x_1 + a_0$ with $a_i \in \mathbb{Z}$ for $i = 0, 1, \ldots, n$ is strictly (weakly) monotone and well-defined if and only if $a_0 \geqslant 0$ and $a_i \underset{(=)}{>} 0$ for all $i \in \{1, \ldots, n\}$.*

Proof. Easy consequence of Lemma 3.5. □

Remark 3.7. Note that all coefficients must be non-negative and that the constraints on the coefficients are exactly the ones one would obtain by the absolute positiveness approach of [16]. Furthermore, these constraints are optimal in the sense that they are both necessary and sufficient for monotonicity and well-definedness.

3.2.2 Quadratic Parametric Polynomials

Next we consider the generic quadratic parametric polynomial function

$$f_{\mathbb{N}}(x_1, \ldots, x_n) = a_0 + \sum_{j=1}^{n} a_j x_j + \sum_{1 \leqslant j \leqslant k \leqslant n} a_{jk} x_j x_k \ \in \mathbb{Z}[x_1, \ldots, x_n] \tag{3.1}$$

Lemma 3.8. *The function $f_{\mathbb{N}}$ is strictly (weakly) monotone and well-defined if and only if $a_0 \geqslant 0$, $a_{jk} \geqslant 0$ and $a_j + a_{jj} \underset{(=)}{\geq} 0$ for all $1 \leqslant j \leqslant k \leqslant n$.*

Proof. By Lemma 3.5, this lemma holds if and only if $f_{\mathbb{N}}(0, \ldots, 0) \geqslant 0$ and $f_{\mathbb{N}}(x_1, \ldots, x_i + 1, \ldots, x_n) \underset{(=)}{\geq} f_{\mathbb{N}}(x_1, \ldots, x_i, \ldots, x_n)$ for all $x_1, \ldots, x_n \in \mathbb{N}$ and all $i \in \{1, \ldots, n\}$. Clearly, $f_{\mathbb{N}}(0, \ldots, 0) \geqslant 0$ is equivalent to $a_0 \geqslant 0$, and the monotonicity condition $f_{\mathbb{N}}(x_1, \ldots, x_i + 1, \ldots, x_n) > f_{\mathbb{N}}(x_1, \ldots, x_i, \ldots, x_n)$ yields

$$\begin{aligned} a_i(x_i + 1) + a_{ii}(x_i + 1)^2 + \sum_{i<k\leqslant n} a_{ik}(x_i + 1)x_k + \sum_{1\leqslant j<i} a_{ji}x_j(x_i + 1) \\ > a_i x_i + a_{ii} x_i^2 + \sum_{i<k\leqslant n} a_{ik} x_i x_k + \sum_{1\leqslant j<i} a_{ji} x_j x_i \end{aligned}$$

which simplifies to

$$a_i + a_{ii} + 2a_{ii}x_i + \sum_{i<k\leqslant n} a_{ik} x_k + \sum_{1\leqslant j<i} a_{ji} x_j > 0$$

This is a linear inequality in $x_1, \ldots, x_n$ that holds for all $x_1, \ldots, x_n \in \mathbb{N}$ if and only if $a_i + a_{ii} > 0$ and all other coefficients are non-negative. Taking the quantification over i into account, this proves the claim for strict monotonicity; for weak monotonicity, we just have to replace $>$ by $\geqslant$ in the above calculation. □

Corollary 3.9. *The function $f_{\mathbb{N}}(x) = ax^2 + bx + c$ with $a, b, c \in \mathbb{Z}$ is strictly (weakly) monotone and well-defined if and only if $a \geqslant 0$, $c \geqslant 0$ and $a + b \underset{(=)}{\geq} 0$.* □

Hence, in a quadratic parametric polynomial all coefficients must be non-negative except the coefficients associated with the linear monomials. They can be negative; for example, the polynomial $2x^2 - x + 1$ satisfies the constraints of Corollary 3.9, which shows that it is both well-defined and strictly monotone.

Remark 3.10. Not only does our approach improve upon the absolute positiveness approach of [16] for quadratic parametric polynomials, but the constraints derived from it are even optimal, that is, necessary and sufficient for monotonicity and well-definedness.

Example 3.11. The polynomial function

$$f_{\mathbb{N}}(x_1, x_2) = 2x_1^2 + 3x_2^2 + x_1 x_2 - x_1 - 2x_2 + 1$$

is both well-defined and strictly monotone according to Lemma 3.8. Yet we can also infer this result in a more modular and probably more intuitive way from Corollary 3.9. To this end, let $f_{\mathbb{N}}(x_1, x_2) = g_1(x_1) + g_2(x_2) + x_1 x_2 + 1$, where

$g_1(x_1) = 2x_1^2 - x_1$ and $g_2(x_2) = 3x_2^2 - 2x_2$. Clearly, by Corollary 3.9, $g_1(x_1)$ and $g_2(x_2)$ are both well-defined and strictly monotone. The same holds for their sum, $g_1(x_1) + g_2(x_2)$, because $g_1(x_1)$ and $g_2(x_2)$ do not share variables. From this we conclude that $f_{\mathbb{N}}$ is then also well-defined and strictly monotone by observing that the addition of terms with non-negative coefficients (in this case: $x_1 x_2$ and 1) is not harmful.

Another thing that is noteworthy about the previous lemma is that it subsumes the result of Lemma 3.6. That is to say, if we set the coefficients a_{jk} of all quadratic monomials in (3.1) to zero, thereby obtaining the linear parametric polynomial function $f'_{\mathbb{N}}(x_1, \ldots, x_n) = a_0 + \sum_{j=1}^{n} a_j x_j$, then the constraints generated by Lemma 3.8 are in fact the ones Lemma 3.6 would produce when applied to $f'_{\mathbb{N}}$. In theory, this means that if we want to prove termination of some TRS, then we do not have to specify a priori whether to interpret a function symbol by a linear or a quadratic parametric polynomial function; we can always go for a quadratic interpretation, and it is solely determined by the constraint solving process (i.e., the process that assigns suitable integers to the abstract coefficients such that all constraints are satisfied) whether the resulting concrete polynomial function is linear or quadratic. In practice, however, this approach has an important drawback; that is, it increases both the number of abstract coefficients and the number of constraints involving these coefficients, which is detrimental to the performance of the constraint solving process.

3.2.3 Cubic Parametric Polynomials

Next we apply our approach to cubic parametric polynomials. First, we consider the univariate polynomial function $f_{\mathbb{N}}(x) = ax^3 + bx^2 + cx + d \in \mathbb{Z}[x]$, for which the monotonicity condition $f_{\mathbb{N}}(x+1) \underset{(}{\geq}_{)} f_{\mathbb{N}}(x)$ for all $x \in \mathbb{N}$ simplifies to

$$\forall x \in \mathbb{N} \quad 3ax^2 + (3a+2b)x + (a+b+c) \underset{(}{\geq}_{)} 0 \tag{3.2}$$

In the interesting case, where $a \neq 0$, the polynomial

$$P(x) = 3ax^2 + (3a+2b)x + (a+b+c)$$

is a quadratic polynomial in x whose geometric representation is a parabola in 2-dimensional space, which has a global minimum at $x_{min} = -(3a+2b)/(6a)$ because the leading coefficient of P must not be negative in order for (3.2) to hold. Since a is involved in the leading coefficient of P, we conclude that a must be positive. Next we focus on strict monotonicity, that is, the solution of the inequality

$$\forall x \in \mathbb{N} \quad 3ax^2 + (3a+2b)x + (a+b+c) > 0 \tag{3.3}$$

Now this inequality holds if and only if either $x_{min} < 0$ and $P(0) > 0$ or $x_{min} \geqslant 0$ and both $P(\lfloor x_{min} \rfloor) > 0$ and $P(\lceil x_{min} \rceil) > 0$. However, these constraints use the floor and ceiling functions, but we would rather have a set of polynomial constraints in a, b and c (which can easily be encoded in SAT or SMT). It is possible, though, to eliminate the floor and ceiling functions from the above constraints, but only at the expense of introducing new variables; e.g., $\lfloor x_{min} \rfloor = n$

for some $n \in \mathbb{Z}$ if and only if $n \leqslant x_{min} < n+1$. Thus, one obtains a set of polynomial constraints in a, b, c and the additional variables. But one can also avoid the introduction of new variables with the following approach. For this purpose, let us examine the roots of P and distinguish two possible cases:

Case 1 P has no roots in $\mathbb{R}$ (both roots are complex numbers),

Case 2 both roots of P are real numbers.

In the first case, (3.3) trivially holds. Moreover, this case is completely characterized by the discriminant of P being negative, i.e., $4b^2 - 3a^2 - 12ac < 0$. In the other case, when both roots r_1 and r_2 are real numbers, the discriminant is non-negative and (3.3) holds if and only if the closed interval $[r_1, r_2]$ does not contain a natural number, i.e., $[r_1, r_2] \cap \mathbb{N} = \emptyset$. While this condition can be fully characterized with the help of the floor and/or ceiling functions, we can also obtain a polynomial characterization as follows. We require the larger of the two roots, that is, r_2, to be negative because then (3.3) is guaranteed to hold. This observation leads to the constraints

$$4b^2 - 3a^2 - 12ac \geqslant 0 \quad \text{and} \quad r_2 = \frac{-(3a+2b) + \sqrt{4b^2 - 3a^2 - 12ac}}{6a} < 0$$

which can be simplified to

$$4b^2 - 3a^2 - 12ac \geqslant 0 \tag{3.4}$$

$$\sqrt{4b^2 - 3a^2 - 12ac} < 3a + 2b \tag{3.5}$$

Due to (3.4), (3.5) holds if and only if $4b^2 - 3a^2 - 12ac < (3a+2b)^2$ and $3a+2b \geqslant 0$, which simplifies to $a+b+c > 0$ and $3a+2b \geqslant 0$. Putting everything together, we conclude from Lemma 3.5 that the function $f_{\mathbb{N}}(x) = ax^3 + bx^2 + cx + d$ is strictly monotone and well-defined if

$$a \geqslant 0 \wedge d \geqslant 0 \wedge (4b^2 - 3a^2 - 12ac < 0 \vee (a+b+c > 0 \wedge 3a+2b \geqslant 0))$$

Observing that $f_{\mathbb{N}}(1) - f_{\mathbb{N}}(0) = a+b+c > 0$ is necessary for strict monotonicity, we arrive at the following equivalent formulation.

Lemma 3.12. *The function $f_{\mathbb{N}}(x) = ax^3 + bx^2 + cx + d \in \mathbb{Z}[x]$ is strictly monotone and well-defined if $a \geqslant 0 \wedge d \geqslant 0 \wedge a+b+c > 0 \wedge (D < 0 \vee 3a+2b \geqslant 0)$, where $D = 4b^2 - 3a^2 - 12ac$.* □

Note that these constraints are only sufficient for monotonicity and well-definedness, they are not necessary. However, they are very close to necessary constraints, as will be explained below.

Weak monotonicity of $ax^3 + bx^2 + cx + d$ is obtained by similar reasoning. The only difference is that in case 2 we differentiate between distinct real roots $r_1 \neq r_2$ and a double root $r_1 = r_2$. In the latter case, which is characterized algebraically by the discriminant of P being zero, (3.2) holds unconditionally (as $P(x) = 3a(x - r_1)^2 \geqslant 0$), whereas in the former case, where the discriminant of P is positive, it suffices to require the larger of the two roots to be negative or zero.

Lemma 3.13. *The function* $f_{\mathbb{N}}(x) = ax^3 + bx^2 + cx + d \in \mathbb{Z}[x]$ *is weakly monotone and well-defined if* $a \geqslant 0 \wedge d \geqslant 0 \wedge a+b+c \geqslant 0 \wedge (D \leqslant 0 \vee 3a+2b \geqslant 0)$, *where* $D = 4b^2 - 3a^2 - 12ac$. □

In case $a = 0$, i.e., $f_{\mathbb{N}}(x) = bx^2 + cx + d$, Lemma 3.12 yields exactly the same constraints as Corollary 3.9, that is, necessary and sufficient constraints. One possible interpretation of this fact is that the simplification we made on our way to Lemma 3.12 did not cast away anything essential. Indeed, that is the case. To this end, we recall that Lemma 3.12 covers the case where the polynomial P has no real roots as well as the case of two negative real roots. The only case (where (3.3) holds) that is not covered is when both roots r_1 and r_2 are positive and $[r_1, r_2] \cap \mathbb{N} = \emptyset$; e.g., the polynomial $2x^3 - 6x^2 + 5x$ is both strictly monotone and well-defined, but it does not satisfy the constraints of Lemma 3.12; in particular, the polynomial $P(x) = 6x^2 - 6x + 1$ has two real roots $r_1, r_2 \in (0, 1)$. However, it turns out that this case is very rare; for example, empirical investigations reveal that in the set of polynomials

$$\{3ax^2 + (3a + 2b)x + (a + b + c) \mid 1 \leqslant a \leqslant 7, -15 \leqslant b, c \leqslant 15 \ (a, b, c \in \mathbb{Z})\}$$

3937 out of a total of 6727 polynomials satisfy (3.3), but only 25 of them are of this special kind. In other words, the constraints of Lemma 3.12 comprise 3912 out of 3937, hence almost all, polynomials; and this is way more than the 1792 ($= 7 \times 16 \times 16$) polynomials that the absolute positiveness approach, where a, b and c are restricted to the non-negative integers, can handle. The following table summarizes all our experiments with varying ranges for a, b and c:

a	b, c	Lemma 3.12
$[1, 7]$	$[-15, 15]$	3912 of 3937
$[1, 7]$	$[-31, 31]$	14055 of 14133
$[1, 15]$	$[-31, 31]$	34718 of 34980

By design, our approach covers two out of the three possible scenarios mentioned above. But which of these scenarios can the absolute positiveness approach deal with? Just like our method, it fails on all instances of the scenario where the polynomial $P(x) = 3ax^2 + (3a + 2b)x + (a + b + c)$ has two positive roots r_1 and r_2, which gives rise to the factorization $P(x) = k(x - r_1)(x - r_2)$ with $k > 0$. This expression is equivalent to $kx^2 - k(r_1 + r_2)x + kr_1r_2$, the linear coefficient of which should be equal to $3a + 2b$. Yet this is not possible if a and b are restricted to the non-negative integers because $-k(r_1 + r_2)$ is a negative number. Concerning the two remaining scenarios, the absolute positiveness approach can handle only some instances of the respective scenarios while failing at others. We present one failing example for either scenario:

- If $a = 1$, $b = -1$ and $c = 1$, then $P(x) = 3x^2 + x + 1$, which has no real roots. Clearly, $P(x)$ is positive for all $x \in \mathbb{N}$. However, the absolute positiveness approach fails because b is negative.
- If $a = 3$, $b = -1$ and $c = -1$, then $P(x) = 9x^2 + 7x + 1$, both roots of which are negative real numbers. Clearly, $P(x)$ is positive for all $x \in \mathbb{N}$, but the absolute positiveness approach fails because b and c are negative.

Generalization to Multivariate Cubic Parametric Polynomials

In this subsection, we elaborate on the question how to generalize the results of Lemmata 3.12 and 3.13 to the multivariate case. In general, this is always possible by a very simple approach that we already introduced in Example 3.11. For this purpose, let $f_{\mathbb{N}}(x_1, \ldots, x_n)$ denote the n-variate generic cubic parametric polynomial function, and let us note that we can write it as

$$f_{\mathbb{N}}(x_1, \ldots, x_n) = \sum_{j=1}^{n} g_j(x_j) + r(x_1, \ldots, x_n) \tag{3.6}$$

where $g_j(x_j)$ denotes the univariate generic cubic parametric polynomial function in x_j without constant term and $r(x_1, \ldots, x_n)$ contains all the remaining terms. Now let us assume that all the $g_j(x_j)$ are both strictly monotone and well-defined. Then the same is true of their sum $\sum_{j=1}^{n} g_j(x_j)$ because of variable-disjointness. But when is this also true of $f_{\mathbb{N}}$? By Lemma 3.5, $f_{\mathbb{N}}$ is strictly monotone and well-defined if and only if $\sum_{j=1}^{n} g_j(0) + r(0, \ldots, 0) \geqslant 0$ and

$$g_i(x_i + 1) - g_i(x_i) + r(x_1, \ldots, x_i + 1, \ldots, x_n) - r(x_1, \ldots, x_i, \ldots, x_n) > 0$$

for all $x_1, \ldots, x_n \in \mathbb{N}$ and all $i \in \{1, \ldots, n\}$. By assumption, we have $g_i(0) \geqslant 0$ and $g_i(x_i + 1) - g_i(x_i) > 0$ for all $x_i \in \mathbb{N}$ and all $i \in \{1, \ldots, n\}$. So the above constraint is guaranteed to hold if $r(0, \ldots, 0) \geqslant 0$ and

$$r(x_1, \ldots, x_i + 1, \ldots, x_n) - r(x_1, \ldots, x_i, \ldots, x_n) \geqslant 0$$

for all $x_1, \ldots, x_n \in \mathbb{N}$ and all $i \in \{1, \ldots, n\}$, that is, if r is weakly monotone and well-defined according to Lemma 3.5. Hence, the n-variate generic cubic parametric polynomial function $f_{\mathbb{N}}(x_1, \ldots, x_n)$ is strictly monotone and well-defined if

- all the $g_j(x_j)$ satisfy the constraints of Lemma 3.12, and
- $r(x_1, \ldots, x_n)$ is weakly monotone and well-defined.

In the same way, but using Lemma 3.13 instead of Lemma 3.12, we obtain weak monotonicity and well-definedness of $f_{\mathbb{N}}$. Concerning weak monotonicity and well-definedness of $r(x_1, \ldots, x_n)$, it is obviously sufficient to restrict all its coefficients to be non-negative. However, using the method proposed in this section, we can even do better. We demonstrate this for $n = 2$ in the example below and remark that the corresponding approach readily generalizes to arbitrarily many variables.

Example 3.14. For the bivariate generic cubic parametric polynomial function

$$f_{\mathbb{N}}(x_1, x_2) = ax_1^3 + bx_1^2x_2 + cx_1x_2^2 + dx_2^3 + ex_1^2 + fx_1x_2 + gx_2^2 + hx_1 + ix_2 + j$$

we write $f_{\mathbb{N}}(x_1, x_2) = g_1(x_1) + g_2(x_2) + r(x_1, x_2)$, where

$$\begin{aligned} g_1(x_1) &= ax_1^3 + ex_1^2 + hx_1 \\ g_2(x_2) &= dx_2^3 + gx_2^2 + ix_2 \\ r(x_1, x_2) &= bx_1^2x_2 + cx_1x_2^2 + fx_1x_2 + j \end{aligned}$$

Then this function is both strictly monotone and well-defined if $ax_1^3 + ex_1^2 + hx_1$ and $dx_2^3 + gx_2^2 + ix_2$ satisfy the constraints of Lemma 3.12, and $r(x_1, x_2)$ is weakly monotone and well-defined. Concerning the latter, one easily derives the following necessary and sufficient conditions using the method proposed in this section: $b, c, j \geqslant 0$ and $b + c + f \geqslant 0$. Thus, f may be negative. Further, note that non-negativeness of b, c and j is necessary for well-definedness of r, and necessity of $b + c + f \geqslant 0$ can, for example, be concluded from $r(1, 1) \geqslant r(0, 1)$.

3.3 Polynomial Interpretations over the Rationals and Reals

In this section, we investigate monotonicity and well-definedness for polynomial interpretations over the rational and real numbers. In particular, we provide complete characterizations of weak and strict monotonicity and relate them to the approach proposed in [45]. Concerning the latter, we argue that the result of Corollary 2.21 can only be established using our approach. Furthermore, we give necessary and sufficient criteria for monotonicity and well-definedness of parametric polynomials of low degree, and we show that, in contrast to polynomial interpretations over the natural numbers, strict monotonicity does not imply weak monotonicity.

According to the definition given in the previous chapter, a weakly (strictly) monotone polynomial interpretation over $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$ is a weakly (strictly) monotone $\mathcal{F}$-algebra $((D_0, \{f_D\}_{f \in \mathcal{F}}), >_{D_0,\delta}, \geqslant_{D_0})$ over the carrier D_0 that associates each function symbol $f \in \mathcal{F}$ with a polynomial f_D with coefficients in D such that f_D is well-defined and weakly (strictly) monotone. Here, as well as in the remainder of this section, weak (strict) monotonicity refers to monotonicity with respect to $\geqslant_{D_0}$ ($>_{D_0,\delta}$), and well-definedness means well-definedness over D_0. Throughout this section, we shall find the following characterizations of weak and strict monotonicity more useful than the ones obtained by specializing Definition 1.2.

Lemma 3.15. *For $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, let f_D be a (not necessarily polynomial) function from D_0^n to D_0, and let $\delta \in D_0$, $\delta > 0$. Then f_D is*

1. *weakly monotone iff $f_D(x_1, \ldots, x_i + h, \ldots, x_n) - f_D(x_1, \ldots, x_i, \ldots, x_n) \geqslant 0$ for all $x_1, \ldots, x_n, h \in D_0$ and all $i \in \{1, \ldots, n\}$.*

2. *strictly monotone iff $f_D(x_1, \ldots, x_i + h, \ldots, x_n) - f_D(x_1, \ldots, x_i, \ldots, x_n) \geqslant \delta$ for all $x_1, \ldots, x_n, h \in D_0$ with $h \geqslant \delta$ and all $i \in \{1, \ldots, n\}$.*

Proof. Obvious from the definition of $>_{D_0,\delta}$ and the definition of monotonicity with respect to $>_{D_0,\delta}$ ($\geqslant_{D_0}$). □

Remark 3.16. Note that the second condition above is formally equivalent to the first one for $\delta = 0$. Hence, weak and strict monotonicity can be treated uniformly.

Corollary 3.17. *In the situation of Lemma 3.15, f_D is strictly monotone iff*

$$f_D(x_1, \dots, x_i + \delta + h, \dots, x_n) - f_D(x_1, \dots, x_i, \dots, x_n) \geqslant \delta$$

for all $x_1, \dots, x_n, h \in D_0$ and all $i \in \{1, \dots, n\}$. □

Intuitively, monotonicity of f_D with respect to $>_{D_0,\delta}$ means that if one increases any of its arguments by at least δ, then the function value increases by at least δ, too. We now show that in general this kind of monotonicity depends on the actual value of δ. That is to say that it may indeed happen that a function is monotone with respect to $>_{D_0,\delta}$ for some value of δ but not so for a different value of δ. In [45], this was already observed for the function

$$f(x) = x - \frac{1}{1 + (x^{\frac{8}{3}} - 2)^2} + \frac{3}{4}$$

which is even non-negative and continuous on $\mathbb{R}_0$, but obviously not a polynomial function. However, the situation is no different for polynomial functions, even for very simple ones, which are indeed relevant for (automated) termination proofs. As an example, let us consider the polynomial function $f_D(x) = x^2$, which was used in Section 2.2 to show that polynomial termination over $\mathbb{Q}$ ($\mathbb{R}_{\mathsf{alg}}$, $\mathbb{R}$) does not imply simple termination (cf. Lemma 2.20 and Corollary 2.21). According to Lemma 3.15, f_D is monotone with respect to $>_{D_0,\delta}$ if and only if

$$(x + h)^2 - x^2 = 2hx + h^2 \geqslant \delta$$

for all $x, h \in D_0$ with $h \geqslant \delta$, which is equivalent to $h^2 \geqslant \delta$ for all $D_0 \ni h \geqslant \delta$. Clearly, the latter holds if and only if it holds for $h = \delta$, that is, if and only if $\delta^2 \geqslant \delta$, which simplifies to $\delta \geqslant 1$ because δ is positive by assumption. Indeed, monotonicity is violated in case δ is less than one; e.g., for $\delta = \frac{1}{2}$, we have $\delta >_{D_0,\delta} 0$, but $f_D(\delta) = \frac{1}{4} >_{D_0,\delta} 0 = f_D(0)$ does not hold.

3.3.1 Weak and Strict Monotonicity

Next we investigate the relationship between weak and strict monotonicity. To begin with, we show how strict monotonicity can be obtained from weak monotonicity.

Lemma 3.18. *In the situation of Lemma 3.15, f_D is strictly monotone if*

1. *f_D is weakly monotone and*
2. *$f_D(x_1, \dots, x_i + \delta, \dots, x_n) - f_D(x_1, \dots, x_i, \dots, x_n) \geqslant \delta$*

for all $x_1, \dots, x_n \in D_0$ and all $i \in \{1, \dots, n\}$.

Proof. Using the characterization of strict monotonicity given in Corollary 3.17, the result is an immediate consequence of (2) and the characterization of weak monotonicity given in Lemma 3.15. □

Note that the conditions mentioned in Lemma 3.18 are only sufficient for strict monotonicity. This is due to the fact that in general strict monotonicity does not imply weak monotonicity, not even for polynomial functions.

Lemma 3.19. *In the situation of Lemma 3.15, strict monotonicity of f_D does not imply weak monotonicity.*

Proof. For $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$ and $\delta \in D_0$, $\delta > 0$, let us consider the function $f_D \colon D_0 \to D_0$, $x \mapsto 2x^2 - x + 1$, which is indeed well-defined over D_0 (cf. Figure 2.1). In addition, f_D is monotone with respect to $>_{D_0,\delta}$ if $\delta = 1$, for example. According to Lemma 3.15, this amounts to showing that $4hx + 2h^2 - h \geqslant 1$ for all $x, h \in D_0$ with $h \geqslant 1$, which does indeed hold as $4hx$ is non-negative and $2h^2 - h = h(2h - 1)$ is at least one if h is at least one. However, f_D is not monotone with respect to $\geqslant_{D_0}$ for any $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$; e.g., $\frac{1}{4} \geqslant_{D_0} 0$ but $f_D\big(\frac{1}{4}\big) = \frac{7}{8} \not\geqslant_{D_0} 1 = f_D(0)$. □

Further, note the similarity between Lemma 3.4 and Lemma 3.18 after setting δ to one and dropping the weak monotonicity condition. This motivates the question whether the latter is needed at all. However, it turns out that weak monotonicity is indeed needed for the soundness of Lemma 3.18. To this end, let us consider the polynomial function $f_D \colon D_0 \to D_0$, $x \mapsto 7x^4 - 28x^3 + 36x^2 - 14x + 2$ for $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, which is well-defined over D_0, as is evident from the graph depicted in Figure 3.1. In addition, it satisfies the second condition of Lemma 3.18 for $\delta = 1$, that is, $f_D(x+1) - f_D(x) \geqslant 1$ for all $x \in D_0$, because the difference $f_D(x+1) - f_D(x)$ can be written as $x \cdot p(x) + 1$, where $p(x) = 28x^2 - 42x + 16$, and $p(x)$ is non-negative for all $x \in D_0$ according to Lemma 3.1. Nevertheless, f_D is not monotone with respect to $>_{D_0,1}$; e.g., we have $\frac{3}{2} >_{D_0,1} 0$ but $f_D\big(\frac{3}{2}\big) = \frac{47}{16} \not>_{D_0,1} 2 = f_D(0)$.

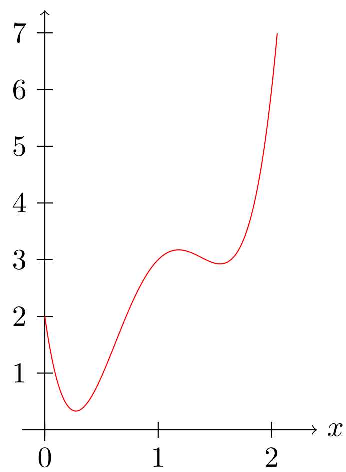

Figure 3.1: The polynomial function $x \mapsto 7x^4 - 28x^3 + 36x^2 - 14x + 2$.

There is another subtlety worth noting about weak and strict monotonicity. Let f be a function from $\mathbb{R}_0^n$ to $\mathbb{R}_0$ that is also well-defined over $\mathbb{R}_{\mathsf{alg},0}$ and $\mathbb{Q}_0$. Then, by definition, monotonicity of f with respect to $\geqslant_{\mathbb{R}_0}$ implies monotonicity with respect to $\geqslant_{\mathbb{R}_{\mathsf{alg},0}}$, which in turn implies monotonicity with respect to $\geqslant_{\mathbb{Q}_0}$. However, in general, the reverse implications do not hold. For example, the function

$$f(x) = \begin{cases} x & \text{if } x \in \mathbb{Q}_0 \\ 0 & \text{otherwise} \end{cases}$$

is obviously monotone with respect to $\geqslant_{\mathbb{Q}_0}$ but not with respect to $\geqslant_{\mathbb{R}_{\mathsf{alg},0}}$; e.g., $\sqrt{2} \geqslant_{\mathbb{R}_{\mathsf{alg},0}} 1$ but $f(\sqrt{2}) = 0 \not\geqslant_{\mathbb{R}_{\mathsf{alg},0}} 1 = f(1)$. Similarly, one can show that monotonicity with respect to $\geqslant_{\mathbb{R}_{\mathsf{alg},0}}$ does not imply monotonicity with respect to $\geqslant_{\mathbb{R}_0}$. Moreover, note that f is monotone with respect to $>_{D_0,\delta}$ for $D = \mathbb{Q}$ (independently of δ) but not for $D = \mathbb{R}_{\mathsf{alg}}$ or $D = \mathbb{R}$. As it turns out, however, this kind of behaviour is solely due to the pathological nature of the function f. For *well-behaved* functions, that is, for functions that are continuous on the non-negative orthant, the situation is not so intricate. In order to see this, we need the following lemma, whose proof is based upon the fact that the behaviour of continuous functions at irrational points is completely defined by the values they take at rational points.

Lemma 3.20. *Let $f\colon \mathbb{R}_0^n \to \mathbb{R}$ be continuous on $\mathbb{R}_0^n$. Then the following statements are equivalent:*

1. *$f(x_1, \dots, x_n) \geqslant 0$ for all $x_1, \dots, x_n \in \mathbb{Q}_0$.*
2. *$f(x_1, \dots, x_n) \geqslant 0$ for all $x_1, \dots, x_n \in \mathbb{R}_{\mathsf{alg},0}$.*
3. *$f(x_1, \dots, x_n) \geqslant 0$ for all $x_1, \dots, x_n \in \mathbb{R}_0$.*

Proof. Clearly, (3) implies (2), which in turn implies (1). So it remains to show the implication from (1) to (3). To this end, let $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}_0^n$ and let $(\vec{x}_k)_{k\in\mathbb{N}}$ be a sequence of n-tuples of non-negative rational numbers $\vec{x}_k \in \mathbb{Q}_0^n$ whose limit is $\vec{x}$. Such a sequence exists because $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$. Then

$$f(\vec{x}) = f(\lim_{k\to\infty} \vec{x}_k) = \lim_{k\to\infty} f(\vec{x}_k)$$

by continuity of f. Thus, $f(\vec{x})$ is the limit of $(f(\vec{x}_k))_{k\in\mathbb{N}}$, which is a sequence of non-negative real numbers by assumption. Hence, $f(\vec{x})$ is non-negative, too. □

Lemma 3.21. *Let f be a continuous function from $\mathbb{R}_0^n$ to $\mathbb{R}_0$ that is well-defined over D_0 with $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}\}$. Then the following statements are equivalent:*

1. *The function f is monotone with respect to $\geqslant_{\mathbb{R}_0}$.*
2. *The function f is monotone with respect to $\geqslant_{D_0}$.*

Proof. By Lemmata 3.15 and 3.20. □

Corollary 3.22. *Let f be a polynomial function from $\mathbb{R}_0^n$ to $\mathbb{R}_0$ induced by a polynomial with coefficients in $\mathbb{Q}$. Then the following statements are equivalent:*

1. *The function f is monotone with respect to $\geqslant_{\mathbb{Q}_0}$.*
2. *The function f is monotone with respect to $\geqslant_{\mathbb{R}_{\mathsf{alg},0}}$.*
3. *The function f is monotone with respect to $\geqslant_{\mathbb{R}_0}$.*

In case f is induced by a polynomial with coefficients in $\mathbb{R}_{\mathsf{alg}}$, at least one of which is irrational, then the last two statements are equivalent. □

Thus, for continuous functions, hence for polynomial functions, one does not have to distinguish between monotonicity with respect to $\geqslant_{\mathbb{Q}_0}$, $\geqslant_{\mathbb{R}_{\mathsf{alg},0}}$ and $\geqslant_{\mathbb{R}_0}$, provided well-definedness over the respective domain holds. A similar statement can be made about strict monotonicity.

Lemma 3.23. *Let f be a continuous function from $\mathbb{R}_0^n$ to $\mathbb{R}_0$ that is well-defined over D_0 with $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}\}$. Then the following statements are equivalent for $D_0 \ni \delta > 0$:*

1. *The function f is monotone with respect to $>_{\mathbb{R}_0,\delta}$.*
2. *The function f is monotone with respect to $>_{D_0,\delta}$.*

Proof. By Corollary 3.17 and Lemma 3.20. □

Corollary 3.24. *Let f be a polynomial function from $\mathbb{R}_0^n$ to $\mathbb{R}_0$ induced by a polynomial with coefficients in $\mathbb{Q}$. Then the following statements are equivalent for $\mathbb{Q}_0 \ni \delta > 0$:*

1. *The function f is monotone with respect to $>_{\mathbb{Q}_0,\delta}$.*
2. *The function f is monotone with respect to $>_{\mathbb{R}_{\mathsf{alg},0},\delta}$.*
3. *The function f is monotone with respect to $>_{\mathbb{R}_0,\delta}$.*

In case f is induced by a polynomial with coefficients in $\mathbb{R}_{\mathsf{alg}}$, at least one of which is irrational, then the last two statements are equivalent for $\mathbb{R}_{\mathsf{alg},0} \ni \delta > 0$. □

3.3.2 Differentiable Functions

Next we turn our attention to functions that are even more well-behaved than continuous functions, namely, differentiable functions. This class of functions was already considered in [45], with the following results. Weak monotonicity can be completely characterized using partial derivatives. This is due to the fact that a function f from $\mathbb{R}_0^n$ to $\mathbb{R}_0$ is monotone in its i-th argument with respect to $\geqslant_{\mathbb{R}_0}$ if and only if it is non-decreasing in its i-th argument (because $\geqslant_{\mathbb{R}_0}$ is just the natural partial order on $\mathbb{R}$). However, non-decreasingness of f in the i-th argument is well-known [8] to be equivalent to the (first-order) partial derivative $\frac{\partial f}{\partial x_i}$ being non-negative (provided it exists). Thus, we obtain the following lemma (cf. also [45, Proposition 2] and [47, Proposition 1]).

Lemma 3.25. *Let f be a function from $\mathbb{R}_0^n$ to $\mathbb{R}_0$ that is continuous and differentiable in all arguments. Then the following statements are equivalent:*

1. *The function f is monotone with respect to $\geqslant_{\mathbb{R}_0}$.*
2. $\frac{\partial f(x_1,\ldots,x_i,\ldots,x_n)}{\partial x_i} \geqslant 0$ *for all $x_1, \ldots, x_n \in \mathbb{R}_0$ and all $i \in \{1, \ldots, n\}$.* □

Corollary 3.26. *Let f be a function from $\mathbb{R}_0^n$ to $\mathbb{R}_0$ that is continuous and differentiable in all arguments, and, in addition, well-defined over D_0 with $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}\}$. Then the following statements are equivalent:*

1. *The function f is monotone with respect to $\geqslant_{\mathbb{R}_0}$.*
2. *The function f is monotone with respect to $\geqslant_{D_0}$.*
3. *$\frac{\partial f(x_1,\ldots,x_i,\ldots,x_n)}{\partial x_i} \geqslant 0$ for all $x_1, \ldots, x_n \in \mathbb{R}_0$ and all $i \in \{1, \ldots, n\}$.*

Proof. By Lemmata 3.25 and 3.21. □

Note that polynomial functions with non-negative real coefficients trivially satisfy the derivative condition of Lemma 3.25, hence they are always weakly monotone. However, there are also polynomial functions that are weakly monotone despite having negative coefficients; e.g., $f(x) = x^3 - x^2 + 2x$, whose derivative $3x^2 - 2x + 2$ is non-negative for all $x \in \mathbb{R}_0$ according to Lemma 3.1.

Finally, comparing Lemma 3.15 and Lemma 3.25 (resp. Corollary 3.26), both of which completely characterize weak monotonicity for differentiable functions, let us mention that the constraints associated with the latter are generally easier to handle than the ones associated with the former.

Concerning strict monotonicity, the following condition is proposed in [45].

Lemma 3.27. *Let f be a function from $\mathbb{R}_0^n$ to $\mathbb{R}_0$ that is continuous and differentiable in all arguments. If $\frac{\partial f(x_1,\ldots,x_i,\ldots,x_n)}{\partial x_i} \geqslant 1$ for all $x_1, \ldots, x_n \in \mathbb{R}_0$ and all $i \in \{1, \ldots, n\}$, then f is monotone with respect to $>_{\mathbb{R}_0,\delta}$ for any positive real number δ. If, in addition, f is well-defined over $\mathbb{Q}_0$ ($\mathbb{R}_{\mathsf{alg},0}$), then f is also monotone with respect to $>_{\mathbb{Q}_0,\delta}$ ($>_{\mathbb{R}_{\mathsf{alg},0},\delta}$) for any positive rational (real algebraic) number δ.* □

In contrast to Lemma 3.15, the derivative condition given in Lemma 3.27 is only sufficient for strict monotonicity; e.g., the function $f(x) = x^2$, which was already shown to be strictly monotone for $\delta \geqslant 1$, cannot be handled by Lemma 3.27. However, Lemma 3.27 has the advantage that it establishes strict monotonicity for any value of δ rather than just for one specific value. In fact, as already mentioned in [45], the derivative condition of Lemma 3.27 is necessary for ensuring strict monotonicity independently of the value of δ.

Furthermore, note that any function that is strictly monotone according to Lemma 3.27 is also weakly monotone by Lemma 3.25 (resp. Corollary 3.26). From this we conclude that Lemma 3.27 is weaker than Lemma 3.18 in the sense that any function that can be shown strictly monotone by the former can also be shown strictly monotone by the latter, but not vice versa, as witnessed (again) by the function $f(x) = x^2$. This is due to the fact that the second condition of Lemma 3.18 is necessary for strict monotonicity, whereas the derivative condition of Lemma 3.27 is not. Moreover, it turns out that the inherent loss of power is considerable. For example, in Section 2.2, we could successfully exploit the function $g(x) = x^2$ to establish polynomial termination of the single rewrite rule $\mathsf{f}(\mathsf{a}) \to \mathsf{f}(\mathsf{g}(\mathsf{a}))$, thus showing that polynomial termination over $\mathbb{Q}$ ($\mathbb{R}_{\mathsf{alg}}$, $\mathbb{R}$) does not imply simple termination (cf. Lemma 2.20 and Corollary 2.21).

However, this result cannot be obtained using the monotonicity criterion of Lemma 3.27 because if the interpretation functions $f(x)$ and $g(x)$ associated with the symbols f and g are strictly monotone according to this lemma, and therefore also weakly monotone, then we must have $f(x) \geqslant x$ and $g(x) \geqslant x$ for all $x \geqslant 0$ (by the mean value theorem of differential calculus, cf. Lemma 3.28 below). Hence, the value of the left-hand side of $\mathsf{f}(\mathsf{a}) \to \mathsf{f}(\mathsf{g}(\mathsf{a}))$ cannot be greater than the value of the right-hand side. So termination cannot be established in this way (not only for polynomial algebras, but for any algebra whatsoever).

Lemma 3.28. *Let f be a function from $\mathbb{R}_0^n$ to $\mathbb{R}_0$ that is continuous and differentiable in all arguments. If $\frac{\partial f(x_1,\ldots,x_i,\ldots,x_n)}{\partial x_i} \geqslant 1$ for all $x_1, \ldots, x_n \in \mathbb{R}_0$ and all $i \in \{1, \ldots, n\}$, then*

$$f(x_1, \ldots, x_i, \ldots, x_n) \geqslant x_i$$

for all $x_1, \ldots, x_n \in \mathbb{R}_0$ and all $i \in \{1, \ldots, n\}$.

Proof. Assume to the contrary that there are $a_1, \ldots, a_n \in \mathbb{R}_0$ and $i \in \{1, \ldots, n\}$ such that $f(a_1, \ldots, a_i, \ldots, a_n) < a_i$, and define a function g from $\mathbb{R}_0$ to $\mathbb{R}_0$ by $g(x_i) := f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$. Then $a_i \neq 0$ by assumption on the range of f and

$$\frac{g(a_i) - g(0)}{a_i - 0} \leqslant \frac{g(a_i)}{a_i} < 1$$

Hence, by Theorem 3.2, there is $0 < \xi < a_i$ such that

$$\frac{\partial g}{\partial x_i}(\xi) = \frac{\partial f}{\partial x_i}(a_1, \ldots, a_{i-1}, \xi, a_{i+1}, \ldots, a_n) < 1$$

However, this contradicts our assumption on the partial derivatives of f. □

In fact, Lemma 3.27 enforces simple termination. To this end, let $\mathcal{F}$ be a signature, and let $\mathcal{A} = (A, \{f_A\}_{f \in \mathcal{F}})$ be an $\mathcal{F}$-algebra (not necessarily polynomial), where $A \subseteq \mathbb{R}_0$ and all interpretation functions are strictly monotone according to Lemma 3.27. Then $\mathcal{A}$ gives rise to a strictly monotone $\mathcal{F}$-algebra $(\mathcal{A}, >_{A,\delta}, \geqslant_A)$ together with $>_{A,\delta}$ and $\geqslant_A$, the restrictions of $>_{\mathbb{R}_0,\delta}$ and $\geqslant_{\mathbb{R}_0}$ to A. More importantly, by Lemma 3.28, $(\mathcal{A}, >_{A,\delta}, \geqslant_A)$ is also a well-founded simple monotone $\mathcal{F}$-algebra (in the sense of [70,77], cf. Definition 3.29 below), and such algebras are well-known to characterize simple termination.

Definition 3.29. A *simple monotone $\mathcal{F}$-algebra* $(\mathcal{A}, >, \geqslant)$ is an $\mathcal{F}$-algebra $\mathcal{A} = (A, \{f_A\}_{f \in \mathcal{F}})$ together with an order $>$ and a partial order $\geqslant$ on A such that $> \cdot \geqslant \subseteq >$ and for each $f \in \mathcal{F}$, f_A is monotone with respect to $>$, and

$$f_A(x_1, \ldots, x_i, \ldots, x_n) \geqslant x_i$$

for all $x_1, \ldots, x_n \in A$ and all $i \in \{1, \ldots, n\}$. A simple monotone algebra is said to be *well-founded* if $>$ is well-founded.

Remark 3.30. In [70, Definition 6.3.1], $\geqslant$ is implicitly assumed to be the reflexive closure of $>$. However, this is not essential. For the validity of the results established on top of this definition, it is sufficient if $>$ and $\geqslant$ satisfy the compatibility condition $> \cdot \geqslant \subseteq >$.

Proposition 3.31 ([70, Proposition 6.3.8]). *A TRS is simply terminating if and only if it admits a compatible well-founded simple monotone algebra.* □

Proposition 3.32 ([70, Proposition 6.3.5]). *Let $\mathcal{R}$ be a TRS over a finite signature $\mathcal{F}$, and let $(\mathcal{A}, >, \geqslant)$ be a compatible simple monotone $\mathcal{F}$-algebra. Then $\mathcal{R}$ is terminating.* □

As a consequence of Proposition 3.31, any TRS that is compatible with the algebra $(\mathcal{A}, >_{A,\delta}, \geqslant_A)$ is necessarily simply terminating.

More importantly, it turns out that the aforementioned advantage that Lemma 3.27 establishes strict monotonicity for any value of δ entails the fact that we can dispense with $>_{A,\delta}$ altogether, replacing it by $>_A$, the restriction of the natural order $>_{\mathbb{R}_0}$ on the non-negative reals to A. This statement is based on the observation that the resulting algebra $(\mathcal{A}, >_A, \geqslant_A)$ is still a simple monotone algebra. To this end, note that well-foundedness of $>_A$ is not (necessarily) required in Definition 3.29 and that an interpretation function is monotone with respect to $>_A$ if it is (strictly) increasing in all its arguments, which is known [8] to be the case if all partial derivatives are positive, a condition that is obviously implied by the derivative condition of Lemma 3.27. Hence, by Proposition 3.32, compatibility of the algebra $(\mathcal{A}, >_A, \geqslant_A)$ with a TRS over a finite signature implies termination. Moreover, it is easy to see that any TRS (resp. any set of rewrite rules) that is compatible with the algebra $(\mathcal{A}, >_{A,\delta}, \geqslant_A)$ is also compatible with $(\mathcal{A}, >_A, \geqslant_A)$ since $>_{A,\delta} \subseteq >_A$, and likewise for weak compatibility. This shows that, in case of finite signatures, the traditional simple monotone algebra approach using the natural order on the reals subsumes the approach of [45] if Lemma 3.27 is used to enforce strict monotonicity.

3.3.3 Parametric Polynomials

In analogy to Section 3.2, we are now going to consider parametric polynomials of low degree with the aim of providing necessary and sufficient constraints in terms of their abstract coefficients such that monotonicity and well-definedness of the resulting concrete polynomials are guaranteed for every instantiation of the coefficients that satisfies the constraints.

Due to Lemma 3.20 and Corollaries 3.22 and 3.24, for polynomial functions, it suffices to consider monotonicity with respect to $>_{D_0,\delta}$ ($\geqslant_{D_0}$) and well-definedness over D_0 for $D = \mathbb{R}$ only; e.g., if f is a polynomial function induced by a polynomial with rational coefficients, then f is well-defined over $\mathbb{Q}_0$ if and only if it is well-defined over $\mathbb{R}_0$ according to Lemma 3.20, and for a positive rational number δ, it is monotone with respect to $>_{\mathbb{Q}_0,\delta}$ ($\geqslant_{\mathbb{Q}_0}$) if and only if it is monotone with respect to $>_{\mathbb{R}_0,\delta}$ ($\geqslant_{\mathbb{R}_0}$) by Corollaries 3.22 and 3.24 (assuming well-definedness). In particular, for parametric polynomials, these monotonicity and well-definedness requirements can readily be expressed in the first-order theory of the reals. For example, the following formula characterizes well-definedness and strict monotonicity (using Lemma 3.15) of the univariate quadratic parametric polynomial $ax^2 + bx + c$:

$$\begin{aligned}\forall x\, \forall h\, ((x \geqslant 0 \implies ax^2 + bx + c \geqslant 0) \wedge \\ (x \geqslant 0 \wedge h \geqslant \delta \implies 2ahx + ah^2 + bh \geqslant \delta))\end{aligned} \tag{3.7}$$

In this formula, all variables, that is, the quantified variables x and h as well as the free variables a, b, c and δ, are supposed to range over the real numbers, and what we seek is an equivalent formula in which no quantified variables appear. For this purpose, we can use any quantifier elimination algorithm for the first-order theory of the reals (resp. the first-order theory of real closed fields), though the calculations are generally tedious to do by hand. However, in our case, it turns out that this process is amenable for automation because of the fact that we are dealing with relatively simple polynomials of low degree. For example, the quantifier elimination tool QEPCAD B [11] implementing the partial cylindrical algebraic decomposition (CAD) algorithm of Collins and Hong [14, 15] produces an equivalent quantifier-free formula for (3.7) in about one second.[1] Despite this fact, we shall formally treat linear and quadratic parametric polynomials below as the obtained results will be used throughout this thesis. But beforehand, let us observe the following lemma relating weak monotonicity and well-definedness.

Lemma 3.33. *For $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, let $f_D\colon D_0^n \to D$ denote the polynomial function associated with a polynomial in $D[x_1, \dots, x_n]$. Then f_D is monotone with respect to $\geqslant_{D_0}$ and well-defined over D_0 if and only if it is monotone with respect to $\geqslant_{D_0}$ and $f_D(0, \dots, 0) \geqslant 0$.*

Proof. Obvious from Lemma 3.15. □

Thus, well-definedness is implicitly taken care of by the single constraint $f_D(0, \dots, 0) \geqslant 0$. In order to ensure weak monotonicity in Lemma 3.33, one may either employ Lemma 3.15 or Lemma 3.25 (resp. Corollary 3.26).

For strict monotonicity and well-definedness, the analogon of Lemma 3.33 does not hold (basically because strict monotonicity does not imply weak monotonicity); e.g., taking $f(x) = 2x^2 - x + 1$, which was already shown to be monotone with respect to $>_{D_0,\delta}$ for $\delta = 1$ using Lemma 3.15, we observe that the function $g(x) = f(x) - 1$ satisfies $g(0) \geqslant 0$ as well as the strict monotonicity condition of Lemma 3.15 (for $\delta = 1$) since $g(x + h) - g(x) = f(x + h) - f(x)$. Nevertheless, g is not well-defined over D_0 because $g\left(\frac{1}{4}\right) = -\frac{1}{8}$. Therefore, we have to explicitly take care of well-definedness.

Linear Parametric Polynomials

Lemma 3.34. *For $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, the linear parametric polynomial function $f_D(x_1, \dots, x_n) = a_n x_n + \dots + a_1 x_1 + a_0$ with $a_i \in D$ for $i = 0, 1, \dots, n$ is monotone with respect to $\geqslant_{D_0}$ and well-defined over D_0 if and only if $a_i \geqslant 0$ for all $i \in \{0, 1, \dots, n\}$.*

[1] On a desktop computer equipped with 1 GB of main memory and a single Intel® Pentium® 4 processor running at a clock rate of 3 GHz.

Proof. Easy consequence of Lemmata 3.33 and 3.15. $\square$

Lemma 3.35. *For $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$ and $\delta \in D_0$, $\delta > 0$, the linear parametric polynomial function $f_D(x_1, \dots, x_n) = a_n x_n + \dots + a_1 x_1 + a_0$ with $a_i \in D$ for $i = 0, 1, \dots, n$ is monotone with respect to $>_{D_0,\delta}$ and well-defined over D_0 if and only if $a_0 \geqslant 0$ and $a_i \geqslant 1$ for all $i \in \{1, \dots, n\}$.*

Proof. According to Lemma 3.15, this lemma holds if and only if f_D is well-defined over D_0 and $f_D(x_1, \dots, x_i + h, \dots, x_n) - f_D(x_1, \dots, x_i, \dots, x_n) \geqslant \delta$ holds for all $x_1, \dots, x_n, h \in D_0$ with $h \geqslant \delta$ and all $i \in \{1, \dots, n\}$. By definition of f_D, the latter condition simplifies to $a_i h \geqslant \delta$ for all $h \geqslant \delta$ and all $i \in \{1, \dots, n\}$, which holds if and only if it holds for $h = \delta$, that is, if $a_i \geqslant 1$ for all $i \in \{1, \dots, n\}$ (since $\delta > 0$). But then $a_0 = f_D(0, \dots, 0) \geqslant 0$ is not only necessary but also sufficient for well-definedness. $\square$

Quadratic Parametric Polynomials

Next we consider the generic quadratic parametric polynomial function

$$f_D(x_1, \dots, x_n) = a_0 + \sum_{j=1}^{n} a_j x_j + \sum_{1 \leqslant j \leqslant k \leqslant n} a_{jk} x_j x_k \ \in D[x_1, \dots, x_n] \tag{3.8}$$

Lemma 3.36. *For $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, the function f_D of (3.8) is monotone with respect to $\geqslant_{D_0}$ and well-defined over D_0 if and only if all its coefficients are non-negative.*

Proof. As a consequence of Lemma 3.20, the function f_D is well-defined over D_0 if and only if it is well-defined over $\mathbb{R}_0$. Then, by Lemma 3.25 (for $D = \mathbb{R}$) and Corollary 3.26 (for $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}\}$), the claim holds if and only if f_D is well-defined over $\mathbb{R}_0$ and

$$\frac{\partial f_D}{\partial x_i}(x_1, \dots, x_i, \dots, x_n) = 2a_{ii} x_i + \sum_{i < k \leqslant n} a_{ik} x_k + \sum_{1 \leqslant j < i} a_{ji} x_j + a_i \geqslant 0$$

for all $x_1, \dots, x_n \in \mathbb{R}_0$ and all $i \in \{1, \dots, n\}$. Being a quantified linear inequality in $x_1, \dots, x_n$, the latter condition holds if and only if all its coefficients are non-negative. Taking the quantification over i into account, this shows that all coefficients of f_D, except for a_0, must be non-negative. But then $a_0 \geqslant 0$ is not only necessary but also sufficient for well-definedness. $\square$

Hence, all coefficients must be non-negative for weak monotonicity and well-definedness. This is no longer the case for the combination of strict monotonicity and well-definedness. To this end, we start with the following result, where we put the focus on strict monotonicity. Well-definedness will be treated subsequently.

Lemma 3.37. *For $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$ and $\delta \in D_0$, $\delta > 0$, the function f_D of (3.8) is monotone with respect to $>_{D_0,\delta}$ and well-defined over D_0 if and only if it is well-defined over D_0, all its coefficients, except for $a_1, \dots, a_n$, are non-negative and $a_{ii}\delta + a_i \geqslant 1$ for all $i \in \{1, \dots, n\}$.*

Proof. By Lemma 3.15, the claim holds if and only if f_D is well-defined over D_0 and $f_D(x_1, \dots, x_i + h, \dots, x_n) - f_D(x_1, \dots, x_i, \dots, x_n) - \delta =$

$$\sum_{i<k\leqslant n} ha_{ik}x_k + \sum_{1\leqslant j<i} ha_{ji}x_j + 2ha_{ii}x_i + (h^2a_{ii} + ha_i - \delta) \geqslant 0 \tag{3.9}$$

holds for all $x_1, \dots, x_n, h \in D_0$ with $h \geqslant \delta$ and all $i \in \{1, \dots, n\}$. Being a quantified linear inequality in $x_1, \dots, x_n$, quantifier elimination with respect to $x_1, \dots, x_n$ amounts to demanding non-negativeness of all coefficients, such that (3.9) holds for all $x_1, \dots, x_n, h \in D_0$ with $h \geqslant \delta$ and all $i \in \{1, \dots, n\}$ if and only if for all $h \in D_0$ with $h \geqslant \delta$ and all $i \in \{1, \dots, n\}$,

- $ha_{ik} \geqslant 0$ for $k = i+1, \dots, n$, and
- $ha_{ji} \geqslant 0$ for $j = 1, \dots, i-1$, and
- $2ha_{ii} \geqslant 0$, and
- $h^2a_{ii} + ha_i - \delta \geqslant 0$.

Now quantifier elimination with respect to h is trivial for the first three items because h is always positive, so we can just discard it. Taking the quantification over i into account, this shows that all coefficients of f_D, except for $a_0, a_1, \dots, a_n$, must be non-negative. But $a_0 = f_D(0, \dots, 0) \geqslant 0$ is obviously necessary for well-definedness. In the fourth item, the quantifier associated with the variable h can be eliminated as follows. Clearly, for $h = \delta$, we obtain the necessary condition $\delta^2a_{ii} + \delta a_i \geqslant \delta$, or equivalently, $\delta a_{ii} + a_i \geqslant 1$. However, this condition is also sufficient because for $h \geqslant \delta$, it implies $ha_{ii} + a_i \geqslant 1$, which in turn implies $h^2a_{ii} + ha_i \geqslant h \geqslant \delta$. □

Hence, for strict monotonicity (and well-definedness), all coefficients must be non-negative, except for the coefficients associated with the linear monomials. In contrast, the strict monotonicity condition of Lemma 3.27 requires the partial derivatives

$$\frac{\partial f_D}{\partial x_i}(x_1, \dots, x_i, \dots, x_n) = 2a_{ii}x_i + \sum_{i<k\leqslant n} a_{ik}x_k + \sum_{1\leqslant j<i} a_{ji}x_j + a_i$$

to be at least one for all non-negative values of $x_1, \dots, x_n$ and all $i \in \{1, \dots, n\}$. Together with $a_0 \geqslant 0$, which is obviously necessary for well-definedness, this means that all coefficients must be non-negative and $a_i \geqslant 1$ for all $i \in \{1, \dots, n\}$. Hence, the latter approach admits negative coefficients only for polynomials of degree three or higher. In addition, note that if all coefficients are non-negative and $a_i \geqslant 1$ for all $i \in \{1, \dots, n\}$, then well-definedness trivially holds, and strict monotonicity holds with respect to any value of δ. However, we can also distill this information from Lemma 3.37 by quantifying over δ. Doing so, we see that the function f_D of (3.8) is well-defined over D_0 and monotone with respect to $>_{D_0,\delta}$ for any δ if and only if all coefficients are non-negative and $a_i \geqslant 1$ for all $i \in \{1, \dots, n\}$.

Now how about well-definedness of f_D? Obviously, by Lemma 3.37, non-well-definedness can only be caused by the coefficients $a_1, \ldots, a_n$ associated with the linear monomials. These coefficients are bounded from below by $a_i \geqslant 1 - a_{ii}\delta$ for all $i \in \{1, \ldots, n\}$. So if all a_{ii} are zero, then f_D has no negative coefficients and well-definedness trivially holds. On the other hand, if some a_{ii} is positive and the corresponding a_i negative, then the term $a_{ii}x_i^2$ eventually dominates the term $a_i x_i$; in particular, we have

$$a_{ii}x_i^2 + a_i x_i = a_{ii}\left(x_i + \frac{a_i}{2a_{ii}}\right)^2 - \frac{a_i^2}{4a_{ii}}$$

Consequently, well-definedness of f_D can always be achieved by choosing its constant coefficient large enough. In this way, one easily derives a sufficient criterion for well-definedness. Quite surprisingly, though, the computation of conditions that are both necessary and sufficient turns out to be quite involved in general, that is, in the multivariate case (even for two indeterminates). The result for the univariate case is as follows.

Lemma 3.38. *For $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$ and $\delta \in D_0$, $\delta > 0$, the function $f_D(x) = ax^2 + bx + c$ with $a, b, c \in D$ is monotone with respect to $>_{D_0,\delta}$ and well-defined over D_0 if and only if $a \geqslant 0 \wedge c \geqslant 0 \wedge (b \geqslant 0 \vee 4ac - b^2 \geqslant 0) \wedge a\delta + b \geqslant 1$.*

Proof. By Lemma 3.37, the claim holds if and only if f_D is well-defined over D_0 and $a \geqslant 0$, $c \geqslant 0$ and $a\delta + b \geqslant 1$. Furthermore, according to Lemmata 3.1 and 3.20, the conditions $a \geqslant 0$, $c \geqslant 0$ and $(b \geqslant 0 \vee 4ac - b^2 \geqslant 0)$ are necessary and sufficient for well-definedness. ☐

In the bivariate case, the function $f_D(x_1, x_2) = ax_1^2 + bx_2^2 + cx_1x_2 + dx_1 + ex_2 + f$ with coefficients in D is monotone with respect to $>_{D_0,\delta}$ and well-defined over D_0 if and only if

$$a, b, c, f \geqslant 0 \wedge a\delta + d \geqslant 1 \wedge b\delta + e \geqslant 1 \wedge \bigvee_{i=1}^{5} \phi_i$$

where the ϕ_i are defined as follows: for $D_1 := 4bf - e^2$ and $D_2 := 4af - d^2$,

$$\phi_1 := d \geqslant 0 \wedge e \geqslant 0 \qquad \phi_2 := d \geqslant 0 \wedge D_1 \geqslant 0 \qquad \phi_3 := e \geqslant 0 \wedge D_2 \geqslant 0$$
$$\phi_4 := D_1 \geqslant 0 \wedge D_2 \geqslant 0 \wedge (2ae - cd \geqslant 0 \vee 2bd - ce \geqslant 0)$$
$$\phi_5 := D_1 > 0 \wedge c^2 f - 4abf + ae^2 - cde + bd^2 \leqslant 0$$

As this result is never used in the remainder of the thesis (unlike the one of Lemma 3.38), we skip its (lengthy) formal proof. We do note, however, that the conditions listed above are equivalent to the ones obtained by the quantifier elimination tool QEPCAD B.

Example 3.39. Consider the polynomial function $f(x) = 2x^2 - x + 1$, for which we have already established well-definedness and strict monotonicity for $\delta = 1$. Now we can use Lemma 3.38 to infer that it is in fact well-defined and strictly monotone for any $\delta \geqslant 1$.

3.4 Negative Coefficients in Polynomial Interpretations

In the previous sections, we have seen that in principle one may use polynomial interpretations with (some) negative coefficients for proving termination of TRSs. Now the obvious question is the following: does there exist a TRS that can be proved terminating by a polynomial interpretation with negative coefficients, but cannot be proved terminating by a polynomial interpretation where the coefficients of all polynomials are non-negative? To elaborate on this question, let us consider the TRS $\mathcal{R}_1$ comprising the two rewrite rules

$$\mathsf{f}(\mathsf{a}) \to \mathsf{f}(\mathsf{b}) \qquad\qquad \mathsf{g}(\mathsf{b}) \to \mathsf{g}(\mathsf{a})$$

We claim that $\mathcal{R}_1$ can be shown polynomially terminating over $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ and $\mathbb{R}$, but only by means of negative coefficients.

Lemma 3.40. *The TRS $\mathcal{R}_1$ is polynomially terminating over $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ and $\mathbb{R}$.*

Proof. The following interpretation establishes polynomial termination over $\mathbb{Q}$:

$$\delta = 1 \qquad \mathsf{a}_{\mathbb{Q}} = 0 \qquad \mathsf{b}_{\mathbb{Q}} = \tfrac{1}{2} \qquad \mathsf{g}_{\mathbb{Q}}(x) = 2x \qquad \mathsf{f}_{\mathbb{Q}}(x) = 6x^2 - 5x + 2$$

First, we show compatibility of this interpretation with the rules of $\mathcal{R}_1$ by observing that the inequalities

$$\mathsf{f}_{\mathbb{Q}}(\mathsf{a}_{\mathbb{Q}}) >_{\mathbb{Q}_0,\delta} \mathsf{f}_{\mathbb{Q}}(\mathsf{b}_{\mathbb{Q}}) \qquad\qquad \mathsf{g}_{\mathbb{Q}}(\mathsf{b}_{\mathbb{Q}}) >_{\mathbb{Q}_0,\delta} \mathsf{g}_{\mathbb{Q}}(\mathsf{a}_{\mathbb{Q}})$$

simplify to $2 >_{\mathbb{Q}_0,1} 1$ and $1 >_{\mathbb{Q}_0,1} 0$. So they do indeed hold according to the definition of $>_{\mathbb{Q}_0,1}$. Next we note that well-definedness and strict monotonicity of $\mathsf{f}_{\mathbb{Q}}$ and $\mathsf{g}_{\mathbb{Q}}$ follow directly from Lemmata 3.35 and 3.38. In particular, $\mathsf{f}_{\mathbb{Q}}$ is well-defined over $\mathbb{Q}_0$ because it has a global minimum at $x_{min} = \frac{5}{12}$, namely $\mathsf{f}_{\mathbb{Q}}(x_{min}) = \frac{23}{24}$, which is positive. This shows polynomial termination over $\mathbb{Q}$. Polynomial termination over $\mathbb{R}_{\mathsf{alg}}$ and $\mathbb{R}$ follows by the same interpretation. □

Next we show that polynomial termination cannot be established, not even incrementally, if negative coefficients are not allowed in the interpretation. To this end, let us consider a strictly monotone polynomial interpretation $\mathcal{P}$ over $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, all of whose interpretation functions are given by polynomials with non-negative coefficients, and let us assume that $\mathcal{P}$ is weakly compatible with both rules of $\mathcal{R}_1$ and (strictly) compatible with at least one rule. Then $\mathcal{P}$ is also weakly monotone (e.g. by Lemma 3.25 or Corollary 3.26, respectively) because all coefficients are assumed to be non-negative. In particular, all interpretation functions are non-decreasing, which is the essential difference to the interpretation given in the proof of Lemma 3.40. Clearly, for (strict) compatibility of $\mathcal{P}$ with a non-empty subset of $\mathcal{R}_1$, the interpretations of the constants a and b must not be equal. Without loss of generality, let $\mathsf{b}_D >_{D_0} \mathsf{a}_D$ (but not necessarily $\mathsf{b}_D >_{D_0,\delta} \mathsf{a}_D$). From this we obtain $\mathsf{f}_D(\mathsf{b}_D) \geqslant_{D_0} \mathsf{f}_D(\mathsf{a}_D)$ by weak monotonicity of f_D. Yet this inequality cannot be strict due to weak compatibility of $\mathcal{P}$ with the first rule of $\mathcal{R}_1$; hence, $\mathsf{f}_D(\mathsf{a}_D) = \mathsf{f}_D(\mathsf{b}_D)$. But then, as a consequence of non-decreasingness, f_D must be constant in the non-empty interval $[\mathsf{a}_D, \mathsf{b}_D]$. However, for a polynomial function, this can only be the case if it is constant everywhere, which is not permissible because of the requirement of strict monotonicity. Hence, there is no polynomial interpretation over $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ or $\mathbb{R}$ that is weakly monotone as well as strictly monotone and proves (incremental) polynomial termination of $\mathcal{R}_1$. In particular, this also reveals that the statement of Lemma 3.40 cannot be shown by a polynomial interpretation using the strict monotonicity condition of Lemma 3.27.

We conclude this section with the remark that there are also TRSs that can be proved terminating by a polynomial interpretation over $\mathbb{N}$ with negative coefficients, but cannot be proved polynomially terminating over $\mathbb{N}$ using only non-negative coefficients. We shall present an example of such a TRS in the next chapter. In fact, for this particular TRS, a stronger statement holds. That is to say that, despite the fact that the system is polynomially terminating over $\mathbb{N}$ (but only with negative coefficients), polynomial termination cannot be established over $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ or $\mathbb{R}$, not even by means of negative coefficients (cf. Lemmata 4.14 and 4.15).

3.5 Conclusion

The topic of this chapter was the investigation of polynomial interpretations with regard to criteria that guarantee monotonicity (weak and/or strict) and well-definedness of the functions induced by the polynomials occurring in such interpretations. For polynomial interpretations over the rational and real (algebraic) numbers, we gave complete characterizations of weak and strict monotonicity, the latter subsuming the current approach proposed in [45]. Most notably, the fact that polynomial termination over $\mathbb{Q}$ ($\mathbb{R}_{\mathsf{alg}}$, $\mathbb{R}$) does not imply simple termination can only be established using our approach. We also presented a simple TRS, consisting of the single rewrite rule $\mathsf{f}(\mathsf{a}) \to \mathsf{f}(\mathsf{g}(\mathsf{a}))$, for which polynomial termination can only be shown using the monotonicity criteria proposed in this chapter. In addition, we investigated the relationship between weak and strict monotonicity, showing that, unlike for polynomial interpretations over the natural numbers, strict monotonicity does not imply weak monotonicity.

Having automation in mind, we put our focus on parametric polynomials of low degree, providing constraints in terms of their abstract coefficients such that monotonicity and well-definedness of the resulting concrete polynomials are guaranteed for every instantiation of the coefficients satisfying the constraints. For polynomial interpretations over the natural numbers, our approach subsumes the one proposed in [16], which is currently used in many termination provers, and, in contrast to the latter, it allows negative numbers in certain coefficients. Similarly, for polynomial interpretations over the rational and real (algebraic) numbers, the constraints obtained by our approach subsume the ones of [45], and negative coefficients can be handled without further ado. Moreover, we showed that there are indeed TRSs that can be proved terminating by polynomial interpretations with negative coefficients, but cannot be proved terminating by polynomial interpretations where all coefficients are non-negative. Some of the criteria proposed in this chapter for polynomial interpretations over the natural numbers have been implemented in the termination prover $\mathsf{T_TT_2}$ [41] (by the tool developers). Experimental results can be found in [58]. In these experiments, checking (weak) compatibility of a polynomial interpretation (over $\mathbb{N}$) with a rewrite rule $\ell \to r$ is done by testing $P_\ell - P_r \underset{(=)}{>} 0$ using the method of absolute positiveness of Chapter 2. Yet, taking a closer look at the latter, one might argue that it is not ideally suited for this purpose (at least not in theory) because, by definition, it only admits polynomials with non-negative coefficients inducing weakly monotone functions, thus precluding a vast majority of permissible polynomials. For future work, we therefore propose to investigate alternatives for the absolute positiveness approach. More concretely, we suggest to consider one that is based on the fact that a (multivariate) polynomial p is (globally) non-negative if it admits a *sum of squares (SOS) decomposition* [64], that is, if it can be written as $p = \sum_i f_i^2$ using a finite number of polynomials f_i. Clearly, if such a decomposition exists, then $p(x) \geqslant 0$ holds for all $x \in \mathbb{R}^n$. The reverse implication does not hold in general (cf. [64]). What makes this approach appealing is the fact that the existence of an SOS decomposition can be decided by solving a *semidefinite programming feasibility problem*, which can be done efficiently, that is, in polynomial time [61, 62], whereas the general problem of testing global non-negativity of a polynomial function is NP-hard (when the degree is at least four, cf. e.g. [62]). So it might be fruitful to approximate the polynomial non-negativity constraints arising from (weak) compatibility of a polynomial interpretation with a set of rewrite rules by a set of SOS constraints rather than using the absolute positiveness approach which is weaker in theory. Note that the relaxation to non-negativity on $\mathbb{R}_0^n$ poses no problem as $p(x) \geqslant 0$ holds for all $x \in \mathbb{R}_0^n$ if and only if $p(x^2) \geqslant 0$ holds for all $x \in \mathbb{R}^n$. Also note that the constraints arising from (strict) compatibility can readily be handled as they can easily be expressed as non-negativity constraints.

The results obtained in this chapter also enable us to prove the decidability result for polynomial interpretations over $\mathbb{Q}$ mentioned in the previous chapter (in Section 2.2). To this end, let $\mathcal{F}$ be a signature, $\{f_{\mathbb{Q}}\}_{f \in \mathcal{F}}$ a collection of polynomials with coefficients in $\mathbb{Q}$, and let δ be a positive rational number. According to Lemma 3.20, for each n-ary function symbol $f \in \mathcal{F}$, $f_{\mathbb{Q}}$ is well-defined over $\mathbb{Q}_0$ if and only if it is well-defined over $\mathbb{R}_0$, with the decidability of the latter being an immediate consequence of Tarski's decidability result for the first-order theory of the reals. Similarly, assuming well-definedness, we conclude from Corollary 3.22 that each interpretation function is monotone with respect to $\geqslant_{\mathbb{Q}_0}$ if and only if it is monotone with respect to $\geqslant_{\mathbb{R}_0}$, which is again decidable by Tarski's result. In the same way, one obtains the decidability of monotonicity with respect to $>_{\mathbb{Q}_0,\delta}$ from Corollary 3.24. Therefore, it is decidable whether the pair $(\{f_{\mathbb{Q}}\}_{f \in \mathcal{F}}, \delta)$ constitutes a valid polynomial interpretation over $\mathbb{Q}$, and weak as well as strict monotonicity of the interpretation is decidable, too. Moreover, (weak) compatibility with a given set of rewrite rules is also decidable because, by Lemma 3.20, the (weak) compatibility constraint $P_\ell - P_r - \delta \geqslant 0$ ($P_\ell - P_r \geqslant 0$) associated with some rewrite rule $\ell \to r$ holds for all $x_1, \dots, x_m \in \mathbb{Q}_0$ if and only if it holds for all $x_1, \dots, x_m \in \mathbb{R}_0$ (where the variables $x_1, \dots, x_m$ correspond to the ones occurring in $\ell \to r$), the latter again being decidable by Tarski's result.

Chapter 4

Polynomial Termination Hierarchy

In Chapter 2, we have introduced four variants of polynomial interpretations, polynomial interpretations over $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ and $\mathbb{R}$. Now the obvious question is:

What is their relationship with regard to termination proving power?

For Knuth-Bendix orders, for example, it is known [19,40,44] that extending the range of the underlying weight function from the natural numbers to the non-negative reals does not result in an increase in termination proving power. As far as (direct) polynomial termination (in the sense of Definitions 2.3 and 2.8) is concerned, a partial answer to the above question was given in 2006 by Lucas [46], who managed to show that there are TRSs that can be shown terminating by polynomial interpretations with rational coefficients, but cannot be shown terminating using polynomials with integer coefficients only. Likewise, he proved that there are TRSs that can be handled by polynomial interpretations with real (algebraic) coefficients, but cannot be handled by polynomial interpretations with rational coefficients. Based on these results and the fact that we have the strict inclusions $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$, there is the common yet unproven belief (expressed in e.g. [10,47]) in the term rewriting community that polynomial interpretations with real coefficients properly subsume polynomial interpretations with rational coefficients, which in turn properly subsume polynomial interpretations with integer coefficients. In this chapter, we shall see, however, that this is not true. In general, the situation turns out to be as depicted in Figure 4.1, which illustrates both our results and the earlier results of Lucas [46]. In particular, we

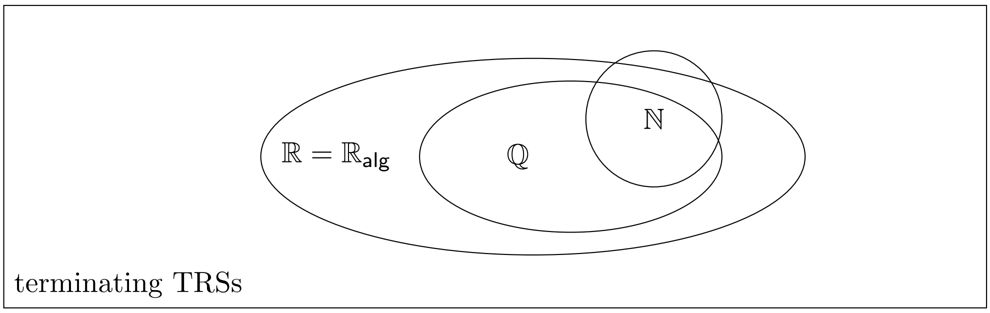

Figure 4.1: Comparison.

prove that polynomial interpretations with rational coefficients are subsumed by polynomial interpretations with real coefficients, which turn out to be equivalent to polynomial interpretations with real algebraic coefficients. Moreover, we show that polynomial interpretations with real or rational coefficients do not subsume polynomial interpretations with integer coefficients. Likewise, we prove that there are TRSs that can be shown terminating by polynomial interpretations with real coefficients as well as by polynomial interpretations with integer coefficients, but cannot be shown terminating using polynomials with rational coefficients only.

We also consider the possibility of establishing termination by using polynomial interpretations in an *incremental* way, an approach already known to Lankford [43, Example 3]. As for (direct) polynomial termination, we give the full picture of the relationship between the four notions of incremental polynomial termination over $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ and $\mathbb{R}$ as introduced in Definition 2.9, showing that it is essentially the same as the one depicted in Figure 4.1 for (direct) polynomial termination. Besides, we also provide evidence that the situation is no different for polynomial interpretations in the context of the DP framework.

The remainder of this chapter is organized as follows. In Section 4.1, we present an auxiliary result concerning polynomial interpretations that will be helpful in Section 4.2, where we show that polynomial interpretations with real algebraic coefficients are equivalent to polynomial interpretations with real coefficients when it comes to proving termination of TRSs (in the DP framework as well as if used as a stand-alone termination method). Then, in Section 4.3, we give the full picture of the relationship between the various notions of polynomial termination, whereas Section 4.4 is dedicated to incremental polynomial termination. Finally, we present some results related to the DP framework in Section 4.5, before concluding in Section 4.6.

Some of the results presented in this chapter originally appeared in the conference paper [54]. A substantially extended paper [57] has been submitted for publication. As far as new contributions are concerned, the solution to the open problem mentioned in [54] whether polynomial termination over $\mathbb{N}$ and $\mathbb{R}$ implies polynomial termination over $\mathbb{Q}$ is provided, thereby completing the picture of the relationship between the various notions of polynomial termination. The results of Sections 4.4 and 4.5, which extend this relationship to incremental polynomial termination and polynomial interpretations in the context of the DP framework, are also new.

4.1 Preliminaries

As described in the previous chapters, in order to prove termination by means of a polynomial interpretation, the following conditions must be satisfied: well-definedness over the respective carrier, monotonicity (weak or strict) and (weak) compatibility with some set of rewrite rules. In this section, we show that for polynomial interpretations over $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ and $\mathbb{R}$ all of these conditions can be expressed as polynomial non-negativity constraints. To this end, let $\mathcal{F}$ be a signature, $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, and consider a collection of polynomials $\{f_D\}_{f \in \mathcal{F}}$ with coefficients in $D' \subseteq D$, where D' is a subring of D and D_0 is the carrier of the interpretation. We claim that all of the aforementioned conditions can be phrased as (conjunctions of) quantified polynomial inequalities of the shape

$$p(x_1, \dots, x_n, \delta) \geqslant 0 \text{ for all } x_1, \dots, x_n \in D_0 \tag{4.1}$$

for some polynomial p in the indeterminates $x_1, \dots, x_n$ and δ with coefficients in the ring D', i.e., $p \in D'[x_1, \dots, x_n, \delta]$.

For well-definedness of the polynomials f_D over D_0, this is obvious from Definition 2.1 and the assumption on the nature of their coefficients, whereas for weak (resp. strict) monotonicity it follows from Lemma 3.15 (resp. Corollary 3.17). In particular, for each n-ary function symbol $f \in \mathcal{F}$, strict monotonicity of f_D in its i-th argument is equivalent to non-negativeness of

$$p(x_1, \dots, x_n, h, \delta) = f_D(x_1, \dots, x_i + \delta + h, \dots, x_n) - f_D(x_1, \dots, x_n) - \delta$$

for all $x_1, \dots, x_n, h \in D_0$. So it remains to show $p \in D'[x_1, \dots, x_n, h, \delta]$. Yet this is an immediate consequence of the usual closure properties of polynomials. To this end, note that f_D is in $D'[x_1, \dots, x_n, h, \delta]$ by assumption. Obviously, both $x_i + \delta + h$ and δ are in $D'[x_1, \dots, x_n, h, \delta]$ as well. Hence, by closure under composition, $f_D(x_1, \dots, x_i + \delta + h, \dots, x_n)$ is then also in $D'[x_1, \dots, x_n, h, \delta]$, and, by closure under addition, the same holds for p. This shows that strict monotonicity can indeed be expressed in the form of (4.1). By similar reasoning, the same is true of weak monotonicity.

Next we note that (weak) compatibility with some rewrite rule $\ell \to r$ amounts to the satisfaction of the non-negativity constraint $P_\ell - P_r - \delta \geqslant 0$ ($P_\ell - P_r \geqslant 0$) for all $x_1, \dots, x_n \in D_0$ (where the variables $x_1, \dots, x_n$ correspond to the ones occurring in $\ell \to r$). By assumption, both P_ℓ and P_r are composed of polynomials with coefficients in the ring D' (because each f_D is a polynomial with coefficients in D'). Hence, by closure under composition, we obtain that P_ℓ and P_r are in $D'[x_1, \dots, x_n] \subseteq D'[x_1, \dots, x_n, \delta]$. Then, by closure under addition, the polynomials $P_\ell - P_r$ and $P_\ell - P_r - \delta$ are in $D'[x_1, \dots, x_n, \delta]$ as well.

We conclude this section with the following corollary. If the value of δ is known, that is, for $\delta \in D_0$, $\delta > 0$, all of the conditions mentioned above can be expressed as (conjunctions of) quantified polynomial inequalities of the shape

$$p(x_1, \dots, x_n) \geqslant 0 \text{ for all } x_1, \dots, x_n \in D_0 \tag{4.2}$$

where p is a polynomial with coefficients in D (as δ is not necessarily in D').

4.2 Real Algebraic Numbers Suffice

As already mentioned at the end of Chapter 2, for polynomial interpretations over $\mathbb{R}$, it suffices to consider polynomials with real algebraic coefficients as interpretations of function symbols according to a result established in [47]. In this section, we extend this result by showing that it even suffices to restrict to the non-negative real algebraic numbers as the carrier for polynomial interpretations with real algebraic coefficients. In other words, we show that polynomial interpretations over $\mathbb{R}$ are equivalent to polynomial interpretations over $\mathbb{R}_{\mathsf{alg}}$

with respect to proving termination of TRSs (in the context of the DP framework as well as if used as a stand-alone termination method). For this purpose, we exploit the following concepts and results from the field of real algebraic geometry [6, 9].

Definition 4.1. A *real closed field* R is an ordered field (i.e., a field equipped with a total order that is compatible with the field operations) such that every positive element of R has a square root in R and every polynomial in $R[x]$ of odd degree has a root in R.

Definition 4.2. Let R be a real closed field and D a subring of R. A *first-order formula in the language of ordered fields with coefficients in D* is a formula written with a finite number of conjunctions, disjunctions, negations, and universal or existential quantifiers on variables, starting from atomic formulas, which are formulas of the kind $f(x_1,\dots,x_n) = 0$, $f(x_1,\dots,x_n) > 0$ or $f(x_1,\dots,x_n) \geqslant 0$, where f is a polynomial with coefficients in D. The *free variables* of a formula are those variables of the polynomials occurring in the formula, which are not quantified. A *sentence* is a formula without free variables.

Theorem 4.3 (Tarski-Seidenberg Transfer Principle)**.** *Suppose that R' is a real closed field that contains the real closed field R. If Φ is a sentence in the language of ordered fields with coefficients in R, then it is true in R if and only if it is true in R'.* □

With these preliminaries at hand, we are now ready to state the main result of this section.

Lemma 4.4. *Let $\mathcal{P}$ be a weakly (strictly) monotone polynomial interpretation over $\mathbb{R}$ such that $\mathcal{S}_1 \subseteq {>_{\mathcal{P}}}$ and $\mathcal{S}_2 \subseteq {\geqslant_{\mathcal{P}}}$ for some finite sets of rewrite rules $\mathcal{S}_1$ and $\mathcal{S}_2$. Then there exists a weakly (strictly) monotone polynomial interpretation $\mathcal{N}$ over $\mathbb{R}_{\mathsf{alg}}$ such that $\mathcal{S}_1 \subseteq {>_{\mathcal{N}}}$ and $\mathcal{S}_2 \subseteq {\geqslant_{\mathcal{N}}}$. The converse statement (obtained by exchanging $\mathbb{R}$ and $\mathbb{R}_{\mathsf{alg}}$) also holds.*

Proof. Let $\mathcal{F}$ denote the signature associated with $\mathcal{S}_1 \cup \mathcal{S}_2$, and let δ be a positive real number and $\mathcal{P} = (\{f_{\mathbb{R}}\}_{f\in\mathcal{F}}, \delta)$ a weakly (strictly) monotone polynomial interpretation over $\mathbb{R}$ such that $\mathcal{S}_1 \subseteq {>_{\mathcal{P}}}$ and $\mathcal{S}_2 \subseteq {\geqslant_{\mathcal{P}}}$. That is, $\mathcal{P}$ satisfies the following conditions:

(a) for each n-ary symbol $f \in \mathcal{F}$, $f_{\mathbb{R}}(x_1,\dots,x_n) \geqslant 0$ for all $x_1,\dots,x_n \in \mathbb{R}_0$,

(b) for each symbol $f \in \mathcal{F}$, $f_{\mathbb{R}}$ is monotone with respect to $\geqslant_{\mathbb{R}_0}$ ($>_{\mathbb{R}_0,\delta}$),

(c) for each rule $\ell \to r \in \mathcal{S}_1$, $P_\ell >_{\mathbb{R}_0,\delta} P_r$ for all $x_1,\dots,x_m \in \mathbb{R}_0$, and

(d) for each rule $s \to t \in \mathcal{S}_2$, $P_s \geqslant_{\mathbb{R}_0} P_t$ for all $y_1,\dots,y_k \in \mathbb{R}_0$.

Here, the variables $x_1,\dots,x_m$ and $y_1,\dots,y_k$ correspond to the ones occurring in $\ell \to r$ and $s \to t$, respectively. As a consequence of the aforementioned result of [47], each interpretation function $f_{\mathbb{R}}$ can be assumed to be a polynomial with real algebraic coefficients. Treating δ as a variable (which we will later quantify existentially) and applying the result of Section 4.1, we note that all four conditions can be phrased as (conjunctions of) quantified polynomial inequalities of the shape "$p(x_1, \ldots, x_n, \delta) \geqslant 0$ for all $x_1, \ldots, x_n \in \mathbb{R}_0$" for some polynomial p with real algebraic coefficients. Moreover, any of these quantified inequalities can readily be expressed as a formula in the language of ordered fields with coefficients in $\mathbb{R}_{\mathsf{alg}}$, where δ is the only free variable. By taking the conjunction of all these formulas, existentially quantifying δ and adding the conjunct $\delta > 0$, we obtain a sentence S in the language of ordered fields with coefficients in $\mathbb{R}_{\mathsf{alg}}$ (because $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{F}$ are finite). By assumption, this sentence is true in the real closed field $\mathbb{R}$. But then it must also be true in the real closed field $\mathbb{R}_{\mathsf{alg}}$ because of Theorem 4.3, which states that S is true in $\mathbb{R}$ if and only if it is true in $\mathbb{R}_{\mathsf{alg}}$. Hence, there exists a positive real *algebraic* number $\delta_{\mathcal{N}}$ such that the pair $\mathcal{N} = (\{f_{\mathbb{R}}\}_{f \in \mathcal{F}}, \delta_{\mathcal{N}})$ constitutes a valid polynomial interpretation over $\mathbb{R}_{\mathsf{alg}}$ that is weakly (strictly) monotone if $\mathcal{P}$ is weakly (strictly) monotone and satisfies $\mathcal{S}_1 \subseteq >_{\mathcal{N}}$ and $\mathcal{S}_2 \subseteq \geqslant_{\mathcal{N}}$.

The converse statement can be shown in the same way. Again, we construct a sentence S in the language of ordered fields with coefficients in $\mathbb{R}_{\mathsf{alg}}$, but this time it is true in $\mathbb{R}_{\mathsf{alg}}$ by assumption. However, by Theorem 4.3, S is true in $\mathbb{R}_{\mathsf{alg}}$ if and only if it is true in $\mathbb{R}$. □

Remark 4.5. Taking a closer look at the result of [47], which is crucially used in the proof of Lemma 4.4, one observes that its proof is not as explicit as it should be, at least not for our purpose. In particular, it does not take the δ into account, which is an integral part of any polynomial interpretation over $\mathbb{R}$. While it is true that δ can be neglected in some cases, this is by no means always the case, as we have already seen in Chapter 3 and will see again in the present chapter. Besides, the proof given in [47] also lacks an explicit treatment of monotonicity. However, these issues can be resolved. A revised proof is given in Appendix A (cf. the proof of Lemma A.1).

Lemma 4.4 gives rise to the following corollaries.

Corollary 4.6. *A TRS is (incrementally) polynomially terminating over $\mathbb{R}$ if and only if it is (incrementally) polynomially terminating over $\mathbb{R}_{\mathsf{alg}}$.* □

Corollary 4.7. *If there exists a weakly monotone polynomial interpretation over $\mathbb{R}$ that succeeds on a given DP problem, then there also exists a weakly monotone polynomial interpretation over $\mathbb{R}_{\mathsf{alg}}$ that succeeds on this DP problem, and vice versa.* □

Thus, polynomial interpretations over $\mathbb{R}$ are equivalent to polynomial interpretations over $\mathbb{R}_{\mathsf{alg}}$ when it comes to proving termination of TRSs. Nevertheless, for the sake of brevity of notation, we will mostly stick to the term "polynomial interpretations over the real numbers" in the remainder of this thesis.

4.3 Direct Polynomial Termination

According to the result established in Section 4.2, for termination analysis, we do not have to distinguish between polynomial interpretations over $\mathbb{R}$ and polynomial interpretations over $\mathbb{R}_{\mathsf{alg}}$. So we are left with polynomial interpretations over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$. In this section, we investigate the relationship between the various notions of polynomial termination (in the sense of Definitions 2.3 and 2.8) induced by them. As far as this relationship is concerned, the following results are known [46].

Theorem 4.8 ([46, Corollary 3]). *There are TRSs that are polynomially terminating over $\mathbb{Q}$ but not over $\mathbb{N}$.* □

Theorem 4.9 ([46, Corollary 2]). *There are TRSs that are polynomially terminating over $\mathbb{R}$ but not over $\mathbb{Q}$ or $\mathbb{N}$.* □

Hence, the extension of the coefficient domain from the integers to the rational numbers entails the possibility to prove some TRSs polynomially terminating, which could not be proved polynomially terminating otherwise. Moreover, a similar statement holds for the extension of the coefficient domain from the rational numbers to the real numbers. Based on these results and the fact that we have the strict inclusions $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$, it is tempting to believe that polynomial interpretations over the reals properly subsume polynomial interpretations over the rationals, which in turn properly subsume polynomial interpretations over the natural numbers. Indeed, as will be shown below, the former proposition can be settled in the affirmative. However, quite surprisingly, the latter proposition does not hold. We show this by means of a TRS that is polynomially terminating over $\mathbb{N}$ but not over $\mathbb{R}$ or $\mathbb{Q}$. Furthermore, we also present a TRS that can be shown polynomially terminating over $\mathbb{N}$ and $\mathbb{R}$ but not over $\mathbb{Q}$, thereby completing the full picture of the relationship between polynomial termination over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$.

4.3.1 Polynomial Termination over $\mathbb{Q}$ vs. $\mathbb{R}$

We start with an easy result showing that polynomial termination over $\mathbb{Q}$ implies polynomial termination over $\mathbb{R}$. Actually, we prove the following more general result, whose proof is based upon the fact that polynomials induce continuous functions, the behaviour of which at irrational points is completely defined by the values they take at rational points (cf. Lemma 3.20).

Lemma 4.10. *Let $\mathcal{P}$ be a weakly (strictly) monotone polynomial interpretation over $\mathbb{Q}$ such that $\mathcal{S}_1 \subseteq {>_\mathcal{P}}$ and $\mathcal{S}_2 \subseteq {\geqslant_\mathcal{P}}$ for some finite sets of rewrite rules $\mathcal{S}_1$ and $\mathcal{S}_2$. Then there exists a weakly (strictly) monotone polynomial interpretation $\mathcal{N}$ over $\mathbb{R}$ such that $\mathcal{S}_1 \subseteq {>_\mathcal{N}}$ and $\mathcal{S}_2 \subseteq {\geqslant_\mathcal{N}}$.*

Proof. Let $\mathcal{F}$ be a signature comprising the symbols occurring in $\mathcal{S}_1$ and $\mathcal{S}_2$, and let $D \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$, $\delta \in D_0$, $\delta > 0$, and $\mathcal{P} = (\{f_D\}_{f \in \mathcal{F}}, \delta)$ such that $\mathcal{P}$ is a weakly (strictly) monotone polynomial interpretation over D that is compatible with $\mathcal{S}_1$ and weakly compatible with $\mathcal{S}_2$. That is, $\mathcal{P}$ satisfies the following conditions:

(a) for each n-ary symbol $f \in \mathcal{F}$, $f_D(x_1, \dots, x_n) \geqslant 0$ for all $x_1, \dots, x_n \in D_0$,

(b) for each symbol $f \in \mathcal{F}$, f_D is monotone with respect to $\geqslant_{D_0}$ ($>_{D_0,\delta}$),

(c) for each rule $\ell \to r \in \mathcal{S}_1$, $P_\ell >_{D_0,\delta} P_r$ for all $x_1, \ldots, x_m \in D_0$, and

(d) for each rule $s \to t \in \mathcal{S}_2$, $P_s \geqslant_{D_0} P_t$ for all $y_1, \ldots, y_k \in D_0$.

Applying the result of Section 4.1, we note that all four conditions can be phrased as (conjunctions of) quantified polynomial inequalities of the shape "$p(x_1, \ldots, x_n) \geqslant 0$ for all $x_1, \ldots, x_n \in D_0$" for some polynomial p with coefficients in D. Now by Lemma 3.20 and the fact that polynomials induce continuous functions all these inequalities do not only hold in D_0 but also in D'_0, for any $D' \in \{\mathbb{Q}, \mathbb{R}_{\mathsf{alg}}, \mathbb{R}\}$ such that $D \subseteq D'$. Hence, the polynomial interpretation $\mathcal{P} = (\{f_D\}_{f \in \mathcal{F}}, \delta)$ is also a valid polynomial interpretation over D' that is weakly (strictly) monotone, compatible with $\mathcal{S}_1$ and weakly compatible with $\mathcal{S}_2$. In particular, letting $D = \mathbb{Q}$ and $D' = \mathbb{R}$ proves the claim. □

Remark 4.11. Note that the proof of Lemma 4.10 actually establishes a stronger result than claimed in the lemma. In particular, for $D = \mathbb{R}_{\mathsf{alg}}$ and $D' = \mathbb{R}$, it represents an alternative proof of the converse statement of Lemma 4.4. Moreover, $\mathcal{S}_1$ and $\mathcal{S}_2$ need not necessarily be finite.

The following result is a direct consequence of Lemma 4.10.

Corollary 4.12. *If a TRS is polynomially terminating over $\mathbb{Q}$, then it is also polynomially terminating over $\mathbb{R}$.* □

The converse statement does not hold according to Theorem 4.9. We conclude this subsection with the following remark stressing the essence of the proof of Lemma 4.10.

Remark 4.13. Not only does the result established in this subsection show that polynomial interpretations over $\mathbb{R}$ ($\mathbb{R}_{\mathsf{alg}}$) subsume polynomial interpretations over $\mathbb{Q}$, but it also reveals that the same interpretation applies.

4.3.2 Polynomial Termination over $\mathbb{N}$ vs. $\mathbb{R}$

According to Theorem 4.9, there are TRSs that are polynomially terminating over $\mathbb{R}$ but not over $\mathbb{N}$. Next we present a TRS that is polynomially terminating over $\mathbb{N}$ but not over $\mathbb{R}$, and hence also not over $\mathbb{Q}$. In order to motivate the construction of this particular TRS, let us first observe that from the viewpoint of number theory there is a fundamental difference between the integers (resp. the natural numbers) and the real or rational numbers. More precisely, the integers are an example of a discrete domain, whereas both $\mathbb{R}$ and $\mathbb{Q}$ are *dense*[1] domains. In the context of polynomial interpretations, the consequences of this major distinction are best explained by an example. Let us consider the polynomial function $x \mapsto 2x^2 - x$ depicted in Figure 4.2, and let us assume that we want to use it as the interpretation of some unary function symbol. Now the point is that this function is permissible in a polynomial interpretation over $\mathbb{N}$, even in a strictly monotone one, as it is both well-defined over $\mathbb{N}$ and strictly monotone

[1] Given two distinct real (rational) numbers a and b, there exists a real (rational) number c in between.

(i.e., monotone with respect to $>_{\mathbb{N}}$). However, viewing it as a function of a non-negative real (resp. rational) variable, we observe that well-definedness over $\mathbb{R}_0$ (resp. $\mathbb{Q}_0$) is violated in the open interval $(0, \frac{1}{2})$ (and monotonicity with respect to $>_{\mathbb{R}_0,\delta}$ (resp. $>_{\mathbb{Q}_0,\delta}$) requires a properly chosen value for δ). Hence, the polynomial function $x \mapsto 2x^2 - x$ is not permissible in any polynomial interpretation over $\mathbb{R}$ (resp. $\mathbb{Q}$).

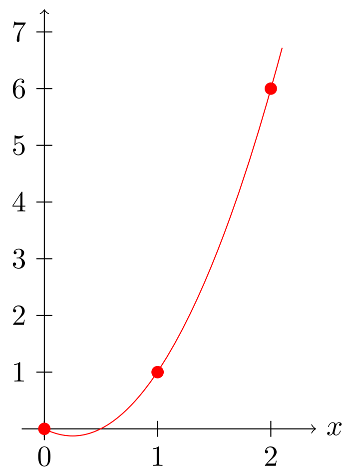

Figure 4.2: The polynomial function $x \mapsto 2x^2 - x$.

Thus, the idea is to design a TRS that enforces an interpretation of this shape for some unary function symbol, and the tool that can be used to achieve this is polynomial interpolation. To elaborate on this, let us consider the following scenario, which is fundamentally based on the assumption that some unary function symbol f is interpreted by a quadratic polynomial $\mathsf{f}(x) = ax^2 + bx + c$ with (unknown) coefficients a, b and c. Then, by polynomial interpolation, these coefficients are uniquely determined by the image of f at three pairwise different locations; in this way, the interpolation constraints $\mathsf{f}(0) = 0$, $\mathsf{f}(1) = 1$ and $\mathsf{f}(2) = 6$ enforce the interpretation $\mathsf{f}(x) = 2x^2 - x$. Next we encode these constraints in terms of the TRS $\mathcal{R}$ consisting of the following rewrite rules, where $\mathsf{s}^n(x)$ abbreviates $\underbrace{\mathsf{s}(\mathsf{s}(\cdots\mathsf{s}}_{n\text{-times}}(x)\cdots))$,

$$
\begin{aligned}
\mathsf{s}(\mathsf{0}) &\to \mathsf{f}(\mathsf{0}) & & \\
\mathsf{s}^2(\mathsf{0}) &\to \mathsf{f}(\mathsf{s}(\mathsf{0})) & \qquad \mathsf{f}(\mathsf{s}(\mathsf{0})) &\to \mathsf{0} \\
\mathsf{s}^7(\mathsf{0}) &\to \mathsf{f}(\mathsf{s}^2(\mathsf{0})) & \qquad \mathsf{f}(\mathsf{s}^2(\mathsf{0})) &\to \mathsf{s}^5(\mathsf{0})
\end{aligned}
$$

and we consider the following two cases: polynomial interpretations over $\mathbb{N}$ on the one hand and polynomial interpretations over $\mathbb{R}$ (resp. $\mathbb{Q}$) on the other hand.

In the context of polynomial interpretations over $\mathbb{N}$, we observe that if we equip the function symbols s and $\mathsf{0}$ with the (natural) interpretations $\mathsf{s}_{\mathbb{N}}(x) = x + 1$ and $\mathsf{0}_{\mathbb{N}} = 0$, then the TRS $\mathcal{R}$ indeed implements the above interpolation constraints.[2] For example, the constraint $\mathsf{f}_{\mathbb{N}}(1) = 1$ is expressed by $\mathsf{f}(\mathsf{s}(\mathsf{0})) \to \mathsf{0}$ and $\mathsf{s}^2(\mathsf{0}) \to \mathsf{f}(\mathsf{s}(\mathsf{0}))$. The former encodes $\mathsf{f}_{\mathbb{N}}(1) > 0$, whereas the latter encodes $\mathsf{f}_{\mathbb{N}}(1) < 2$. Moreover, the rule $\mathsf{s}(\mathsf{0}) \to \mathsf{f}(\mathsf{0})$ encodes $\mathsf{f}_{\mathbb{N}}(0) < 1$, which is equivalent to $\mathsf{f}_{\mathbb{N}}(0) = 0$ in the domain of the natural numbers. Thus, this interpolation constraint can be expressed by a single rewrite rule, whereas the other two constraints require two rules each. Summing up, by virtue of the method of polynomial interpolation, we have reduced the problem of enforcing a specific interpretation for some unary function symbol to the problem of enforcing a natural semantics for the symbols s and $\mathsf{0}$.

Next we elaborate on the ramifications of considering the TRS $\mathcal{R}$ in the context of polynomial interpretations over $D \in \{\mathbb{Q}, \mathbb{R}\}$. For this purpose, let us assume that the symbols s and $\mathsf{0}$ are interpreted by $\mathsf{s}_D(x) = x + s_0$ and $\mathsf{0}_D = 0$, so that s has some kind of *successor function* semantics. Then the compatibility constraints associated with the TRS $\mathcal{R}$ are as follows:

$$\begin{aligned} s_0 - \delta &\geqslant \mathsf{f}_D(0) & & \\ 2s_0 - \delta &\geqslant \mathsf{f}_D(s_0) & \qquad \mathsf{f}_D(s_0) &\geqslant 0 + \delta \\ 7s_0 - \delta &\geqslant \mathsf{f}_D(2s_0) & \qquad \mathsf{f}_D(2s_0) &\geqslant 5s_0 + \delta \end{aligned}$$

Hence, $\mathsf{f}_D(0)$ is confined to the closed interval $[0, s_0 - \delta]$, whereas $\mathsf{f}_D(s_0)$ is confined to $[0 + \delta, 2s_0 - \delta]$ and $\mathsf{f}_D(2s_0)$ to $[5s_0 + \delta, 7s_0 - \delta]$. Basically, this means that these constraints do not uniquely determine the function f_D. In other words, the method of polynomial interpolation does not readily apply to the case of polynomial interpretations over $\mathbb{R}$ and $\mathbb{Q}$. However, we can make it work. To this end, we observe that if $s_0 = \delta$, then the above system of inequalities actually turns into a system of equations, which can be viewed as a set of interpolation constraints (parameterized by s_0) that uniquely determine f_D:

$$\mathsf{f}_D(0) = 0 \qquad \mathsf{f}_D(s_0) = s_0 \qquad \mathsf{f}_D(2s_0) = 6s_0$$

Clearly, if $s_0 = \delta = 1$, then the symbol f is fixed to the interpretation $2x^2 - x$, as was the case in the context of polynomial interpretations over $\mathbb{N}$ (note that in the latter case $\delta = 1$ is implicit because of the equivalence $x >_{\mathbb{N}} y$ iff $x \geqslant_{\mathbb{N}} y + 1$). Hence, we draw the conclusion that once we can manage to design a TRS that enforces $s_0 = \delta$, we can again leverage the method of polynomial interpolation to enforce a specific interpretation for some unary function symbol. Moreover, we remark that the actual value of s_0 is irrelevant for achieving our goal. That is to say that s_0 only serves as a scale factor in the interpolation constraints determining f_D. Clearly, if $s_0 \neq 1$, then f_D is not fixed to the interpretation $2x^2 - x$. However, it is still fixed to an interpretation of the same (desired) shape. But more on this later.

In the preceding paragraphs we have presented the basic method that we use in order to show that polynomial termination over $\mathbb{N}$ does not imply polynomial termination over $\mathbb{R}$. The construction given was based on several assumptions, the essential ones of which are listed below:

(a) The symbol s had to be interpreted by a linear polynomial $x \mapsto x + s_0$.

[2] In fact, one can even show that $\mathsf{s}_{\mathbb{N}}(x) = x + 1$ is sufficient for this purpose.

(b) The condition $s_0 = \delta$ was required to hold.

(c) The function symbol f had to be interpreted by a quadratic polynomial.

Now the point is that one can get rid of all these assumptions by adding suitable rewrite rules to the TRS $\mathcal{R}$. The resulting TRS, which we will refer to as $\mathcal{R}_2$, consists of the rules given in Table 4.1. In this TRS, the rules (4.9) and (4.10)

$\mathsf{s}(\mathsf{0}) \to \mathsf{f}(\mathsf{0})$	(4.3)	$\mathsf{f}(\mathsf{g}(x)) \to \mathsf{g}(\mathsf{g}(\mathsf{f}(x)))$	(4.9)
$\mathsf{s}^2(\mathsf{0}) \to \mathsf{f}(\mathsf{s}(\mathsf{0}))$	(4.4)	$\mathsf{g}(\mathsf{s}(x)) \to \mathsf{s}(\mathsf{s}(\mathsf{g}(x)))$	(4.10)
$\mathsf{s}^7(\mathsf{0}) \to \mathsf{f}(\mathsf{s}^2(\mathsf{0}))$	(4.5)	$\mathsf{g}(x) \to \mathsf{h}(x, x)$	(4.11)
$\mathsf{f}(\mathsf{s}(\mathsf{0})) \to \mathsf{0}$	(4.6)	$\mathsf{s}(x) \to \mathsf{h}(\mathsf{0}, x)$	(4.12)
$\mathsf{f}(\mathsf{s}^2(\mathsf{0})) \to \mathsf{s}^5(\mathsf{0})$	(4.7)	$\mathsf{s}(x) \to \mathsf{h}(x, \mathsf{0})$	(4.13)
$\mathsf{f}(\mathsf{s}^2(x)) \to \mathsf{h}(\mathsf{f}(x), \mathsf{g}(\mathsf{h}(x, x)))$	(4.8)	$\mathsf{h}(\mathsf{f}(x), \mathsf{g}(x)) \to \mathsf{f}(\mathsf{s}(x))$	(4.14)

Table 4.1: The TRS $\mathcal{R}_2$.

serve the purpose of ensuring the first of the above items. Informally, (4.10) constrains the interpretation of the symbol s to a linear polynomial by simple reasoning about the degrees of the left- and right-hand side polynomials, and (4.9) does the same thing with respect to g. Because both interpretations are linear, compatibility with (4.10) can only be achieved if the leading coefficient of the interpretation of s is one.

Concerning item (c) above, we remark that the tricky part is to enforce the upper bound of two on the degree of the polynomial f_D interpreting the symbol f (in a polynomial interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$). We use the following approach. Let $\mathsf{f}_D(x)$ be a polynomial of degree $n \geqslant 1$, and let $s_0 \neq 0$. Then Taylor's theorem [60] for polynomials tells us that

$$\mathsf{f}_D(x+s_0) = \sum_{k=0}^{n} \frac{s_0^k}{k!}\mathsf{f}_D^{(k)}(x) = \mathsf{f}_D(x)+s_0\mathsf{f}'_D(x)+\frac{s_0^2}{2}\mathsf{f}''_D(x)+\cdots+\frac{s_0^n}{n!}\mathsf{f}_D^{(n)}(x) \quad (4.15)$$

Hence, $\deg(\mathsf{f}_D(x+s_0)-\mathsf{f}_D(x)) = \deg(\mathsf{f}_D(x))-1 = n-1$ since $s_0 \neq 0$. From this we infer that $\mathsf{f}_D(x)$ is at most quadratic if $\mathsf{f}_D(x+s_0)-\mathsf{f}_D(x)$ is at most linear. That is, $\mathsf{f}_D(x)$ is at most quadratic if there exists a linear function $\mathsf{g}_D(x)$ such that $\mathsf{g}_D(x) \geqslant \mathsf{f}_D(x+s_0)-\mathsf{f}_D(x)$, or equivalently, $\mathsf{f}_D(x)+\mathsf{g}_D(x) \geqslant \mathsf{f}_D(x+s_0)$, for all non-negative values of x. This can be encoded in terms of rule (4.14) as soon as the interpretation of h corresponds to the addition of two numbers. And this is exactly the purpose of the rules (4.11), (4.12) and (4.13). More precisely, by linearity of the interpretation of g, we infer from (4.11) that the interpretation of h must have the linear shape $h_1x + h_2y + h_0$. Furthermore, compatibility with (4.12) and (4.13) implies $h_1 = h_2 = 1$ due to item (a) above. Hence, the interpretation of h really models the addition of two numbers (modulo adding a constant). So the above construction can indeed be used to set the desired upper bound on $\deg(\mathsf{f}_D(x))$. More generally, it can be used to set *arbitrary* upper bounds on the degree of an interpretation, and it readily allows to establish lower bounds as well.

Next we comment on how to enforce the second of the above assumptions. As it turns out, the hard part is to enforce the condition $s_0 \leqslant \delta$. The idea is as follows. First, we consider the rule (4.4), observing that if f is interpreted by a quadratic polynomial f_D and s by the linear polynomial $x + s_0$, then (the interpretation of) its right-hand side will eventually become larger than its left-hand side with growing s_0, thus violating compatibility. In this way, s_0 is bounded from above, and the faster the growth of f_D, the lower the bound. The problem with this statement, however, is that it is only true if f_D is fixed (which is a priori not the case); otherwise, for any given value of s_0, one can always find a quadratic polynomial f_D such that compatibility with (4.4) is satisfied. The parabolic curve associated with f_D only has to be flat enough. So in order to prevent this, we have to somehow control the growth of f_D. Now that is where the rule (4.8) comes into play, which basically expresses that if one increases the argument of f_D by a certain amount (i.e., $2s_0$), then the value of the function is guaranteed to increase by a certain minimum amount as well. Thus, this rule establishes a lower bound on the growth of f_D. And it turns out that if f_D has just the right amount of growth, then we can readily establish the desired upper bound δ for s_0.

Finally, having presented all the relevant details of our construction, it remains to formally prove our main claim that the TRS $\mathcal{R}_2$ is polynomially terminating over $\mathbb{N}$ but not over $\mathbb{R}$ or $\mathbb{Q}$.

Lemma 4.14. *The TRS $\mathcal{R}_2$ is polynomially terminating over $\mathbb{N}$.*

Proof. We consider the following interpretation:

$$\mathsf{0}_\mathbb{N} = 0 \quad \mathsf{s}_\mathbb{N}(x) = x + 1 \quad \mathsf{f}_\mathbb{N}(x) = 2x^2 - x \quad \mathsf{g}_\mathbb{N}(x) = 4x + 4 \quad \mathsf{h}_\mathbb{N}(x, y) = x + y$$

By Lemma 3.6 and Corollary 3.9, all functions are well-defined over $\mathbb{N}$ and strictly monotone (i.e., monotone with respect to $>_\mathbb{N}$). In addition, this interpretation is compatible with $\mathcal{R}_2$ because the resulting inequalities

$$\begin{array}{cc}
1 >_\mathbb{N} 0 & 32x^2 + 60x + 28 >_\mathbb{N} 32x^2 - 16x + 20 \\
2 >_\mathbb{N} 1 & 4x + 8 >_\mathbb{N} 4x + 6 \\
7 >_\mathbb{N} 6 & 4x + 4 >_\mathbb{N} 2x \\
1 >_\mathbb{N} 0 & x + 1 >_\mathbb{N} x \\
6 >_\mathbb{N} 5 & x + 1 >_\mathbb{N} x \\
2x^2 + 7x + 6 >_\mathbb{N} 2x^2 + 7x + 4 & 2x^2 + 3x + 4 >_\mathbb{N} 2x^2 + 3x + 1
\end{array}$$

are clearly satisfied for all natural numbers x. □

Lemma 4.15. *The TRS $\mathcal{R}_2$ is not polynomially terminating over $\mathbb{R}$.*

Proof. Let us assume that $\mathcal{R}_2$ is polynomially terminating over $\mathbb{R}$ and derive a contradiction. To begin with, we observe that compatibility with (4.10) implies

$$\deg(\mathsf{g}_\mathbb{R}(x)) \cdot \deg(\mathsf{s}_\mathbb{R}(x)) \geqslant \deg(\mathsf{s}_\mathbb{R}(x)) \cdot \deg(\mathsf{s}_\mathbb{R}(x)) \cdot \deg(\mathsf{g}_\mathbb{R}(x))$$

together with the requirement that all interpretation functions must be strictly monotone, which means that $\mathsf{s}_{\mathbb{R}}(x)$ and $\mathsf{g}_{\mathbb{R}}(x)$ cannot be constant polynomials, that is, both $\deg(\mathsf{s}_{\mathbb{R}}(x))$ and $\deg(\mathsf{g}_{\mathbb{R}}(x))$ must be at least 1. But then the above inequality simplifies to $\deg(\mathsf{s}_{\mathbb{R}}(x)) \leqslant 1$. Hence, we obtain $\deg(\mathsf{s}_{\mathbb{R}}(x)) = 1$ and, by applying the same reasoning to (4.9), $\deg(\mathsf{g}_{\mathbb{R}}(x)) = 1$. So the symbols s and g must be interpreted by linear polynomials $\mathsf{s}_{\mathbb{R}}(x) = s_1 x + s_0$ and $\mathsf{g}_{\mathbb{R}}(x) = g_1 x + g_0$, where $s_0, g_0 \in \mathbb{R}_0$ and $s_1, g_1 \geqslant_{\mathbb{R}} 1$ due to Lemma 3.35. Then compatibility with (4.10) translates to the inequality

$$g_1 s_1 x + g_1 s_0 + g_0 >_{\mathbb{R}_0,\delta} s_1^2 g_1 x + s_1^2 g_0 + s_1 s_0 + s_0 \tag{4.16}$$

which must hold for all non-negative real numbers x. This implies the following condition on the respective leading coefficients: $g_1 s_1 \geqslant_{\mathbb{R}} s_1^2 g_1$. Due to $s_1, g_1 \geqslant_{\mathbb{R}} 1$, this can only hold if $s_1 = 1$. Hence, $\mathsf{s}_{\mathbb{R}}(x) = x + s_0$. This result simplifies (4.16) to $g_1 s_0 >_{\mathbb{R}_0,\delta} 2s_0$, which implies $g_1 s_0 >_{\mathbb{R}} 2s_0$. From this we conclude that $s_0 >_{\mathbb{R}} 0$ and $g_1 >_{\mathbb{R}} 2$.

Now suppose that the function symbol f were also interpreted by a linear polynomial $\mathsf{f}_{\mathbb{R}}$. Then we could apply the same reasoning to the rule (4.9) because it is structurally equivalent to (4.10), thus inferring $g_1 = 1$. However, this would contradict $g_1 >_{\mathbb{R}} 2$. Therefore, $\mathsf{f}_{\mathbb{R}}$ cannot be linear.

Next we turn our attention to the rewrite rules (4.11), (4.12) and (4.13). Because $\mathsf{g}_{\mathbb{R}}$ is a linear polynomial function, compatibility with (4.11) constrains the function $h\colon \mathbb{R}_0 \to \mathbb{R}_0$, $x \mapsto \mathsf{h}_{\mathbb{R}}(x, x)$ to be at most linear. This can only be the case if the interpretation $\mathsf{h}_{\mathbb{R}}$ contains no terms of degree two or higher. In other words, $\mathsf{h}_{\mathbb{R}}(x, y) = h_1 x + h_2 y + h_0$, where $h_0 \in \mathbb{R}_0$ and $h_1, h_2 \geqslant_{\mathbb{R}} 1$ according to Lemma 3.35. Since $\mathsf{s}_{\mathbb{R}}(x) = x + s_0$, compatibility with (4.13) implies $h_1 = 1$, and compatibility with (4.12) implies $h_2 = 1$; thus, $\mathsf{h}_{\mathbb{R}}(x, y) = x + y + h_0$.

Using the obtained information in the compatibility constraint associated with (4.14), we derive the inequality

$$\mathsf{g}_{\mathbb{R}}(x) + h_0 >_{\mathbb{R}_0,\delta} \mathsf{f}_{\mathbb{R}}(x + s_0) - \mathsf{f}_{\mathbb{R}}(x)$$

which must hold for all non-negative real numbers x. But this can only be the case if $\deg(\mathsf{g}_{\mathbb{R}}(x) + h_0) \geqslant \deg(\mathsf{f}_{\mathbb{R}}(x + s_0) - \mathsf{f}_{\mathbb{R}}(x))$, which simplifies to $1 \geqslant \deg(\mathsf{f}_{\mathbb{R}}(x)) - 1$ since $s_0 \neq 0$ and $\mathsf{f}_{\mathbb{R}}$ is not a constant polynomial (cf. (4.15)). Consequently, $\mathsf{f}_{\mathbb{R}}$ must be a quadratic polynomial function, that is, $\mathsf{f}_{\mathbb{R}}(x) = ax^2 + bx + c$ with $a >_{\mathbb{R}} 0$, $c \geqslant_{\mathbb{R}} 0$ and $a\delta + b \geqslant_{\mathbb{R}} 1$ according to Lemma 3.38.

Next we consider the compatibility constraint associated with the rule (4.8), from which we deduce an important auxiliary result. After unraveling the definitions of $>_{\mathbb{R}_0,\delta}$ and the interpretation functions, this constraint simplifies to

$$4as_0 x + 4as_0^2 + 2bs_0 \geqslant_{\mathbb{R}} 2g_1 x + g_1 h_0 + g_0 + h_0 + \delta \quad \text{for all } x \in \mathbb{R}_0,$$

which implies the following condition on the respective leading coefficients: $4as_0 \geqslant_{\mathbb{R}} 2g_1$; from this and $g_1 >_{\mathbb{R}} 2$ we conclude

$$as_0 >_{\mathbb{R}} 1 \tag{4.17}$$

and note that $as_0 = \mathsf{f}'_{\mathbb{R}}(\frac{s_0}{2}) - \mathsf{f}'_{\mathbb{R}}(0)$, where $\mathsf{f}'_{\mathbb{R}}$ denotes the derivative of $\mathsf{f}_{\mathbb{R}}$. Hence, the expression as_0 characterizes the change of the slopes of the tangents to $\mathsf{f}_{\mathbb{R}}$

at the points $(0, \mathsf{f}_{\mathbb{R}}(0))$ and $(\frac{s_0}{2}, \mathsf{f}_{\mathbb{R}}(\frac{s_0}{2}))$, and thus (4.17) actually sets a lower bound on the growth of $\mathsf{f}_{\mathbb{R}}$.

Now let us consider the combined compatibility constraint imposed by the rules (4.4) and (4.6), namely $\mathsf{0}_{\mathbb{R}} + 2s_0 >_{\mathbb{R}_0,\delta} \mathsf{f}_{\mathbb{R}}(\mathsf{s}_{\mathbb{R}}(\mathsf{0}_{\mathbb{R}})) >_{\mathbb{R}_0,\delta} \mathsf{0}_{\mathbb{R}}$, which implies $\mathsf{0}_{\mathbb{R}} + 2s_0 \geqslant_{\mathbb{R}} \mathsf{0}_{\mathbb{R}} + 2\delta$ by definition of $>_{\mathbb{R}_0,\delta}$. Thus, we obtain $s_0 \geqslant_{\mathbb{R}} \delta$. In fact, we even have $s_0 = \delta$, which can be derived from the compatibility constraint of rule (4.4) using the conditions $s_0 \geqslant_{\mathbb{R}} \delta$, $a\delta + b \geqslant_{\mathbb{R}} 1$ and $as_0 + b \geqslant_{\mathbb{R}} 1$, the combination of the former two conditions:

$$
\begin{array}{rcl}
\mathsf{0}_{\mathbb{R}} + 2s_0 & >_{\mathbb{R}_0,\delta} & \mathsf{f}_{\mathbb{R}}(\mathsf{s}_{\mathbb{R}}(\mathsf{0}_{\mathbb{R}})) \\
\mathsf{0}_{\mathbb{R}} + 2s_0 - \delta & \geqslant_{\mathbb{R}} & \mathsf{f}_{\mathbb{R}}(\mathsf{s}_{\mathbb{R}}(\mathsf{0}_{\mathbb{R}})) \\
& = & a(\mathsf{0}_{\mathbb{R}} + s_0)^2 + b(\mathsf{0}_{\mathbb{R}} + s_0) + c \\
& = & a\mathsf{0}_{\mathbb{R}}^2 + \mathsf{0}_{\mathbb{R}}(2as_0 + b) + as_0^2 + bs_0 + c \\
& \geqslant_{\mathbb{R}} & a\mathsf{0}_{\mathbb{R}}^2 + \mathsf{0}_{\mathbb{R}} + as_0^2 + bs_0 + c \\
& \geqslant_{\mathbb{R}} & \mathsf{0}_{\mathbb{R}} + as_0^2 + bs_0 \\
& \geqslant_{\mathbb{R}} & \mathsf{0}_{\mathbb{R}} + as_0^2 + (1 - a\delta)s_0 \\
& = & \mathsf{0}_{\mathbb{R}} + as_0(s_0 - \delta) + s_0
\end{array}
$$

Hence, we conclude that $\mathsf{0}_{\mathbb{R}} + 2s_0 - \delta \geqslant_{\mathbb{R}} \mathsf{0}_{\mathbb{R}} + as_0(s_0 - \delta) + s_0$, or equivalently, $s_0 - \delta \geqslant_{\mathbb{R}} as_0(s_0 - \delta)$. Yet because of (4.17) and since $s_0 \geqslant_{\mathbb{R}} \delta$, this inequality can only be satisfied if:

$$s_0 = \delta \tag{4.18}$$

This result has immediate consequences concerning the interpretation of the constant $\mathsf{0}$. To see this, let us consider the rule (4.12), whose compatibility constraint simplifies to $s_0 \geqslant_{\mathbb{R}} \mathsf{0}_{\mathbb{R}} + h_0 + \delta$. Because of (4.18) and the fact that $\mathsf{0}_{\mathbb{R}}$ and h_0 must be non-negative, the only possibility is $\mathsf{0}_{\mathbb{R}} = h_0 = 0$.

Moreover, condition (4.18) is the key to the proof of this lemma. To this end, we consider the inequalities

$$
\begin{array}{rlrl}
s_0 >_{\mathbb{R}_0,s_0} & \mathsf{f}_{\mathbb{R}}(0) & & \\
2s_0 >_{\mathbb{R}_0,s_0} & \mathsf{f}_{\mathbb{R}}(s_0) & \qquad \mathsf{f}_{\mathbb{R}}(s_0) >_{\mathbb{R}_0,s_0} & 0 \\
7s_0 >_{\mathbb{R}_0,s_0} & \mathsf{f}_{\mathbb{R}}(2s_0) & \qquad \mathsf{f}_{\mathbb{R}}(2s_0) >_{\mathbb{R}_0,s_0} & 5s_0
\end{array}
$$

arising from compatibility with (4.3) – (4.7). By definition of $>_{\mathbb{R}_0,s_0}$, these inequalities induce the following system of equations:

$$\mathsf{f}_{\mathbb{R}}(0) = 0 \qquad \mathsf{f}_{\mathbb{R}}(s_0) = s_0 \qquad \mathsf{f}_{\mathbb{R}}(2s_0) = 6s_0$$

After unraveling the definition of $\mathsf{f}_{\mathbb{R}}$ and substituting $z := as_0$, we get a system of linear equations in the unknowns z, b and c

$$c = 0 \qquad z + b = 1 \qquad 4z + 2b = 6$$

which has the unique solution $z = 2$, $b = -1$ and $c = 0$. Hence, $\mathsf{f}_{\mathbb{R}}$ must have the shape $\mathsf{f}_{\mathbb{R}}(x) = ax^2 - x = ax\left(x - \frac{1}{a}\right)$ in every compatible polynomial interpretation over $\mathbb{R}$. However, this function is not a permissible interpretation for the function symbol f because it is not well-defined over $\mathbb{R}_0$. In particular, it is negative in the open interval $\left(0, \frac{1}{a}\right)$; e.g., $\mathsf{f}_{\mathbb{R}}\left(\frac{1}{2a}\right) = -\frac{1}{4a}$. Therefore, the TRS $\mathcal{R}_2$ is not polynomially terminating over $\mathbb{R}$. □

Remark 4.16. In this proof, the interpretation of f is fixed to $\mathsf{f}_{\mathbb{R}}(x) = ax^2 - x$, which violates well-definedness over $\mathbb{R}_0$. However, this function is obviously well-defined over $\mathbb{R}_m$ for a properly chosen negative real number m. So what happens if we take this $\mathbb{R}_m$ instead of $\mathbb{R}_0$ as the carrier of a polynomial interpretation? To answer this question, let us consider some negative real number $x_0 \in \mathbb{R}_m$. Then $\mathsf{f}_{\mathbb{R}}(x_0) >_{\mathbb{R}} 0$ such that $\mathsf{f}_{\mathbb{R}}(\delta) - \mathsf{f}_{\mathbb{R}}(x_0) = \delta - \mathsf{f}_{\mathbb{R}}(x_0) <_{\mathbb{R}} \delta$, which means that $\mathsf{f}_{\mathbb{R}}$ violates monotonicity with respect to the order $>_{\mathbb{R}_m,\delta}$.

Moreover, note that, by construction, $\mathcal{R}_2$ is an example of a TRS that can be proved terminating by a polynomial interpretation over $\mathbb{N}$ with negative coefficients, but cannot be proved polynomially terminating over $\mathbb{N}$ using only non-negative coefficients.

The previous lemma, together with Corollary 4.12, yields the following corollary.

Corollary 4.17. *The TRS $\mathcal{R}_2$ is not polynomially terminating over $\mathbb{Q}$.* □

Finally, combining the material presented above, we obtain the following corollary, the main result of this subsection.

Corollary 4.18. *There are TRSs that are polynomially terminating over $\mathbb{N}$ but not over $\mathbb{R}$ or $\mathbb{Q}$.* □

We conclude this subsection with a remark on the actual choice of the polynomial serving as the interpretation of the function symbol f.

Remark 4.19. As explained at the beginning of this subsection, the TRS $\mathcal{R}_2$ was designed to enforce an interpretation for f, which is permissible in a polynomial interpretation over $\mathbb{N}$ but not over $\mathbb{R}$ ($\mathbb{Q}$). The interpretation of our choice was the polynomial $2x^2 - x$. However, we could have chosen any other polynomial as long as it is well-defined and strictly monotone over $\mathbb{N}$ but not over $\mathbb{R}_0$ ($\mathbb{Q}_0$), for example, the polynomial $5x^3 - 15x^2 + 11x$, whose graph is depicted in Figure 4.3. The methods introduced in this subsection are general enough to handle any such polynomial. So the actual choice is not that important.

4.3.3 Polynomial Termination over $\mathbb{N}$ and $\mathbb{R}$ vs. $\mathbb{Q}$

This subsection is devoted to showing that polynomial termination over $\mathbb{N}$ and $\mathbb{R}$ does not imply polynomial termination over $\mathbb{Q}$. The proof is constructive, so we give a concrete TRS having the desired properties. In order to motivate the construction underlying this particular system, let us consider the following quantified polynomial inequality

$$\forall x \ (2x^2 - x) \cdot p(a) \geqslant 0 \qquad (*)$$

where $p \in \mathbb{Z}[a]$ is a univariate integer polynomial in the indeterminate a, all of whose roots are assumed to be irrational, and which is positive for some non-negative integer value of a. To be concrete, let us consider $p(a) = a^2 - 2$ and try to satisfy ($*$) in $\mathbb{N}$, $\mathbb{Q}_0$ and $\mathbb{R}_0$, respectively. First, we observe that $a := \sqrt{2}$

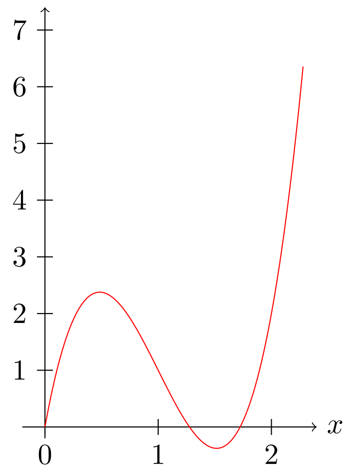

Figure 4.3: The polynomial function $x \mapsto 5x^3 - 15x^2 + 11x$.

is a satisfying assignment in $\mathbb{R}_0$. Besides, ($*$) is also satisfiable in $\mathbb{N}$ by assigning $a := 2$, for example, and observing that the polynomial $2x^2 - x$ is non-negative for all $x \in \mathbb{N}$. However, ($*$) cannot be satisfied in $\mathbb{Q}_0$ as non-negativeness of $2x^2 - x$ does not hold for all $x \in \mathbb{Q}_0$ and p has no rational roots. To sum up, ($*$) is satisfiable in $\mathbb{N}$ and $\mathbb{R}_0$ but not in $\mathbb{Q}_0$. Thus, the basic idea now is to design a TRS containing some rewrite rule whose compatibility constraint reduces to a polynomial inequality similar in nature to ($*$). For this purpose, we rewrite the inequality $(2x^2 - x) \cdot (a^2 - 2) \geqslant 0$ to

$$2a^2x^2 + 2x \geqslant 4x^2 + a^2x$$

because now both the left- and right-hand side can be viewed as a composition of several polynomial functions, all of which are strictly monotone and well-defined. In particular, we identify the following constituents: $\mathsf{h}(x, y) = x + y$, $\mathsf{r}(x) = 2x$, $\mathsf{p}(x) = x^2$ and $\mathsf{k}(x) = a^2x$. Thus, the above inequality can be written in the form

$$\mathsf{h}(\mathsf{r}(\mathsf{k}(\mathsf{p}(x))), \mathsf{r}(x)) \geqslant \mathsf{h}(\mathsf{r}(\mathsf{r}(\mathsf{p}(x))), \mathsf{k}(x)) \tag{$**$}$$

which can easily be modeled as a rewrite rule. (Note that $\mathsf{r}(x)$ is not really necessary as $\mathsf{r}(x) = \mathsf{h}(x, x)$, but it gives rise to a shorter encoding.) Then we also need rewrite rules that enforce the desired interpretations for the function symbols h, r, p and k. For this purpose, we leverage the techniques presented in the previous subsection, in particular the method of polynomial interpolation. The resulting TRS, which we will refer to as $\mathcal{R}_3$, is shown in Table 4.2. Each of the blocks in this table serves a specific purpose. The largest block consists of the rules (4.19) – (4.28) and is basically a slightly modified version of the TRS $\mathcal{R}_2$ of the previous subsection. These rules ensure that the symbol s has the semantics of a successor function $x \mapsto x + s_0$. Moreover, for any compatible (and strictly monotone) polynomial interpretation over $\mathbb{Q}$ (resp. $\mathbb{R}$), it is guaranteed that s_0 is equal to δ, the minimal step width of the order $>_{\mathbb{Q}_0,\delta}$ (resp. $>_{\mathbb{R}_0,\delta}$). In

$$\begin{aligned}
\mathsf{f}(\mathsf{g}(x)) &\to \mathsf{g}^2(\mathsf{f}(x)) && (4.19)\\
\mathsf{g}(\mathsf{s}(x)) &\to \mathsf{s}^2(\mathsf{g}(x)) && (4.20)\\
\mathsf{s}(x) &\to \mathsf{h}(0, x) && (4.21)\\
\mathsf{s}(x) &\to \mathsf{h}(x, 0) && (4.22)\\
\mathsf{f}(0) &\to 0 && (4.23)\\
\mathsf{s}^3(0) &\to \mathsf{f}(\mathsf{s}(0)) && (4.24)\\
\mathsf{f}(\mathsf{s}(0)) &\to \mathsf{s}(0) && (4.25)\\
\mathsf{h}(\mathsf{f}(x), \mathsf{g}(x)) &\to \mathsf{f}(\mathsf{s}(x)) && (4.26)\\
\mathsf{f}(\mathsf{s}^2(x)) &\to \mathsf{h}(\mathsf{f}(x), \mathsf{g}(\mathsf{h}(x, x))) && (4.27)\\
\mathsf{g}(x) &\to \mathsf{h}(\mathsf{h}(\mathsf{h}(\mathsf{h}(x, x), x), x), x) && (4.28)
\end{aligned}$$

$$\begin{aligned}
\mathsf{s}(0) &\to \mathsf{r}(0) && (4.34)\\
\mathsf{s}^3(0) &\to \mathsf{r}(\mathsf{s}(0)) && (4.35)\\
\mathsf{r}(\mathsf{s}(0)) &\to \mathsf{s}(0) && (4.36)\\
\mathsf{g}(x) &\to \mathsf{r}(x) && (4.37)
\end{aligned}$$

$$\begin{aligned}
\mathsf{s}(0) &\to \mathsf{k}(0) && (4.29)\\
\mathsf{s}^2(\mathsf{p}^2(\mathsf{a})) &\to \mathsf{s}(\mathsf{k}(\mathsf{p}(\mathsf{a}))) && (4.30)\\
\mathsf{s}(\mathsf{k}(\mathsf{p}(\mathsf{a}))) &\to \mathsf{p}^2(\mathsf{a}) && (4.31)\\
\mathsf{g}(x) &\to \mathsf{k}(x) && (4.32)\\
\mathsf{a} &\to 0 && (4.33)
\end{aligned}$$

$$\begin{aligned}
\mathsf{s}(0) &\to \mathsf{p}(0) && (4.38)\\
\mathsf{s}^2(0) &\to \mathsf{p}(\mathsf{s}(0)) && (4.39)\\
\mathsf{p}(\mathsf{s}(0)) &\to 0 && (4.40)\\
\mathsf{s}^5(0) &\to \mathsf{p}(\mathsf{s}^2(0)) && (4.41)\\
\mathsf{p}(\mathsf{s}^2(0)) &\to \mathsf{s}^3(0) && (4.42)\\
\mathsf{h}(\mathsf{p}(x), \mathsf{g}(x)) &\to \mathsf{p}(\mathsf{s}(x)) && (4.43)
\end{aligned}$$

$$\mathsf{s}(\mathsf{h}(\mathsf{r}(\mathsf{k}(\mathsf{p}(x))), \mathsf{r}(x))) \to \mathsf{h}(\mathsf{r}^2(\mathsf{p}(x)), \mathsf{k}(x)) \qquad (4.44)$$

Table 4.2: The TRS $\mathcal{R}_3$.

the previous subsection, these conditions were identified as the key requirements for the method of polynomial interpolation to work in this setting. Finally, this block also enforces $\mathsf{h}(x, y) = x + y$. The next block, consisting of the rules (4.34) – (4.37), makes use of polynomial interpolation to achieve $\mathsf{r}(x) = 2x$. Likewise, the block (4.38) – (4.43) equips the symbol p with the semantics of a squaring function. And the block (4.29) – (4.33) enforces the desired semantics for the symbol k, i.e., a linear function $x \mapsto k_1 x$ whose slope k_1 is proportional to the square of the interpretation of the constant a. Finally, (4.44) encodes the main idea presented at the beginning of this subsection (cf. (∗∗)).

Lemma 4.20. *The TRS $\mathcal{R}_3$ is polynomially terminating over $\mathbb{N}$ and $\mathbb{R}$.*

Proof. For polynomial termination over $\mathbb{N}$, the following interpretation applies:

$$0_{\mathbb{N}} = 0 \qquad \mathsf{s}_{\mathbb{N}}(x) = x + 1 \qquad \mathsf{f}_{\mathbb{N}}(x) = 3x^2 - 2x + 1 \qquad \mathsf{g}_{\mathbb{N}}(x) = 6x + 6$$
$$\mathsf{h}_{\mathbb{N}}(x, y) = x + y \qquad \mathsf{p}_{\mathbb{N}}(x) = x^2 \qquad \mathsf{r}_{\mathbb{N}}(x) = 2x \qquad \mathsf{k}_{\mathbb{N}}(x) = 4x \qquad \mathsf{a}_{\mathbb{N}} = 2$$

By Lemma 3.6 and Corollary 3.9, all interpretation functions are well-defined over $\mathbb{N}$ and strictly monotone (i.e., monotone with respect to $>_{\mathbb{N}}$). Concerning compatibility, which is easily verified, we remark that the rule (4.44) gives rise to the constraint

$$8x^2 + 2x + 1 >_{\mathbb{N}} 4x^2 + 4x \qquad \iff \qquad 4x^2 - 2x + 1 >_{\mathbb{N}} 0$$

which holds for all $x \in \mathbb{N}$. For polynomial termination over $\mathbb{R}$, we let $\delta = 1$, but we have to modify the interpretation as $4x^2 - 2x + 1 >_{\mathbb{R}_0,\delta} 0$ does not hold for all $x \in \mathbb{R}_0$. Taking $\mathsf{a}_{\mathbb{R}} = \sqrt{2}$, $\mathsf{k}_{\mathbb{R}}(x) = 2x$ and the above interpretations for the other function symbols establishes polynomial termination over $\mathbb{R}$. Note that then the constraint $4x^2 + 2x + 1 >_{\mathbb{R}_0,\delta} 4x^2 + 2x$ associated with (4.44) trivially holds, and monotonicity with respect to $>_{\mathbb{R}_0,\delta}$ and well-definedness over $\mathbb{R}_0$ follow directly from Lemmata 3.35 and 3.38. □

Lemma 4.21. *The TRS $\mathcal{R}_3$ is not polynomially terminating over $\mathbb{Q}$.*

Proof. Let us assume that $\mathcal{R}_3$ is polynomially terminating over $\mathbb{Q}$ and derive a contradiction. Adapting the reasoning in the proof of Lemma 4.15, we infer from compatibility with (a subset of) the block of rules (4.19) – (4.28) that $\mathsf{s}_{\mathbb{Q}}(x) = x + s_0$, $\mathsf{g}_{\mathbb{Q}}(x) = g_1 x + g_0$, $\mathsf{h}_{\mathbb{Q}}(x, y) = x + y + h_0$ and $\mathsf{f}_{\mathbb{Q}}(x) = ax^2 + bx + c$, subject to the following constraints:

$$s_0 >_{\mathbb{Q}} 0 \qquad g_1 >_{\mathbb{Q}} 2 \qquad g_0, h_0 \in \mathbb{Q}_0 \qquad a >_{\mathbb{Q}} 0 \qquad c \geqslant_{\mathbb{Q}} 0 \qquad a\delta + b \geqslant_{\mathbb{Q}} 1$$

Next we consider the compatibility constraints associated with (4.27) and (4.28), from which we deduce an important auxiliary result. Compatibility with (4.28) implies the condition $g_1 \geqslant_{\mathbb{Q}} 5$ on the respective leading coefficients since $\mathsf{h}_{\mathbb{Q}}(x, y) = x + y + h_0$, and compatibility with (4.27) simplifies to

$$4as_0 x + 4as_0^2 + 2bs_0 \geqslant_{\mathbb{Q}} 2g_1 x + g_1 h_0 + g_0 + h_0 + \delta \quad \text{for all } x \in \mathbb{Q}_0,$$

from which we infer $4as_0 \geqslant_{\mathbb{Q}} 2g_1$. Together with $g_1 \geqslant_{\mathbb{Q}} 5$, this implies $as_0 >_{\mathbb{Q}} 2$.

Now let us consider the combined compatibility constraint imposed by the rules (4.24) and (4.25), namely $\mathsf{0}_{\mathbb{Q}} + 3s_0 >_{\mathbb{Q}_0,\delta} \mathsf{f}_{\mathbb{Q}}(\mathsf{s}_{\mathbb{Q}}(\mathsf{0}_{\mathbb{Q}})) >_{\mathbb{Q}_0,\delta} \mathsf{0}_{\mathbb{Q}} + s_0$, which implies $\mathsf{0}_{\mathbb{Q}} + 3s_0 \geqslant_{\mathbb{Q}} \mathsf{0}_{\mathbb{Q}} + s_0 + 2\delta$ by definition of $>_{\mathbb{Q}_0,\delta}$. Thus, we obtain $s_0 \geqslant_{\mathbb{Q}} \delta$. In fact, we even have $s_0 = \delta$, which can be derived from the compatibility constraint of (4.24) using the conditions $s_0 \geqslant_{\mathbb{Q}} \delta$, $a\delta + b \geqslant_{\mathbb{Q}} 1$, $as_0 + b \geqslant_{\mathbb{Q}} 1$, the combination of the former two conditions, and $\mathsf{f}_{\mathbb{Q}}(\mathsf{0}_{\mathbb{Q}}) \geqslant_{\mathbb{Q}} \mathsf{0}_{\mathbb{Q}} + \delta$, the compatibility constraint associated with (4.23):

$$\begin{aligned}
\mathsf{0}_{\mathbb{Q}} + 3s_0 - \delta \quad &\geqslant_{\mathbb{Q}} \quad \mathsf{f}_{\mathbb{Q}}(\mathsf{s}_{\mathbb{Q}}(\mathsf{0}_{\mathbb{Q}})) \;=\; \mathsf{f}_{\mathbb{Q}}(\mathsf{0}_{\mathbb{Q}}) + 2a\mathsf{0}_{\mathbb{Q}} s_0 + as_0^2 + bs_0 \\
&\geqslant_{\mathbb{Q}} \quad \mathsf{0}_{\mathbb{Q}} + \delta + as_0^2 + bs_0 \\
&\geqslant_{\mathbb{Q}} \quad \mathsf{0}_{\mathbb{Q}} + \delta + as_0^2 + (1 - a\delta)s_0 \;=\; \mathsf{0}_{\mathbb{Q}} + s_0 + \delta + as_0(s_0 - \delta)
\end{aligned}$$

Hence, we conclude that $\mathsf{0}_{\mathbb{Q}} + 3s_0 - \delta \geqslant_{\mathbb{Q}} \mathsf{0}_{\mathbb{Q}} + s_0 + \delta + as_0(s_0 - \delta)$, or equivalently, $2(s_0 - \delta) \geqslant_{\mathbb{Q}} as_0(s_0 - \delta)$. But since $as_0 >_{\mathbb{Q}} 2$ and $s_0 \geqslant_{\mathbb{Q}} \delta$, this inequality can only hold if

$$s_0 = \delta \tag{4.45}$$

This result has immediate consequences concerning the interpretation of the constant $\mathsf{0}$. To see this, let us consider the inequality $s_0 \geqslant_{\mathbb{Q}} \mathsf{0}_{\mathbb{Q}} + h_0 + \delta$ obtained from compatibility with (4.21). Because of (4.45) and the fact that $\mathsf{0}_{\mathbb{Q}}$ and h_0 must be non-negative, the only possibility is $\mathsf{0}_{\mathbb{Q}} = h_0 = 0$.

Moreover, as in the proof of Lemma 4.15, condition (4.45) is the key to the proof of the lemma at hand. To this end, let us consider the compatibility constraints associated with the rules (4.38) – (4.42). By definition of $>_{\mathbb{Q}_0,s_0}$, these constraints give rise to the following system of equations:

$$\mathsf{p}_{\mathbb{Q}}(0) = 0 \qquad \mathsf{p}_{\mathbb{Q}}(s_0) = s_0 \qquad \mathsf{p}_{\mathbb{Q}}(2s_0) = 4s_0$$

Viewing these equations as polynomial interpolation constraints, we conclude that no linear polynomial can satisfy them (because $s_0 \neq 0$). Hence, $\mathsf{p}_{\mathbb{Q}}$ must at least be quadratic. Moreover, by compatibility with (4.43), $\mathsf{p}_{\mathbb{Q}}$ is at most quadratic (using the same reasoning as for (4.26), cf. the proof of Lemma 4.15 or (4.15)). So we let $\mathsf{p}_{\mathbb{Q}}(x) = p_2x^2 + p_1x + p_0$ in the equations above and infer the (unique) solution $p_0 = p_1 = 0$ and $p_2s_0 = 1$, i.e., $\mathsf{p}_{\mathbb{Q}}(x) = p_2x^2$ with $p_2 > 0$.

Next we consider the compatibility constraints of the rules (4.34) – (4.36), from which we deduce the interpolation constraints $\mathsf{r}_{\mathbb{Q}}(0) = 0$ and $\mathsf{r}_{\mathbb{Q}}(s_0) = 2s_0$. Because $\mathsf{g}_{\mathbb{Q}}$ is linear, $\mathsf{r}_{\mathbb{Q}}$ must be linear, too, for compatibility with (4.37). Hence, by polynomial interpolation, $\mathsf{r}_{\mathbb{Q}}(x) = 2x$. Likewise, $\mathsf{k}_{\mathbb{Q}}$ must be linear for compatibility with (4.32), i.e., $\mathsf{k}_{\mathbb{Q}}(x) = k_1x + k_0$. In particular, k_0 must be zero because of compatibility with (4.29). Then the compatibility constraints imposed by (4.30) and (4.31) yield

$$p_2^3\mathsf{a}_{\mathbb{Q}}^4 + 2s_0 - \delta \geqslant_{\mathbb{Q}} k_1p_2\mathsf{a}_{\mathbb{Q}}^2 + s_0 \geqslant_{\mathbb{Q}} p_2^3\mathsf{a}_{\mathbb{Q}}^4 + \delta$$

But $s_0 = \delta$, hence $k_1p_2\mathsf{a}_{\mathbb{Q}}^2 = p_2^3\mathsf{a}_{\mathbb{Q}}^4$, from which we obtain $k_1 = p_2^2\mathsf{a}_{\mathbb{Q}}^2$ since $\mathsf{a}_{\mathbb{Q}}$ cannot be zero due to compatibility with (4.33). In other words, $\mathsf{k}_{\mathbb{Q}}(x) = p_2^2\mathsf{a}_{\mathbb{Q}}^2x$.

Finally, we consider the compatibility constraint associated with the rule (4.44), which simplifies to

$$(2p_2x^2 - x)((p_2\mathsf{a}_{\mathbb{Q}})^2 - 2) \geqslant_{\mathbb{Q}} 0 \quad \text{for all } x \in \mathbb{Q}_0.$$

However, this inequality is unsatisfiable as the polynomial $2p_2x^2 - x$ is negative for some $x \in \mathbb{Q}_0$ and $(p_2\mathsf{a}_{\mathbb{Q}})^2 - 2$ cannot be zero because both p_2 and $\mathsf{a}_{\mathbb{Q}}$ must be rational numbers. □

Combining Lemma 4.20 and Lemma 4.21, we obtain the main result of this subsection.

Corollary 4.22. *There are TRSs that are polynomially terminating over $\mathbb{N}$ and $\mathbb{R}$ but not over $\mathbb{Q}$.* □

We conclude this section with two additional observations. First, we present alternative proofs of Theorems 4.8 and 4.9, which shows the inhabitation of the areas with the symbols $\mathbb{Q}$ and $\mathbb{R}$ in Figure 4.1. Second, we show that the use of non-linear interpretations in the proofs of the main results of this section is essential.

As far as Theorem 4.8 is concerned, we remark that we have already encountered two TRSs in the course of this thesis that prove its claim. What is more, the actual proofs have already been given as well, though not explicitly stated as such. The first of these systems is the (one-rule) TRS $\mathcal{R}_0$ of Example 2.19 in Chapter 2. This TRS is polynomially terminating over $\mathbb{Q}$ according to Lemma 2.20 but not over $\mathbb{N}$ as it is not simply terminating (cf. Section 2.2).

The other system exhibiting these properties is the TRS $\mathcal{R}_1$ introduced in Section 3.4 of Chapter 3, which is well-known in term rewriting as it evidences the fact that there are TRSs that are simply terminating but not totally terminating (cf. [50], for example). Because of this, and due to the fact that polynomial termination over $\mathbb{N}$ implies total termination, which we mentioned in Section 2.2, the TRS $\mathcal{R}_1$ cannot be polynomially terminating over $\mathbb{N}$. However, it is polynomially terminating over $\mathbb{Q}$ according to Lemma 3.40. Alternatively, the former statement can be concluded from the following observation. In every compatible polynomial interpretation over $\mathbb{N}$, we have $\mathsf{a}_\mathbb{N} > \mathsf{b}_\mathbb{N}$ or $\mathsf{a}_\mathbb{N} \leqslant \mathsf{b}_\mathbb{N}$. Strict monotonicity of $\mathsf{f}_\mathbb{N}$ and $\mathsf{g}_\mathbb{N}$ yields $\mathsf{g}_\mathbb{N}(\mathsf{a}_\mathbb{N}) > \mathsf{g}_\mathbb{N}(\mathsf{b}_\mathbb{N})$ or $\mathsf{f}_\mathbb{N}(\mathsf{a}_\mathbb{N}) \leqslant \mathsf{f}_\mathbb{N}(\mathsf{b}_\mathbb{N})$. In both cases compatibility with some rule of $\mathcal{R}_1$ is violated. So the two TRSs $\mathcal{R}_0$ and $\mathcal{R}_1$ can indeed be used to provide alternative proofs for Theorem 4.8. In particular, the corresponding proofs show that the strict inclusion holds even for *ground* TRSs. Besides, both proofs are considerably shorter and simpler than the original proof in [46, pp. 62–67].

The techniques introduced in this section also facilitate an alternative proof of Theorem 4.9. It is based on the TRS $\mathcal{R}_4$ consisting of the rules (4.9) – (4.13), (4.46) and (4.47):

$$\mathsf{k}(\mathsf{k}(x)) \to \mathsf{h}(x,x) \tag{4.46}$$

$$\mathsf{s}(\mathsf{h}(x,x)) \to \mathsf{k}(\mathsf{k}(x)) \tag{4.47}$$

By construction, this TRS is polynomially terminating over $\mathbb{R}$ but not over $\mathbb{Q}$ or $\mathbb{N}$. This is achieved by the same method as in the original proof of [46], namely by forcing the leading coefficient of the interpretation of the unary function symbol k to be an irrational number.

Lemma 4.23. *The TRS $\mathcal{R}_4$ is polynomially terminating over $\mathbb{R}$.*

Proof. By the following interpretation:

$$\delta = 1 \qquad \mathsf{0}_\mathbb{R} = 0 \qquad \mathsf{s}_\mathbb{R}(x) = x + 4 \qquad \mathsf{f}_\mathbb{R}(x) = x^2$$
$$\mathsf{g}_\mathbb{R}(x) = 3x + 5 \qquad \mathsf{h}_\mathbb{R}(x,y) = x + y \qquad \mathsf{k}_\mathbb{R}(x) = \sqrt{2}x + 1 \qquad \square$$

Informally, the reason why $\mathcal{R}_4$ is not polynomially terminating over $\mathbb{Q}$ or $\mathbb{N}$ is as follows (the formal proof will be given in the next section, where it will follow from a more general result, cf. Lemma 4.29). As in the proof of Lemma 4.15, the rules (4.9) – (4.13) ensure that the symbol s has an interpretation of the shape $\mathsf{s}_{\mathbb{Q}/\mathbb{N}}(x) = x + s_0$, while the interpretation of h is $\mathsf{h}_{\mathbb{Q}/\mathbb{N}}(x,y) = x + y + h_0$. In particular, both interpretations are linear. As a consequence, compatibility with (4.47) implies that the interpretation of k is at most linear as well, that is, $\mathsf{k}_{\mathbb{Q}/\mathbb{N}}(x) = k_1 x + k_0$. Then the compatibility constraints imposed by (4.46) and (4.47) give rise to the following conditions on the respective leading coefficients: $2 \geqslant k_1^2 \geqslant 2$. Hence, $k_1 = \sqrt{2}$, which is not a rational number. Therefore, the TRS $\mathcal{R}_4$ cannot be polynomially terminating over $\mathbb{Q}$ or $\mathbb{N}$.

Combining all results presented so far gives rise to Figure 4.4, which gives the full picture of the relationship between polynomial termination over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$. (Note that all areas are inhabited, including the trivial one lying in the intersection of $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$.)

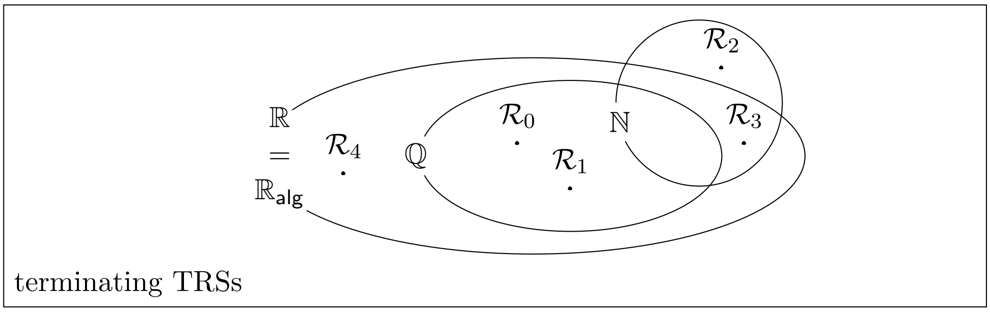

Figure 4.4: Summary.

Finally, we review our results in the context of automated termination analysis, where linear polynomial interpretations play a predominant role. This naturally raises the question as to what extent the restriction to linear polynomial interpretations influences the relationships depicted in Figure 4.4, and in what follows we shall see that it changes considerably. More precisely, the areas inhabited by the TRSs $\mathcal{R}_2$ and $\mathcal{R}_3$ become empty, such that polynomial termination by a linear polynomial interpretation over $\mathbb{N}$ implies polynomial termination by a linear polynomial interpretation over $\mathbb{Q}$, which in turn implies polynomial termination by a linear polynomial interpretation over $\mathbb{R}$. The latter follows directly from Corollary 4.12 and Remark 4.13, whereas the former is shown below.

Lemma 4.24. *Polynomial termination by a linear polynomial interpretation over $\mathbb{N}$ implies polynomial termination by a linear polynomial interpretation over $\mathbb{Q}$.*

Proof. Let $\mathcal{R}$ be a TRS that is compatible with a strictly monotone polynomial interpretation $\mathcal{I}$ over $\mathbb{N}$, where each n-ary function symbol f is interpreted by a linear polynomial $a_n x_n + \cdots + a_1 x_1 + a_0$ with integer coefficients. We show that the same interpretation also establishes polynomial termination over $\mathbb{Q}$ with the value of δ set to one. By Lemma 3.6, the coefficients of the respective interpretation functions have to satisfy $a_0 \geqslant 0$ and $a_i \geqslant 1$ for all $i \in \{1, \dots, n\}$ in order to guarantee strict monotonicity and well-definedness over $\mathbb{N}$. Hence, by Lemma 3.35, we also have well-definedness over $\mathbb{Q}_0$ and monotonicity with respect to the order $>_{\mathbb{Q}_0,1}$. Moreover, by compatibility of $\mathcal{I}$ and $\mathcal{R}$, each rewrite rule $\ell \to r \in \mathcal{R}$ satisfies

$$P_\ell - P_r >_{\mathbb{N}} 0 \quad \text{for all } x_1, \dots, x_m \in \mathbb{N}. \tag{4.48}$$

Due to the fact that linear polynomial functions are closed under addition and composition, $P_\ell - P_r$ is a linear integer polynomial $c_m x_m + \cdots + c_1 x_1 + c_0$, such that (4.48) holds if and only if $c_0 \geqslant 1$ and $c_i \geqslant 0$ for all $i \in \{1, \dots, m\}$. However, then we also have

$$P_\ell - P_r >_{\mathbb{Q}_0,1} 0 \quad \text{for all } x_1, \dots, x_m \in \mathbb{Q}_0,$$

which shows that the TRS $\mathcal{R}$ is compatible with the strictly monotone linear polynomial interpretation $(\mathcal{I}, \delta) = (\mathcal{I}, 1)$ over $\mathbb{Q}$. □

Hence, polynomial termination by a linear polynomial interpretation over $\mathbb{N}$ implies polynomial termination by a linear polynomial interpretation over $\mathbb{Q}$, which in turn implies polynomial termination by a linear polynomial interpretation over $\mathbb{R}$, and both inclusions are proper due to the results of [46], which were obtained using linear polynomial interpretations.

4.4 Incremental Polynomial Termination

In this section, we consider the possibility of establishing termination by using polynomial interpretations in an incremental way (in the sense of Definition 2.9). As in the previous section, we give the full picture of the relationship between the three notions of incremental polynomial termination over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$, showing that it is essentially the same as the one depicted in Figure 4.4 for direct polynomial termination. We can even reuse some of the TRSs mentioned in that figure (resp. the results established on top of them), namely the TRSs $\mathcal{R}_0$, $\mathcal{R}_1$ and $\mathcal{R}_4$, for this purpose. However, we have to replace the TRSs $\mathcal{R}_2$ and $\mathcal{R}_3$ as the proofs of Lemmata 4.15 and 4.21 break down if we allow incremental termination proofs. In more detail, the proof of Lemma 4.15 does not extend because the TRS $\mathcal{R}_2$ is incrementally polynomially terminating over $\mathbb{Q}$ (and thus also over $\mathbb{R}$ according to Lemma 4.10 and Corollary 4.27 below). This can be seen by considering the interpretation[3]

$$\mathsf{0}_{\mathbb{Q}} = 0 \quad \mathsf{s}_{\mathbb{Q}}(x) = x + 1 \quad \mathsf{f}_{\mathbb{Q}}(x) = x^2 + x \quad \mathsf{g}_{\mathbb{Q}}(x) = 2x + \tfrac{5}{2} \quad \mathsf{h}_{\mathbb{Q}}(x, y) = x + y$$

with $\delta = 1$. Using the results of Chapter 3, it is easy to verify that the latter is both weakly and strictly monotone, thus constituting an extended monotone algebra. The rewrite rules of $\mathcal{R}_2$ give rise to the following inequalities:

$$\begin{aligned}
1 &\geqslant_{\mathbb{Q}} 0 & 4x^2 + 12x + \tfrac{35}{4} &\geqslant_{\mathbb{Q}} 4x^2 + 4x + \tfrac{15}{2} \\
2 &\geqslant_{\mathbb{Q}} 2 & 2x + \tfrac{9}{2} &\geqslant_{\mathbb{Q}} 2x + \tfrac{9}{2} \\
7 &\geqslant_{\mathbb{Q}} 6 & 2x + \tfrac{5}{2} &\geqslant_{\mathbb{Q}} 2x \\
2 &\geqslant_{\mathbb{Q}} 0 & x + 1 &\geqslant_{\mathbb{Q}} x \\
6 &\geqslant_{\mathbb{Q}} 5 & x + 1 &\geqslant_{\mathbb{Q}} x \\
x^2 + 5x + 6 &\geqslant_{\mathbb{Q}} x^2 + 5x + \tfrac{5}{2} & x^2 + 3x + \tfrac{5}{2} &\geqslant_{\mathbb{Q}} x^2 + 3x + 2
\end{aligned}$$

Removing the rules from $\mathcal{R}_2$ for which the corresponding constraint remains true after strengthening $\geqslant_{\mathbb{Q}}$ to $>_{\mathbb{Q}_0,\delta}$ leaves us with (4.4), (4.10) and (4.14), which are easily handled, e.g. by the interpretation

$$\delta = 1 \quad \mathsf{0}_{\mathbb{Q}} = 0 \quad \mathsf{s}_{\mathbb{Q}}(x) = x + 1 \quad \mathsf{f}_{\mathbb{Q}}(x) = x \quad \mathsf{g}_{\mathbb{Q}}(x) = 3x \quad \mathsf{h}_{\mathbb{Q}}(x, y) = x + y + 2$$

Similarly, the TRS $\mathcal{R}_3$ can be shown to be incrementally polynomially terminating over $\mathbb{Q}$.

As the notions of polynomial termination and incremental polynomial termination coincide for one-rule TRSs, the alternative proof of Theorem 4.8 on the basis of the one-rule TRS $\mathcal{R}_0$ given in the previous section immediately yields the following result.

[3] I thank Harald Zankl for finding this interpretation.

Corollary 4.25. *There are TRSs that are polynomially terminating over* $\mathbb{Q}^*$ *but not over* $\mathbb{N}^*$. □

In fact, the TRS $\mathcal{R}_0$ proves the stronger statement that there are TRSs which are polynomially terminating over $\mathbb{Q}$ but not incrementally polynomially terminating over $\mathbb{N}$.

Remark 4.26. Also the proof of Theorem 4.8 based on the TRS $\mathcal{R}_1$ can be used to establish this result (using the observation that the interpretations of the constants a and b must not be equal in order to make progress, that is, in order to remove at least one rule).

In analogy to Corollary 4.12, incremental polynomial termination over $\mathbb{Q}$ implies incremental polynomial termination over $\mathbb{R}$. This result is a direct consequence of Lemma 4.10 (as was the case for Corollary 4.12).

Corollary 4.27. *If a TRS is polynomially terminating over* $\mathbb{Q}^*$*, then it is also polynomially terminating over* $\mathbb{R}^*$. □

Next we use the TRS $\mathcal{R}_4$ of the previous section to show that the converse of Corollary 4.27 does not hold. From Lemma 4.23 we already know that $\mathcal{R}_4$ is polynomially terminating over $\mathbb{R}$. So it remains to show that it is not incrementally polynomially terminating over $\mathbb{Q}$. We also show that it is neither incrementally polynomially terminating over $\mathbb{N}$. But first we present the following auxiliary result on a subset of its rules.

Lemma 4.28. *Let* $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$*, and let* $\mathcal{P}$ *be a strictly monotone polynomial interpretation over* D *that is weakly compatible with the rules* (4.9) – (4.13). *Then the interpretations of the symbols* s*,* h *and* g *have the shape*

$$\mathsf{s}_D(x) = x + s_0 \qquad \mathsf{h}_D(x, y) = x + y + h_0 \qquad \mathsf{g}_D(x) = g_1 x + g_0$$

where all coefficients are non-negative and $g_1 \geqslant 2$*. Moreover, the interpretation of the symbol* f *is at least quadratic.*

In order to reuse the proof of this result in the next section, where we only require weak (rather than strict) monotonicity of the interpretation $\mathcal{P}$, the account presented here is as general as possible (i.e., making the least assumptions on $\mathcal{P}$).

Proof. Let $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$, and let $\mathcal{P}$ be a polynomial interpretation over D that is weakly compatible with (4.9) – (4.13), and in which the unary symbols f, g and s are interpreted by non-constant polynomials $\mathsf{f}_D(x)$, $\mathsf{g}_D(x)$ and $\mathsf{s}_D(x)$. (Note that strict monotonicity of $\mathcal{P}$ obviously implies these conditions.) Then the degrees of these polynomials must be at least 1, such that weak compatibility with (4.10) implies

$$\deg(\mathsf{g}_D(x)) \cdot \deg(\mathsf{s}_D(x)) \geqslant \deg(\mathsf{s}_D(x)) \cdot \deg(\mathsf{s}_D(x)) \cdot \deg(\mathsf{g}_D(x))$$

which simplifies to $\deg(\mathsf{s}_D(x)) \leqslant 1$. Hence, we obtain $\deg(\mathsf{s}_D(x)) = 1$ and, by applying the same reasoning to (4.9), $\deg(\mathsf{g}_D(x)) = 1$.

So the function symbols s and g must be interpreted by linear polynomials $\mathsf{s}_D(x) = s_1x + s_0$ and $\mathsf{g}_D(x) = g_1x + g_0$, where $s_0, s_1, g_0, g_1 \in D_0$ due to well-definedness over D_0 and $s_1, g_1 > 0$ to make them non-constant. Then the weak compatibility constraint imposed by (4.10) gives rise to the inequality

$$g_1s_1x + g_1s_0 + g_0 \geqslant_{D_0} s_1^2g_1x + s_1^2g_0 + s_1s_0 + s_0 \tag{4.49}$$

which must hold for all $x \in D_0$. This implies the following condition on the respective leading coefficients: $g_1s_1 \geqslant s_1^2g_1$. Due to $s_1, g_1 > 0$, this can only hold if $s_1 \leqslant 1$. Now suppose that the function symbol f were also interpreted by a linear polynomial f_D. Then we could apply the same reasoning to the rule (4.9) because it is structurally equivalent to (4.10), thus inferring $g_1 \leqslant 1$. So f_D cannot be linear if $g_1 > 1$.

Next we consider the rewrite rules (4.11), (4.12) and (4.13). As g_D is linear, weak compatibility with (4.11) implies that the function $\mathsf{h}_D(x, x)$ is at most linear as well. This can only be the case if the interpretation h_D is a linear polynomial function $\mathsf{h}_D(x, y) = h_1x + h_2y + h_0$, where $h_0, h_1, h_2 \in D_0$ due to well-definedness over D_0. Since $\mathsf{s}_D(x) = s_1x+s_0$, weak compatibility with (4.13) implies $s_1 \geqslant h_1$, and weak compatibility with (4.12) implies $s_1 \geqslant h_2$. Similarly, we obtain $g_1 \geqslant h_1 + h_2$ from weak compatibility with (4.11).

Now if $s_1, h_1, h_2 \geqslant 1$, conditions that are implied by strict monotonicity of s_D and h_D according to Lemmata 3.6 (for $D = \mathbb{N}$) and 3.35 (for $D \in \{\mathbb{Q}, \mathbb{R}\}$), then we obtain $s_1 = h_1 = h_2 = 1$ and $g_1 \geqslant 2$, such that

$$\mathsf{s}_D(x) = x + s_0 \qquad \mathsf{h}_D(x, y) = x + y + h_0 \qquad \mathsf{g}_D(x) = g_1x + g_0$$

with $g_1 \geqslant 2$, which shows that f_D cannot be linear. Due to the fact that all of the above assumptions (on the interpretations of the symbols f, g, h and s) follow from strict monotonicity of $\mathcal{P}$, this concludes the proof. □

With the help of this lemma it is easy to show that the TRS $\mathcal{R}_4$ is not incrementally polynomially terminating over $\mathbb{Q}$ or $\mathbb{N}$.

Lemma 4.29. *The TRS $\mathcal{R}_4$ is not polynomially terminating over $\mathbb{Q}^*$ or $\mathbb{N}^*$.*

Proof. Let $D \in \{\mathbb{N}, \mathbb{Q}\}$, and let $\mathcal{P}$ be a strictly monotone polynomial interpretation over D that is weakly compatible with $\mathcal{R}_4$. Then, by Lemma 4.28, the interpretations of the symbols s, h and g have the shape

$$\mathsf{s}_D(x) = x + s_0 \qquad \mathsf{h}_D(x, y) = x + y + h_0 \qquad \mathsf{g}_D(x) = g_1x + g_0$$

As the interpretations of the symbols s and h are linear, weak compatibility with (4.47) implies that the interpretation of k is at most linear as well. Then, letting $\mathsf{k}_D(x) = k_1x + k_0$, the weak compatibility constraints associated with (4.46) and (4.47) give rise to the following conditions on the respective leading coefficients: $2 \geqslant k_1^2 \geqslant 2$. Hence, $k_1 = \sqrt{2}$, which is not a rational number. So we conclude that there is no strictly monotone polynomial interpretation over $\mathbb{N}$ or $\mathbb{Q}$ that is weakly compatible with the TRS $\mathcal{R}_4$. This implies that $\mathcal{R}_4$ is not incrementally polynomially terminating over $\mathbb{N}$ or $\mathbb{Q}$. □

Combining Lemma 4.23 and Lemma 4.29, we obtain the following result.

Corollary 4.30. *There are TRSs that are polynomially terminating over $\mathbb{R}^*$ but not over $\mathbb{Q}^*$ or $\mathbb{N}^*$.* □

As a further consequence of Lemmata 4.23 and 4.29, we see that the TRS $\mathcal{R}_4$ is polynomially terminating over $\mathbb{R}$ but not over $\mathbb{Q}$ or $\mathbb{N}$, which provides the alternative proof of Theorem 4.9 that was already sketched in the previous section.

4.4.1 Incremental Polynomial Termination over $\mathbb{N}$ and $\mathbb{R}$ vs. $\mathbb{Q}$

Next we establish the analogon of Corollary 4.22 in the incremental setting. That is, we show that incremental polynomial termination over $\mathbb{N}$ and $\mathbb{R}$ does not imply incremental polynomial termination over $\mathbb{Q}$. Again, we give a concrete TRS having the desired properties, but unfortunately, as was already mentioned in the introduction of this section, we cannot reuse the TRS $\mathcal{R}_3$ directly. Nevertheless, we can and do reuse the principle idea underlying the construction of $\mathcal{R}_3$ (cf. ($**$)). However, we use a different method than polynomial interpolation in order to enforce the desired interpretations for the involved function symbols. To this end, let us consider the (auxiliary) TRS $\mathcal{S}$ consisting of the rules (4.9) – (4.13) and (4.50) – (4.55):

$$\mathsf{k}(x) \to \mathsf{h}(x, x) \tag{4.50}$$
$$\mathsf{s}^3(\mathsf{h}(x, x)) \to \mathsf{k}(x) \tag{4.51}$$
$$\mathsf{h}(\mathsf{f}(x), \mathsf{k}(x)) \to \mathsf{f}(\mathsf{s}(x)) \tag{4.52}$$
$$\mathsf{f}(\mathsf{s}^2(x)) \to \mathsf{h}(\mathsf{f}(x), \mathsf{k}(\mathsf{h}(x, x))) \tag{4.53}$$
$$\mathsf{f}(\mathsf{s}(x)) \to \mathsf{h}(\mathsf{f}(x), \mathsf{s}(\mathsf{0})) \tag{4.54}$$
$$\mathsf{s}^2(\mathsf{0}) \to \mathsf{h}(\mathsf{f}(\mathsf{s}(\mathsf{0})), \mathsf{s}(\mathsf{0})) \tag{4.55}$$

The purpose of this TRS is to equip the symbol s (f) with the semantics of a successor (squaring) function and to ensure that the interpretation of the symbol h corresponds to the addition of two numbers. Besides, this TRS will not only be helpful in this subsection but also in the next one.

Lemma 4.31. *Let $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$, and let $\mathcal{P}$ be a strictly monotone polynomial interpretation over D that is weakly compatible with the TRS $\mathcal{S}$. Then*

$$\mathsf{0}_D = 0 \quad \mathsf{s}_D(x) = x + s_0 \quad \mathsf{h}_D(x, y) = x + y$$
$$\mathsf{g}_D(x) = g_1 x + g_0 \quad \mathsf{k}_D(x) = 2x + k_0 \quad \mathsf{f}_D(x) = a x^2$$

where $a s_0 = 1$, $g_1 \geqslant 2$ and all coefficients are non-negative.

Proof. By Lemma 4.28, the interpretations of the symbols s, h and g have the shape $\mathsf{s}_D(x) = x + s_0$, $\mathsf{h}_D(x, y) = x + y + h_0$ and $\mathsf{g}_D(x) = g_1 x + g_0$, where all coefficients are non-negative and $g_1 \geqslant 2$. Moreover, the interpretation of f is at least quadratic.

Applying this partial interpretation in (4.50) and (4.51), we obtain, by weak compatibility, the inequalities

$$2x + h_0 + 3s_0 \geqslant_{D_0} \mathsf{k}_D(x) \geqslant_{D_0} 2x + h_0 \quad \text{for all } x \in D_0,$$

which imply $\mathsf{k}_D(x) = 2x + k_0$ with $k_0 \geqslant 0$ (due to well-definedness over D_0).

Next we consider the rule (4.53) from which we infer that $\mathsf{s}_D(x) \neq x$ because otherwise weak compatibility would be violated; hence, $s_0 > 0$. Then, by weak compatibility with (4.52), we obtain the inequality

$$\mathsf{k}_D(x) + h_0 \geqslant_{D_0} \mathsf{f}_D(x + s_0) - \mathsf{f}_D(x) \quad \text{for all } x \in D_0.$$

Now this can only be the case if $\deg(\mathsf{k}_D(x)+h_0) \geqslant \deg(\mathsf{f}_D(x+s_0)-\mathsf{f}_D(x))$, which simplifies to $1 \geqslant \deg(\mathsf{f}_D(x)) - 1$ since $s_0 \neq 0$ and f_D is at least quadratic (hence not constant, cf. also (4.15)). Consequently, f_D must be a quadratic polynomial function, that is, $\mathsf{f}_D(x) = ax^2 + bx + c$ with $a > 0$ (due to well-definedness over D_0). Then the inequalities arising from weak compatibility with (4.52) and (4.53) simplify to

$$\begin{aligned} 2x + k_0 + h_0 &\geqslant_{D_0} 2as_0x + as_0^2 + bs_0 \\ 4as_0x + 4as_0^2 + 2bs_0 &\geqslant_{D_0} 4x + 3h_0 + k_0 \end{aligned}$$

both of which must hold for all $x \in D_0$. Hence, by looking at the leading coefficients, we infer that $as_0 = 1$. Furthermore, weak compatibility with (4.54) is satisfied if and only if the inequality

$$2as_0x + as_0^2 + bs_0 \geqslant_{D_0} \mathsf{0}_D + s_0 + h_0$$

holds for all $x \in D_0$. For $x = 0$, and using the condition $as_0 = 1$, we conclude that $bs_0 \geqslant_{D_0} \mathsf{0}_D + h_0 \geqslant_{D_0} 0$, which implies that $b \geqslant 0$ as $s_0 > 0$.

Using all the information gathered above, the compatibility constraint associated with (4.55) gives rise to the inequality $0 \geqslant_{D_0} \mathsf{f}_D(\mathsf{0}_D) + 2\,\mathsf{0}_D + bs_0 + h_0$, all of whose summands on the right-hand side are non-negative as $b \geqslant 0$ and all interpretation functions must be well-defined over D_0. Consequently, we must have $\mathsf{0}_D = h_0 = b = c = \mathsf{f}_D(\mathsf{0}_D) = 0$. □

In order to establish the main result of this subsection, we extend the TRS $\mathcal{S}$ by the rewrite rules given in Table 4.3, calling the resulting system $\mathcal{R}_5$. As in Section 4.3, each block serves a specific purpose. The one made up of (4.56) – (4.58) enforces the desired semantics for the symbol r, that is, a linear function $x \mapsto 2x$ that doubles its input, while the block (4.60) – (4.64) enforces a linear function $x \mapsto q_1x$ for the symbol q whose slope q_1 is proportional to the square of the interpretation of the constant m. Finally, (4.59) encodes the main idea of the construction, as mentioned above.

Lemma 4.32. *The TRS $\mathcal{R}_5$ is polynomially terminating over $\mathbb{N}^*$ and $\mathbb{R}^*$.*

Proof. For polynomial termination over $\mathbb{N}^*$, we start with the interpretation

$$\mathsf{0}_\mathbb{N} = 0 \quad \mathsf{s}_\mathbb{N}(x) = x + 1 \quad \mathsf{f}_\mathbb{N}(x) = x^2 \quad \mathsf{g}_\mathbb{N}(x) = 3x + 5$$
$$\mathsf{h}_\mathbb{N}(x, y) = x + y \quad \mathsf{k}_\mathbb{N}(x) = 2x + 2 \quad \mathsf{q}_\mathbb{N}(x) = 4x \quad \mathsf{r}_\mathbb{N}(x) = 2x \quad \mathsf{m}_\mathbb{N} = 2$$

$$\mathsf{k}(x) \to \mathsf{r}(x) \quad (4.56)$$
$$\mathsf{s}(\mathsf{r}(x)) \to \mathsf{h}(x, x) \quad (4.57)$$
$$\mathsf{h}(0, 0) \to \mathsf{r}(0) \quad (4.58)$$

$$\mathsf{h}(\mathsf{r}(\mathsf{q}(\mathsf{f}(x))), \mathsf{r}(x)) \to \mathsf{h}(\mathsf{r}^2(\mathsf{f}(x)), \mathsf{q}(x)) \quad (4.59)$$

$$\mathsf{g}^2(x) \to \mathsf{q}(x) \quad (4.60)$$
$$\mathsf{h}(0, 0) \to \mathsf{q}(0) \quad (4.61)$$
$$\mathsf{f}(\mathsf{f}(\mathsf{m})) \to \mathsf{q}(\mathsf{f}(\mathsf{m})) \quad (4.62)$$
$$\mathsf{h}(0, \mathsf{q}(\mathsf{f}(\mathsf{m}))) \to \mathsf{h}(\mathsf{f}(\mathsf{f}(\mathsf{m})), 0) \quad (4.63)$$
$$\mathsf{m} \to \mathsf{s}(0) \quad (4.64)$$

Table 4.3: The TRS $\mathcal{R}_5$ (without the $\mathcal{S}$-rules).

By Lemma 3.6 and Corollary 3.9, all interpretation functions are well-defined over $\mathbb{N}$ and strictly monotone (i.e., monotone with respect to $>_{\mathbb{N}}$) as well as weakly monotone (i.e., monotone with respect to $\geqslant_{\mathbb{N}}$). Moreover, it is easy to verify that this interpretation is weakly compatible with $\mathcal{R}_5$. In particular, the rule (4.59) gives rise to the constraint

$$8x^2 + 2x \geqslant_{\mathbb{N}} 4x^2 + 4x \qquad \Longleftrightarrow \qquad 2x^2 - x \geqslant_{\mathbb{N}} 0$$

which holds for all $x \in \mathbb{N}$. After removing the rules from $\mathcal{R}_5$ for which (strict) compatibility holds, we are left with the rules (4.54), (4.55), (4.58), (4.59) and (4.61) – (4.63), all of which can be handled (that is, removed at once) by the following linear interpretation:

$$0_{\mathbb{N}} = 0 \quad \mathsf{s}_{\mathbb{N}}(x) = 7x + 2 \quad \mathsf{h}_{\mathbb{N}}(x, y) = x + 2y + 1$$
$$\mathsf{f}_{\mathbb{N}}(x) = 4x + 2 \quad \mathsf{q}_{\mathbb{N}}(x) = 4x \quad \mathsf{r}_{\mathbb{N}}(x) = x \quad \mathsf{m}_{\mathbb{N}} = 0$$

For polynomial termination over $\mathbb{R}^*$, we consider the interpretation

$$\delta = 1 \quad 0_{\mathbb{R}} = 0 \quad \mathsf{s}_{\mathbb{R}}(x) = x + 1 \quad \mathsf{f}_{\mathbb{R}}(x) = x^2 \quad \mathsf{g}_{\mathbb{R}}(x) = 3x + 5$$
$$\mathsf{h}_{\mathbb{R}}(x, y) = x + y \quad \mathsf{k}_{\mathbb{R}}(x) = 2x + 2 \quad \mathsf{q}_{\mathbb{R}}(x) = 2x \quad \mathsf{r}_{\mathbb{R}}(x) = 2x \quad \mathsf{m}_{\mathbb{R}} = \sqrt{2}$$

which is both weakly and strictly monotone according to Lemmata 3.34 – 3.36, and 3.38. So all interpretation functions are well-defined over $\mathbb{R}_0$ and monotone with respect to $>_{\mathbb{R}_0,\delta}$ and $\geqslant_{\mathbb{R}_0}$. Moreover, one easily verifies that this interpretation is weakly compatible with $\mathcal{R}_5$. In particular, the constraint $4x^2 + 2x \geqslant_{\mathbb{R}_0} 4x^2 + 2x$ associated with (4.59) trivially holds. After removing the rules from $\mathcal{R}_5$ for which (strict) compatibility holds (i.e., for which the corresponding constraint remains true after strengthening $\geqslant_{\mathbb{R}_0}$ to $>_{\mathbb{R}_0,\delta}$), we are left with (4.54), (4.55), (4.58), (4.59) and (4.61) – (4.64), all of which can be removed at once by the following linear interpretation:

$$\delta = 1 \quad 0_{\mathbb{R}} = 0 \quad \mathsf{s}_{\mathbb{R}}(x) = 6x + 2 \quad \mathsf{f}_{\mathbb{R}}(x) = 3x + 2$$
$$\mathsf{h}_{\mathbb{R}}(x, y) = x + 2y + 1 \quad \mathsf{q}_{\mathbb{R}}(x) = 2x \quad \mathsf{r}_{\mathbb{R}}(x) = x \quad \mathsf{m}_{\mathbb{R}} = 3$$

□

Lemma 4.33. *The TRS $\mathcal{R}_5$ is not polynomially terminating over $\mathbb{Q}^*$.*

Proof. Let $\mathcal{P}$ be a strictly monotone polynomial interpretation over $\mathbb{Q}$ that is weakly compatible with $\mathcal{R}_5$. According to Lemma 4.31, the symbols $\mathsf{0}$, s, f, g, h and k are interpreted as follows:

$$\mathsf{0}_{\mathbb{Q}} = 0 \quad \mathsf{s}_{\mathbb{Q}}(x) = x + s_0 \quad \mathsf{h}_{\mathbb{Q}}(x,y) = x + y$$
$$\mathsf{g}_{\mathbb{Q}}(x) = g_1 x + g_0 \quad \mathsf{k}_{\mathbb{Q}}(x) = 2x + k_0 \quad \mathsf{f}_{\mathbb{Q}}(x) = ax^2$$

where $s_0, g_1, a > 0$ and $g_0, k_0 \geqslant 0$.

As the interpretation of k is linear, weak compatibility with the rule (4.56) implies that the interpretation of r is at most linear as well, i.e., $\mathsf{r}_{\mathbb{Q}}(x) = r_1 x + r_0$ with $r_0 \geqslant 0$ and $2 \geqslant r_1 \geqslant 0$. We also have $r_1 \geqslant 2$ due to weak compatibility with (4.57) and $0 \geqslant r_0$ due to weak compatibility with (4.58); hence, $\mathsf{r}_{\mathbb{Q}}(x) = 2x$.

Similarly, by linearity of $\mathsf{g}_{\mathbb{Q}}$ and weak compatibility with (4.60), the interpretation of q must have the shape $\mathsf{q}_{\mathbb{Q}}(x) = q_1 x + q_0$. Then weak compatibility with (4.61) yields $0 \geqslant q_0$; hence, $\mathsf{q}_{\mathbb{Q}}(x) = q_1 x$, $q_1 \geqslant 0$. Next we note that weak compatibility with (4.62) and (4.63) implies that $\mathsf{f}_{\mathbb{Q}}(\mathsf{f}_{\mathbb{Q}}(\mathsf{m}_{\mathbb{Q}})) = \mathsf{q}_{\mathbb{Q}}(\mathsf{f}_{\mathbb{Q}}(\mathsf{m}_{\mathbb{Q}}))$, which evaluates to $a^3 \mathsf{m}_{\mathbb{Q}}^4 = a\, q_1 \mathsf{m}_{\mathbb{Q}}^2$. From this we infer that $q_1 = a^2 \mathsf{m}_{\mathbb{Q}}^2$ as $a > 0$ and $\mathsf{m}_{\mathbb{Q}} \geqslant s_0 > 0$ due to weak compatibility with (4.64); i.e., $\mathsf{q}_{\mathbb{Q}}(x) = a^2 \mathsf{m}_{\mathbb{Q}}^2 x$.

Finally, we consider the weak compatibility constraint associated with (4.59), which simplifies to

$$(2ax^2 - x)((a\,\mathsf{m}_{\mathbb{Q}})^2 - 2) \geqslant 0 \quad \text{for all } x \in \mathbb{Q}_0.$$

However, this inequality is unsatisfiable as the polynomial $2ax^2 - x$ is negative for some $x \in \mathbb{Q}_0$ and $(a\,\mathsf{m}_{\mathbb{Q}})^2 - 2$ cannot be zero because both a and $\mathsf{m}_{\mathbb{Q}}$ must be rational numbers. So we conclude that there is no strictly monotone polynomial interpretation over $\mathbb{Q}$ that is weakly compatible with the TRS $\mathcal{R}_5$. This implies that $\mathcal{R}_5$ is not incrementally polynomially terminating over $\mathbb{Q}$. □

Together, Lemma 4.32 and Lemma 4.33 yield the main result of this subsection.

Corollary 4.34. *There are TRSs that are polynomially terminating over $\mathbb{N}^*$ and $\mathbb{R}^*$ but not over $\mathbb{Q}^*$.* □

4.4.2 Incremental Polynomial Termination over $\mathbb{N}$ vs. $\mathbb{R}$

In this subsection, we show that there are TRSs that are incrementally polynomially terminating over $\mathbb{N}$ but not over $\mathbb{R}$. For this purpose, we extend the TRS $\mathcal{S}$ of the previous subsection by the single rewrite rule $\mathsf{f}(x) \to x$ and call the resulting system $\mathcal{R}_6$.

Lemma 4.35. *The TRS $\mathcal{R}_6$ is polynomially terminating over $\mathbb{N}^*$.*

Proof. First, we consider the interpretation

$$\mathsf{0}_{\mathbb{N}} = 0 \qquad \mathsf{s}_{\mathbb{N}}(x) = x + 1 \qquad \mathsf{f}_{\mathbb{N}}(x) = x^2$$
$$\mathsf{h}_{\mathbb{N}}(x,y) = x + y \qquad \mathsf{g}_{\mathbb{N}}(x) = 3x + 5 \qquad \mathsf{k}_{\mathbb{N}}(x) = 2x + 2$$

which is both weakly and strictly monotone according to Lemma 3.6 and Corollary 3.9, and one easily verifies that this interpretation is weakly compatible with $\mathcal{R}_6$. In particular, the constraint $x^2 \geqslant_{\mathbb{N}} x$ associated with $\mathsf{f}(x) \to x$ holds for all $x \in \mathbb{N}$. Removing the rules from $\mathcal{R}_6$ for which (strict) compatibility holds leaves us with the rules (4.54), (4.55) and $\mathsf{f}(x) \to x$, which are easily handled, e.g. by the linear interpretation

$$\mathsf{0}_{\mathbb{N}} = 0 \qquad \mathsf{s}_{\mathbb{N}}(x) = 3x + 2 \qquad \mathsf{f}_{\mathbb{N}}(x) = 2x + 1 \qquad \mathsf{h}_{\mathbb{N}}(x, y) = x + y \qquad \square$$

Lemma 4.36. *The TRS $\mathcal{R}_6$ is not polynomially terminating over $\mathbb{R}^*$ or $\mathbb{Q}^*$.*

Proof. Let $D \in \{\mathbb{Q}, \mathbb{R}\}$, and let $\mathcal{P}$ be a polynomial interpretation over D that is weakly compatible with $\mathcal{R}_6$, and in which the interpretation of the function symbol f has the shape $\mathsf{f}_D(x) = ax^2$ with $a > 0$. Then the weak compatibility constraint $ax^2 \geqslant_{D_0} x$ associated with $\mathsf{f}(x) \to x$ does not hold for all $x \in D_0$ because the polynomial $ax^2 - x = ax\left(x - \frac{1}{a}\right)$ is negative in the open interval $\left(0, \frac{1}{a}\right)$. As the above assumption on the interpretation of f follows from Lemma 4.31 if $\mathcal{P}$ is strictly monotone, we conclude that there is no strictly monotone polynomial interpretation over $\mathbb{R}$ or $\mathbb{Q}$ that is weakly compatible with the TRS $\mathcal{R}_6$. This implies that $\mathcal{R}_6$ is not incrementally polynomially terminating over $\mathbb{R}$ or $\mathbb{Q}$. $\square$

Together, Lemma 4.35 and Lemma 4.36 yield the main result of this subsection.

Corollary 4.37. *There are TRSs that are polynomially terminating over $\mathbb{N}^*$ but not over $\mathbb{R}^*$ or $\mathbb{Q}^*$.* $\square$

The results presented in this section are summarized in Figure 4.5, which gives the full picture of the relationship between incremental polynomial termination over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$.

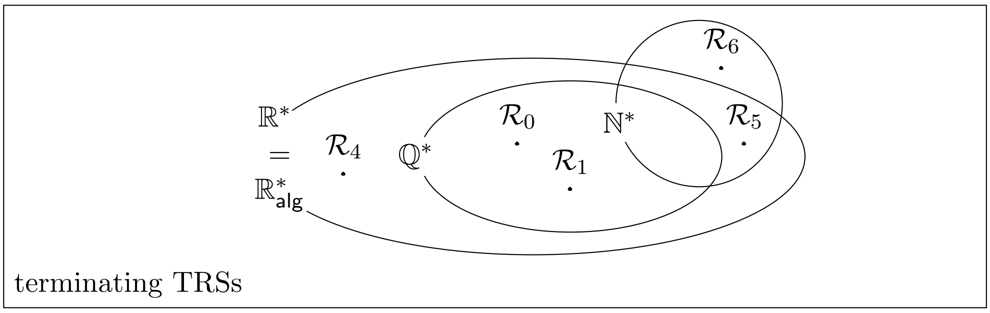

Figure 4.5: Summary.

4.5 Polynomial Interpretations in the DP Framework

In this section, we review the results obtained in the previous section in light of the DP framework. Technically, the difference between the use of polynomial interpretations in incremental termination proofs and their use in the DP framework is that the former requires strict monotonicity of the interpretations, whereas the latter only demands weak monotonicity. Now the goal of this section is to adapt the results obtained in Section 4.4 by modifying the corresponding TRSs in such a way as to eliminate any dependence on strict monotonicity in the respective proof arguments. Fortunately, there are not many occurrences where strict monotonicity is crucial, and all of them can easily be handled. The idea is as follows. Suppose that a unary function symbol f is interpreted by a linear polynomial $f_1 x + f_0$ and consider the rewrite rule $\mathsf{f}(x) \to x$. Then weak compatibility translates to the inequality $f_1 x + f_0 \geqslant x$, which implies $f_1 \geqslant 1$. So the rule $\mathsf{f}(x) \to x$ enforces strict monotonicity of the interpretation of the symbol f. Moreover, note that we can also use the rule $\mathsf{s}(\mathsf{f}(x)) \to x$ for this purpose if the interpretation of the symbol s is fixed to the shape $x \mapsto x + s_0$ (which is the case for all TRSs considered in this section).

Before we use this idea to strengthen the results obtained in the previous section, let us recall that, by Lemma 4.10, polynomial interpretations over the rational numbers are subsumed by polynomial interpretations over the real numbers. In particular, in the setting of the DP framework, we obtain the following corollary from Lemma 4.10.

Corollary 4.38. *If there exists a weakly monotone polynomial interpretation over $\mathbb{Q}$ that succeeds on a given DP problem, then there also exists a weakly monotone polynomial interpretation over $\mathbb{R}$ that succeeds on this DP problem.* □

In order to strengthen the results related to the TRSs $\mathcal{R}_4$, $\mathcal{R}_5$ and $\mathcal{R}_6$, we now apply the construction described above and extend these systems by the following rewrite rules, referring to the resulting systems as $\mathcal{R}_4'$, $\mathcal{R}_5'$ and $\mathcal{R}_6'$:

$$\mathsf{s}(x) \to x \tag{4.65}$$
$$\mathsf{g}(x) \to x \tag{4.66}$$
$$\mathsf{f}(\mathsf{g}(x)) \to x \tag{4.67}$$
$$\mathsf{s}(\mathsf{h}(x, \mathsf{0})) \to x \tag{4.68}$$
$$\mathsf{s}(\mathsf{h}(\mathsf{0}, x)) \to x \tag{4.69}$$

To begin with, we show that this allows us to drop the requirement of strict monotonicity in Lemma 4.28.

Lemma 4.39. *Let $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$, and let $\mathcal{P}$ be a weakly monotone polynomial interpretation over D that is weakly compatible with the rules (4.9) – (4.13) and (4.65) – (4.69). Then*

$$\mathsf{s}_D(x) = x + s_0 \qquad \mathsf{h}_D(x, y) = x + y + h_0 \qquad \mathsf{g}_D(x) = g_1 x + g_0$$

where all coefficients are non-negative and $g_1 \geqslant 2$. Moreover, $\mathsf{f}_D(x)$ is at least quadratic.

Proof. Due to weak compatibility with (4.65), (4.66) and (4.67), the interpretations $\mathsf{f}_D(x)$, $\mathsf{g}_D(x)$ and $\mathsf{s}_D(x)$ of the unary symbols f, g and s cannot be constant polynomials. But then the proof of Lemma 4.28 shows that the symbols s, g and h must be interpreted by linear polynomials $\mathsf{s}_D(x) = s_1 x + s_0$,

$\mathsf{g}_D(x) = g_1x + g_0$ and $\mathsf{h}_D(x,y) = h_1x + h_2y + h_0$, subject to the following constraints:

$$1 \geqslant s_1 > 0 \quad s_1 \geqslant h_1 \geqslant 0 \quad s_1 \geqslant h_2 \geqslant 0 \quad g_1 \geqslant h_1 + h_2 \quad g_1 > 0 \quad s_0, g_0, h_0 \geqslant 0$$

Moreover, f_D cannot be linear if $g_1 > 1$.

In this situation, weak compatibility with the rule (4.65) implies that $s_1 \geqslant 1$ since $\mathsf{s}_D(x) = s_1x + s_0$; hence, $\mathsf{s}_D(x) = x + s_0$. Using the latter interpretation in the compatibility constraint associated with (4.68), we infer that $h_1 \geqslant 1$. Likewise, by weak compatibility with (4.69), we obtain $h_2 \geqslant 1$. But then we must have $h_1 = h_2 = 1$ since $s_1 = 1$, and therefore $g_1 \geqslant h_1 + h_2 = 2$, such that $\mathsf{s}_D(x) = x + s_0$, $\mathsf{h}_D(x,y) = x + y + h_0$ and $\mathsf{g}_D(x) = g_1x + g_0$ with $g_1 \geqslant 2$, which shows that $\mathsf{f}_D(x)$ is at least quadratic. □

As a consequence of this lemma, we can also drop the requirement of strict monotonicity in Lemma 4.31 by adding the rules (4.65) – (4.69) to the TRS $\mathcal{S}$ of Section 4.4, thereby obtaining the TRS $\mathcal{S}'$.

Lemma 4.40. *Let $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$, and let $\mathcal{P}$ be a weakly monotone polynomial interpretation over D that is weakly compatible with the TRS $\mathcal{S}'$. Then*

$$\mathsf{0}_D = 0 \quad \mathsf{s}_D(x) = x + s_0 \quad \mathsf{h}_D(x,y) = x + y$$
$$\mathsf{g}_D(x) = g_1x + g_0 \quad \mathsf{k}_D(x) = 2x + k_0 \quad \mathsf{f}_D(x) = ax^2$$

where $as_0 = 1$, $g_1 \geqslant 2$ and all coefficients are non-negative.

Proof. By the proof of Lemma 4.31, replacing Lemma 4.28 by Lemma 4.39. □

This is all we need to get rid of the requirement of strict monotonicity in the results obtained for the TRSs $\mathcal{R}_4$, $\mathcal{R}_5$ and $\mathcal{R}_6$ in Lemmata 4.29, 4.33 and 4.36.

Lemma 4.41. *For the TRSs $\mathcal{R}_4'$, $\mathcal{R}_5'$ and $\mathcal{R}_6'$ the following statements hold:*

1. *There is no weakly monotone polynomial interpretation over $\mathbb{Q}$ or $\mathbb{N}$ that is weakly compatible with $\mathcal{R}_4'$.*
2. *There is no weakly monotone polynomial interpretation over $\mathbb{Q}$ that is weakly compatible with $\mathcal{R}_5'$.*
3. *There is no weakly monotone polynomial interpretation over $\mathbb{R}$ or $\mathbb{Q}$ that is weakly compatible with $\mathcal{R}_6'$.*

Proof. The first claim follows from the proof of Lemma 4.29 after replacing Lemma 4.28 by Lemma 4.39 (and relaxing strict monotonicity to weak monotonicity). Similarly, the other claims follow from the proofs of Lemmata 4.33 and 4.36, respectively, after replacing Lemma 4.31 by Lemma 4.40. □

Hence, we obtain the following corollaries in the DP framework.

Corollary 4.42. *For the TRSs $\mathcal{R}_4'$, $\mathcal{R}_5'$ and $\mathcal{R}_6'$ the following statements hold:*

1. *There is no weakly monotone polynomial interpretation over $\mathbb{Q}$ or $\mathbb{N}$ that succeeds on the DP problem $(\_, \mathcal{R}'_4)$.*

2. *There is no weakly monotone polynomial interpretation over $\mathbb{Q}$ that succeeds on the DP problem $(\_, \mathcal{R}'_5)$.*

3. *There is no weakly monotone polynomial interpretation over $\mathbb{R}$ or $\mathbb{Q}$ that succeeds on the DP problem $(\_, \mathcal{R}'_6)$.* □

Nevertheless, these TRSs do admit certain polynomial interpretations, even strictly monotone ones, which can be leveraged to establish (incremental) polynomial termination of all systems.

Lemma 4.43. *For the TRSs $\mathcal{R}'_4$, $\mathcal{R}'_5$ and $\mathcal{R}'_6$ the following statements hold:*

1. *$\mathcal{R}'_4$ is polynomially terminating over $\mathbb{R}$ (hence also over $\mathbb{R}^*$).*

2. *$\mathcal{R}'_5$ is polynomially terminating over $\mathbb{N}^*$ and $\mathbb{R}^*$.*

3. *$\mathcal{R}'_6$ is polynomially terminating over $\mathbb{N}^*$.*

Proof. For the first claim, the same interpretation as in the proof of Lemma 4.23 applies. The second claim follows by the same interpretations as in the proof of Lemma 4.32, whereas the third one follows by the interpretations given in the proof of Lemma 4.35. □

As far as the TRSs $\mathcal{R}_0$ and $\mathcal{R}_1$ are concerned, adding the rules (4.65) – (4.69) does not have the desired effect of yielding TRSs that preclude weakly monotone polynomial interpretations over $\mathbb{N}$ while admitting interpretations, preferably even strictly monotone ones, over $\mathbb{Q}$ (and $\mathbb{R}$). However, such TRSs do exist. For example, the TRS $\mathcal{R}_7$ consisting of (4.9) – (4.13), (4.65) – (4.69), (4.70) and (4.71):

$$\mathsf{h}(\mathsf{k}(x), \mathsf{k}(x)) \to \mathsf{h}(\mathsf{h}(x, x), x) \tag{4.70}$$
$$\mathsf{s}(\mathsf{h}(\mathsf{h}(x, x), x)) \to \mathsf{h}(\mathsf{k}(x), \mathsf{k}(x)) \tag{4.71}$$

Lemma 4.44. *There is no weakly monotone polynomial interpretation over $\mathbb{N}$ that is weakly compatible with the TRS $\mathcal{R}_7$.*

Proof. Let $\mathcal{P}$ be a weakly monotone polynomial interpretation over $\mathbb{N}$ that is weakly compatible with $\mathcal{R}_7$. Then, by Lemma 4.39, we have $\mathsf{s}_{\mathbb{N}}(x) = x + s_0$ and $\mathsf{h}_{\mathbb{N}}(x, y) = x + y + h_0$, such that the weak compatibility constraints associated with (4.70) and (4.71) give rise to the inequalities

$$3x + 2h_0 + s_0 \geqslant_{\mathbb{N}} 2\,\mathsf{k}_{\mathbb{N}}(x) + h_0 \geqslant_{\mathbb{N}} 3x + 2h_0$$

which must hold for all $x \in \mathbb{N}$. Hence, the interpretation of k must have the shape $\mathsf{k}_{\mathbb{N}}(x) = \frac{3}{2}x + k_0$, which is not an integer polynomial. □

Corollary 4.45. *There is no weakly monotone polynomial interpretation over $\mathbb{N}$ that succeeds on the DP problem $(\_, \mathcal{R}_7)$.* □

Nevertheless, there are strictly monotone polynomial interpretations over $\mathbb{Q}$ (and $\mathbb{R}$) that are compatible with $\mathcal{R}_7$.

Lemma 4.46. *The TRS $\mathcal{R}_7$ is polynomially terminating over $\mathbb{Q}$.*

Proof. By the following interpretation:

$$\delta = 1 \qquad \mathsf{0}_{\mathbb{Q}} = 0 \qquad \mathsf{s}_{\mathbb{Q}}(x) = x + 4 \qquad \mathsf{f}_{\mathbb{Q}}(x) = x^2$$
$$\mathsf{g}_{\mathbb{Q}}(x) = 3x + 5 \qquad \mathsf{h}_{\mathbb{Q}}(x, y) = x + y \qquad \mathsf{k}_{\mathbb{Q}}(x) = \tfrac{3}{2}x + 1$$

□

4.6 Conclusion

In this chapter, we investigated the relationship of polynomial interpretations with real, real algebraic, rational and integer coefficients with regard to termination proving power.

Concerning the notion of direct polynomial termination, we presented three new results, the first of which shows that polynomial interpretations with real coefficients subsume polynomial interpretations with rational coefficients, the second of which shows that polynomial interpretations with real or rational coefficients do not properly subsume polynomial interpretations with integer coefficients, a result that comes somewhat unexpected, and the third of which shows that there are TRSs that can be proved terminating by polynomial interpretations with real and with integer coefficients but not with rational coefficients. Besides, we also showed that polynomial interpretations with real coefficients are equivalent to polynomial interpretations with real algebraic coefficients, a result confirming that automatic termination tools do not have to be concerned with transcendental (real) numbers. Together with the earlier results of Lucas [46], the results presented in this chapter give the full picture of the relationship between the various instances of polynomial interpretations.

We also considered the possibility of establishing termination by using polynomial interpretations in an incremental fashion, showing that the relationship between the induced notions of incremental polynomial termination (over $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$ and $\mathbb{R}$) equals the one obtained for direct polynomial termination. Furthermore, we indicated how to adapt these results to the DP framework, thereby obtaining evidence that the relationship is no different in that setting (for standard reduction pairs based on polynomial interpretations, as described in Section 1.4). However, in this respect more work needs to be done. Most notably, it remains to consider reduction pairs that incorporate usable rules with (implicit) *argument filters* [26] (induced by polynomial interpretations).

Part II

Matrix Interpretations

Chapter 5

Introduction and Outline

The second part of this thesis is devoted to matrix interpretations. Since their inception in 2006, matrix interpretations have evolved into one of the most important (that is, powerful) methods for termination analysis and complexity analysis of term rewrite systems. While originally introduced by Hofbauer and Waldmann as a stand-alone method for establishing termination proofs in the context of string rewriting [32, 33], allowing them to solve challenging termination problems like $\{\mathsf{aa} \to \mathsf{bc}, \mathsf{bb} \to \mathsf{ac}, \mathsf{cc} \to \mathsf{ab}\}$, problem #104 on the RTA list of open problems,[1] it was not long until Endrullis *et al.* [20] generalized (one particular instance of) the matrix method to term rewriting and also incorporated it into the DP framework [3, 25–28, 71], the state-of-the-art framework for establishing termination of TRSs.

The matrix method is based on the well-known paradigm of interpreting terms into a domain equipped with a suitable well-founded (order) relation. In the original approach of [20], the authors consider the set of vectors of natural numbers as underlying domain, together with a well-founded order that relates two vectors if and only if there is a strict decrease in the respective first components and a weak decrease in all other components. Function symbols are interpreted by suitable linear mappings represented by square matrices of natural numbers.

The order chosen in [20] is not the only possible base order for matrix interpretations. There do exist other well-founded orders on vectors of natural numbers that induce variants of matrix interpretations which are also suitable for termination analysis. We study several such orders in Appendix B, some of which give rise to matrix interpretations that are equally (but not more) powerful for proving termination than the ones of [20]. This research originally appeared in [55]. (We present it in the appendix because it does not yield more powerful kinds of matrix interpretations. Nevertheless, it does give further insight as to why the approach of [20] performs well in practice, and it also contains a generalization of the latter coming along with a more powerful implementation.)

Recently, another generalization appeared in [17] that employs matrices of natural numbers as underlying domain and interprets each function symbol by a linear matrix polynomial. In principle, this approach also allows for non-linear matrix polynomials.

In [1, 22, 76] the method of Endrullis *et al.* was lifted to the non-negative rational and real (algebraic) numbers using the same technique that was already used to lift polynomial interpretations from the natural numbers to the rationals

[1] `http://rtaloop.mancoosi.univ-paris-diderot.fr`

and reals (cf. Section 2.1 and [30]). Thus, one distinguishes three variants of matrix interpretations, matrix interpretations over the real, rational and natural numbers, and the obvious question is:

What is their relationship with regard to termination proving power?

Giving a complete answer to this question is the first major goal of this second part of the thesis. This is achieved in Chapter 6, where we also show how the choice of the matrix dimension affects termination proving power.

Beyond termination analysis, matrix interpretations are also apt for analyzing the *derivational complexity* of TRSs, where the aim is to obtain (quantitative) information about the maximal length of rewrite sequences (or derivations) in terms of the size of their initial term. As term rewriting is a formal model of computation and algorithms of polynomial complexity are widely accepted as *feasible*, one is especially interested in polynomial derivational complexity. However, in general, matrix interpretations induce exponential upper bounds on the derivational complexity of compatible TRSs. In order to obtain polynomial upper bounds, the matrices occurring in a matrix interpretation have to satisfy certain (additional) properties, the study of which is the main objective of Chapter 7.

In the remainder of this chapter, we give a formal account of matrix interpretations, introducing all relevant concepts, definitions and terminology. But first we recall a few basic notions from linear algebra.

5.1 Linear Algebra

For any commutative ring R (e.g., $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}_{\mathsf{alg}}$, $\mathbb{R}$), we denote the ring of all *n-dimensional square matrices* over R by $R^{n \times n}$. The $n \times n$ *identity matrix* is denoted by I_n and the $n \times n$ *zero matrix* by 0_n. We simply write I and 0 if n is clear from the context. In case R is equipped with a partial order $\geqslant$, the componentwise extension of this order to $R^{n \times n}$ is also denoted by $\geqslant$. We call a matrix *non-negative* if all its entries are non-negative and denote the set of all non-negative n-dimensional square matrices of $\mathbb{Z}^{n \times n}$ ($\mathbb{R}^{n \times n}$) by $\mathbb{N}^{n \times n}$ ($\mathbb{R}_0^{n \times n}$). As usual, we write A^{T} for the *transpose* of a matrix (vector) A. For a (column) vector $\vec{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$, $(\vec{x})_i$ denotes its i-th component x_i. Likewise, A_{ij} denotes the entry in the i-th row and j-th column of a matrix A, and A_{j-} (A_{-j}) refers to the j-th row (column).

For a square matrix $A \in R^{n \times n}$, the *characteristic polynomial* $\chi_A(\lambda)$ is defined as $\det(\lambda I_n - A)$, where det denotes the (matrix) determinant. It is a monic polynomial of degree n with coefficients in R. The equation $\chi_A(\lambda) = 0$ is called the *characteristic equation* of A. The solutions of this equation, that is, the *roots* of $\chi_A(\lambda)$, are precisely the *eigenvalues* of A. If R is contained in an algebraically closed field (where each polynomial of degree n with coefficients in the field is guaranteed to have exactly n roots), then A has exactly n (not necessarily distinct) eigenvalues in this field. A non-zero vector x is an *eigenvector* of A if $Ax = \lambda x$ for some eigenvalue λ of A.

We say that a polynomial $p \in R[x]$ *annihilates* a square matrix $A \in R^{n \times n}$ if $p(A) = 0$. The Cayley-Hamilton theorem [66] states that A satisfies its own characteristic equation, that is, χ_A annihilates A (it holds for square matrices over commutative rings). Let R be a field and consider the set $\{ p \in R[x] \mid p(A) = 0 \}$ of annihilating polynomials of a matrix $A \in R^{n \times n}$. This set is generated by the *minimal polynomial* $\mathrm{m}_A(x)$ of A, which is the unique monic polynomial of minimum degree that annihilates A (cf. e.g. [35]). Any polynomial that annihilates A is a (polynomial) multiple of $\mathrm{m}_A(x)$. In other words, if $p(A) = 0$ for some $p \in R[x]$, then $\mathrm{m}_A(x)$ divides $p(x)$. In particular, $\mathrm{m}_A(x)$ divides the characteristic polynomial of A, and $\mathrm{m}_A(\lambda) = 0$ if and only if λ is an eigenvalue of A.

5.2 Matrix Interpretations

Following [1, 20, 76], we define matrix interpretations as follows. Let $\mathcal{F}$ denote a signature. For *matrix interpretations over* $\mathbb{R}$, we fix a *dimension* $n \in \mathbb{N} \setminus \{0\}$, some positive real number δ and use the set $\mathbb{R}_0^n$ as the carrier of an $\mathcal{F}$-algebra $\mathcal{M}$, together with the orders $>_\delta$ and $\geqslant$ on $\mathbb{R}_0^n$:

$$(x_1, \dots, x_n)^{\mathrm{T}} >_\delta (y_1, \dots, y_n)^{\mathrm{T}} \iff x_1 >_{\mathbb{R},\delta} y_1 \text{ and } x_i \geqslant_{\mathbb{R}} y_i \text{ for } i = 2, \dots, n$$
$$(x_1, \dots, x_n)^{\mathrm{T}} \geqslant (y_1, \dots, y_n)^{\mathrm{T}} \iff x_i \geqslant_{\mathbb{R}} y_i \text{ for } i = 1, \dots, n$$

Here, as in Section 2.1, $x >_{\mathbb{R},\delta} y$ if and only if $x \geqslant_{\mathbb{R}} y + \delta$. Each k-ary function symbol $f \in \mathcal{F}$ is interpreted by a linear function of the shape

$$f_{\mathcal{M}} \colon (\mathbb{R}_0^n)^k \to \mathbb{R}_0^n, (\vec{x}_1, \dots, \vec{x}_k) \mapsto F_1 \vec{x}_1 + \cdots + F_k \vec{x}_k + \vec{f}$$

where $\vec{x}_1, \dots, \vec{x}_k$ are (column) vectors of variables, $F_1, \dots, F_k \in \mathbb{R}_0^{n \times n}$ and $\vec{f} \in \mathbb{R}_0^n$. In this way, $(\mathcal{M}, >_\delta, \geqslant)$ forms a weakly monotone $\mathcal{F}$-algebra. If, in addition, the top left entry $(F_i)_{11}$ of each matrix F_i is at least one, then we call $\mathcal{M}$ a *monotone* matrix interpretation over $\mathbb{R}$, in which case $(\mathcal{M}, >_\delta, \geqslant)$ becomes an extended monotone $\mathcal{F}$-algebra. Note that in any case we have $>_{\mathcal{M}} \subseteq \geqslant_{\mathcal{M}}$ since $>_\delta \subseteq \geqslant$ (independently of δ). That is, compatibility of $\mathcal{M}$ with a set of rewrite rules implies weak compatibility with that set of rules, as was the case for polynomial interpretations.

We obtain matrix interpretations over $\mathbb{R}_{\mathsf{alg}}$ by restricting the carrier to the set of vectors of non-negative real algebraic numbers. Similarly, matrix interpretations over $\mathbb{Q}$ operate on the carrier $\mathbb{Q}_0^n$. For matrix interpretations over $\mathbb{N}$, one uses the carrier $\mathbb{N}^n$ and $\delta = 1$, such that

$$(x_1, \dots, x_n)^{\mathrm{T}} >_\delta (y_1, \dots, y_n)^{\mathrm{T}} \iff x_1 >_{\mathbb{N}} y_1 \text{ and } x_i \geqslant_{\mathbb{N}} y_i \text{ for } i = 2, \dots, n$$

For polynomial interpretations, we showed in Section 4.2 that it suffices to consider the real algebraic numbers instead of the entire set of real numbers, thus establishing the equivalence between polynomial interpretations over $\mathbb{R}$ and polynomial interpretations over $\mathbb{R}_{\mathsf{alg}}$ for proving termination of TRSs. Observing that the technique employed in Section 4.2 readily applies to matrix interpretations as well, we conclude that matrix interpretations over $\mathbb{R}$ are equivalent to matrix interpretations over $\mathbb{R}_{\mathsf{alg}}$ (cf. also [59]). In the context of string rewriting, this was also observed in [22]. So transcendental numbers are not relevant for termination proofs based on matrix interpretations. Nevertheless, for the sake of brevity of notation, we shall stick to the term “matrix interpretations over the real numbers”.

In analogy to polynomial interpretations, matrix interpretations can be used as a stand-alone termination method or in the context of the DP framework. Specializing the definitions and results given for monotone algebras in Section 1.3, we note that in the former case the termination of a TRS can either be shown directly by a compatible monotone matrix interpretation (cf. Theorem 1.6 and Corollary 1.7) or incrementally by a sequence of monotone matrix interpretations, each of which removes some rewrite rules until eventually all rewrite rules have been removed. In the latter case, when applying matrix interpretations in the DP framework, the algebras induced by them are only required to be weakly monotone. In this context, we say that a matrix interpretation $\mathcal{M}$ *succeeds* on a given DP problem if the weakly monotone algebra $(\mathcal{M}, >_\delta, \geqslant)$ succeeds on it.

Chapter 6

Matrix Termination Hierarchy

In this chapter, we clarify the relationship between matrix interpretations over the real, rational and natural numbers. We also clarify the ramifications of matrix dimension on termination proving power.

As a starting point, it is instructive to restrict to one-dimensional matrix interpretations, that is, *linear* polynomial interpretations, for which the relationship between the induced notions of polynomial termination is known (cf. Chapter 4 and [46, 54]) and can be pictured as in Figure 6.1. That is, linear

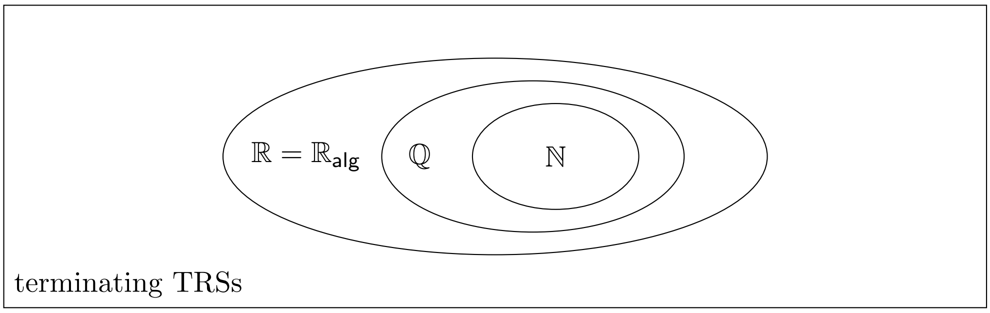

Figure 6.1: Linear polynomial interpretations.

polynomial interpretations over the real numbers subsume linear polynomial interpretations over the rational numbers, which in turn subsume linear polynomial interpretations over the natural numbers. Both inclusions are proper. In [46] this is evidenced by the TRSs $\mathcal{R}_{\mathbb{Q}}$ and $\mathcal{R}_{\mathbb{R}}$, the first of which can be shown terminating by a linear polynomial interpretation over the rational numbers but not over the natural numbers. Similarly, the second system can be shown terminating by a linear polynomial interpretation over the reals but not over the rationals. Unfortunately, the usefulness of both $\mathcal{R}_{\mathbb{Q}}$ and $\mathcal{R}_{\mathbb{R}}$ is limited to dimension one (cf. [48]) because, without restricting the dimension, both systems can be handled with 2-dimensional matrix interpretations over the natural numbers, cf. Appendix A. (The same is true for the TRSs $\mathcal{R}_0$ and $\mathcal{R}_1$ occurring in Figure 4.4 of Chapter 4, as can easily be verified with any termination prover supporting matrix interpretations, e.g., $\mathsf{T_TT_2}$ [41]. For the TRS $\mathcal{R}_4$, we could not even find a compatible matrix interpretation.) In this context, we also mention related work appearing in [23], where a relative termination problem in the form of a string rewrite system is presented that can be handled with matrix interpretations over the rationals but not with matrix interpretations over the natural numbers. However, *relative* termination is essential in this example because the relative component is the key ingredient for precluding matrix interpretations over the natural numbers. As the latter component consists of a single non-terminating rule, the entire example does not readily generalize to (real) termination problems. Besides, there is no evidence in [23] demonstrating the benefit of using irrational numbers in matrix interpretations. Thus, we conclude that new techniques are required to clarify the relationship between matrix interpretations over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$.

One of the main results of this chapter is to show that the termination hierarchy depicted in Figure 6.1 does in fact extend from one-dimensional matrix interpretations to arbitrary matrix interpretations. That is, matrix interpretations over the reals are more powerful with respect to proving termination than matrix interpretations over the rationals, which are in turn more powerful than matrix interpretations over the natural numbers. In particular, we show that this relationship does not only hold in connection with matrix interpretations as a stand-alone termination technique but also in the setting of the DP framework. Moreover, our results point out the limitations of a recent attempt [48] to simulate matrix interpretations over the rationals with matrix interpretations over the natural numbers (of higher dimension).

We also investigate the ramifications of matrix dimension on termination proving power. Clearly, by increasing the dimension, one can never lose power (in theory; in practice the increased search space may prohibit finding a termination proof). But what is the exact shape of the inherent dimension hierarchy? A partial answer to this question was given in [23], where the authors show that the hierarchy is infinite. Yet no exact information is provided as to which levels are actually inhabited. We close this gap in the second part of this chapter, thus giving a complete answer to the question raised above. For this purpose, we establish a hierarchy of matrix interpretations with respect to matrix dimension and show it to be infinite, with each level properly subsuming its predecessor. In other words, we show that matrix interpretations of dimension $(n+1)$ are strictly more powerful for proving termination than n-dimensional matrix interpretations (for any $n \geqslant 1$). The construction we use for this purpose is entirely different from the one proposed in [23]. Apart from the fact that it allows to infer the exact shape of the dimension hierarchy, it has the additional advantage that it produces witnesses (that is, TRSs) that are substantially smaller than the ones of [23]. To be precise, the construction employed in [23] gives rise to a family of SRSs $(\mathcal{S}_d)_{d \geqslant 2}$ having the property that any of its members $\mathcal{S}_{2d}$ (of even index) cannot be handled with matrix interpretations of dimension d or less (as a consequence of the Amitsur-Levitzki theorem [2]), but can be handled with dimension $d' = 2d + 3$. Each system $\mathcal{S}_d$ consists of the following rules over the finite alphabet $\Sigma_d = \{\mathsf{s}, \mathsf{1}, \ldots, \mathsf{d}, \mathsf{f}\}$:

$$\mathsf{s}\, e_k\, \mathsf{f} \to \mathsf{s}\, o_k\, \mathsf{f}$$

for all $1 \leqslant k \leqslant \frac{d!}{2}$. Here, $e_1, e_2, \ldots$ $(o_1, o_2, \ldots)$ is any enumeration of even (odd)[1] permutations of the symbols $\{\mathsf{1}, \ldots, \mathsf{d}\}$. Hence, the number of rewrite rules

[1] A permutation is called even (odd) if it can be written as a composition of an even (odd) number of transpositions.

in $\mathcal{S}_d$ exhibits factorial growth in the dimension d. In contrast, the systems created by our approach have constant size and the dimension d' is *optimal*, i.e., $d' = d + 1$.

The remainder of this chapter is organized as follows. In Section 6.1, we show that matrix interpretations over the reals are more powerful than matrix interpretations over the rationals, which are in turn more powerful than matrix interpretations over the natural numbers. Subsequently, we present our results on the dimension hierarchy related to matrix interpretations in Section 6.2, before concluding with suggestions for future research in Section 6.3.

The material presented in this chapter appeared in [56].

6.1 Domain Hierarchy

In this section, we first show that matrix interpretations over $\mathbb{R}$ subsume matrix interpretations over $\mathbb{Q}$, which in turn subsume matrix interpretations over $\mathbb{N}$. Then, in Sections 6.1.1 and 6.1.2, both inclusions are proved to be proper.

Lemma 6.1. *Let $\mathcal{M}$ be an n-dimensional matrix interpretation over $\mathbb{N}$ (not necessarily monotone), and let $\mathcal{S}_1$ and $\mathcal{S}_2$ be finite sets of rewrite rules such that $\mathcal{S}_1 \subseteq >_{\mathcal{M}}$ and $\mathcal{S}_2 \subseteq \geqslant_{\mathcal{M}}$. Then there exists an n-dimensional matrix interpretation $\mathcal{N}$ over $\mathbb{Q}$ such that $\mathcal{S}_1 \subseteq >_{\mathcal{N}}$ and $\mathcal{S}_2 \subseteq \geqslant_{\mathcal{N}}$. Moreover, $\mathcal{N}$ is monotone if and only if $\mathcal{M}$ is monotone.*

Proof. Let $\mathcal{F}$ denote the signature associated with $\mathcal{S}_1 \cup \mathcal{S}_2$. Then, by assumption, $\mathcal{M}$ associates each k-ary function symbol $f \in \mathcal{F}$ with a linear function $f_{\mathcal{M}}(\vec{x}_1, \dots, \vec{x}_k) = F_1\vec{x}_1 + \cdots + F_k\vec{x}_k + \vec{f}$, where $F_1, \dots, F_k \in \mathbb{N}^{n \times n}$ and $\vec{f} \in \mathbb{N}^n$, such that $\mathcal{S}_1 \subseteq >_{\mathcal{M}}$ and $\mathcal{S}_2 \subseteq \geqslant_{\mathcal{M}}$. Based on this interpretation, we define the matrix interpretation $\mathcal{N}$ by letting $\delta = 1$ and taking the same interpretation functions, i.e., $f_{\mathcal{N}}(\vec{x}_1, \dots, \vec{x}_k) = f_{\mathcal{M}}(\vec{x}_1, \dots, \vec{x}_k)$ for all $f \in \mathcal{F}$. Then $\mathcal{N}$ is well-defined, and it is monotone if and only if $\mathcal{M}$ is monotone. As to compatibility of $\mathcal{N}$ with $\mathcal{S}_1$, let us consider an arbitrary rewrite rule $\ell \to r \in \mathcal{S}_1$ and show that $\ell >_{\mathcal{M}} r$ implies $\ell >_{\mathcal{N}} r$, i.e., $[\alpha]_{\mathcal{N}}(\ell) >_\delta [\alpha]_{\mathcal{N}}(r)$ for all variable assignments α. Because of linearity of the interpretation functions, we can write $[\alpha]_{\mathcal{N}}(\ell) = L_1\vec{x}_1 + \cdots + L_m\vec{x}_m + \vec{\ell}$ and $[\alpha]_{\mathcal{N}}(r) = R_1\vec{x}_1 + \cdots + R_m\vec{x}_m + \vec{r}$, where the assignment α maps the variables $x_1, \dots, x_m$ occurring in ℓ, r to $\vec{x}_i = \alpha(x_i)$ for $i = 1, \dots, m$. Thus, it remains to show that the inequality

$$L_1\vec{x}_1 + \cdots + L_m\vec{x}_m + \vec{\ell} >_\delta R_1\vec{x}_1 + \cdots + R_m\vec{x}_m + \vec{r}$$

holds for all $\vec{x}_1, \dots, \vec{x}_m \in \mathbb{Q}_0^n$. This is exactly the case if $L_i \geqslant R_i$ for $i = 1, \dots, m$ and $\vec{\ell} >_\delta \vec{r}$, i.e., $\ell_i \geqslant r_i$ for $i = 2, \dots, n$ and $\ell_1 \geqslant r_1 + \delta = r_1 + 1$. Indeed, all these conditions follow from compatibility of $\mathcal{M}$ with $\ell \to r$ because, by the same reasoning as above (and since the interpretation functions of $\mathcal{M}$ and $\mathcal{N}$ coincide), $\ell >_{\mathcal{M}} r$ holds in $(\mathcal{M}, >, \geqslant)$ if and only if

$$L_1\vec{x}_1 + \cdots + L_m\vec{x}_m + \vec{\ell} > R_1\vec{x}_1 + \cdots + R_m\vec{x}_m + \vec{r}$$

holds for all $\vec{x}_1, \dots, \vec{x}_m \in \mathbb{N}^n$, which implies $L_i \geqslant R_i$ for $i = 1, \dots, m$ and $\vec{\ell} > \vec{r}$, that is, $\ell_i \geqslant r_i$ for $i = 2, \dots, n$ and $\ell_1 >_{\mathbb{N}} r_1$, the latter being equivalent to $\ell_1 \geqslant r_1 + 1$ as $\vec{\ell}, \vec{r} \in \mathbb{N}^n$. This shows compatibility of $\mathcal{N}$ with $\mathcal{S}_1$. Weak compatibility with $\mathcal{S}_2$ follows in the same way. $\square$

The essence of the proof of this lemma is that any matrix interpretation over $\mathbb{N}$ can be conceived as a matrix interpretation over $\mathbb{Q}$. Likewise, any matrix interpretation over $\mathbb{Q}$ can be conceived as a matrix interpretation over $\mathbb{R}$.

Lemma 6.2. *Let $\mathcal{M}$ be an n-dimensional matrix interpretation over $\mathbb{Q}$ (not necessarily monotone), and let $\mathcal{S}_1$ and $\mathcal{S}_2$ be finite sets of rewrite rules such that $\mathcal{S}_1 \subseteq >_{\mathcal{M}}$ and $\mathcal{S}_2 \subseteq \geqslant_{\mathcal{M}}$. Then there exists an n-dimensional matrix interpretation $\mathcal{N}$ over $\mathbb{R}$ such that $\mathcal{S}_1 \subseteq >_{\mathcal{N}}$ and $\mathcal{S}_2 \subseteq \geqslant_{\mathcal{N}}$. Moreover, $\mathcal{N}$ is monotone if and only if $\mathcal{M}$ is monotone.*

Proof. Similar to the proof of Lemma 6.1, with $\mathcal{N}$ defined by $\delta_{\mathcal{N}} = \delta_{\mathcal{M}} = \delta$ and $f_{\mathcal{N}}(\vec{x}_1, \ldots, \vec{x}_k) = f_{\mathcal{M}}(\vec{x}_1, \ldots, \vec{x}_k)$ for all $f \in \mathcal{F}$. $\square$

As an immediate consequence of the previous lemmata, we obtain the following corollary stating that matrix interpretations over $\mathbb{N}$ are no more powerful than matrix interpretations over $\mathbb{Q}$, which are in turn no more powerful than matrix interpretations over $\mathbb{R}$.

Corollary 6.3. *Let $\mathcal{R}$ be a TRS and $(\mathcal{P}, \mathcal{S})$ a DP problem.*

1. *If there is an (incremental) termination proof for $\mathcal{R}$ using monotone matrix interpretations over $\mathbb{N}$ (resp. $\mathbb{Q}$), then there is also one using monotone matrix interpretations over $\mathbb{Q}$ (resp. $\mathbb{R}$).*

2. *If a matrix interpretation over $\mathbb{N}$ (resp. $\mathbb{Q}$) succeeds on $(\mathcal{P}, \mathcal{S})$, then there is also a matrix interpretation over $\mathbb{Q}$ (resp. $\mathbb{R}$) of the same dimension that succeeds on $(\mathcal{P}, \mathcal{S})$.* $\square$

In the remainder of this section, we show that the converse statements do not hold.

6.1.1 Matrix Interpretations over the Rational Numbers

In order to show that matrix interpretations over $\mathbb{Q}$ are indeed more powerful than matrix interpretations over $\mathbb{N}$, let us first consider the TRS $\mathcal{S}$ consisting of the following rewrite rules:[2]

$$x + \mathsf{a} \to x \tag{6.1}$$
$$x + \mathsf{a} \to (x + \mathsf{b}) + \mathsf{b} \tag{6.2}$$
$$\mathsf{a} + x \to x \tag{6.3}$$
$$\mathsf{a} + x \to \mathsf{b} + (\mathsf{b} + x) \tag{6.4}$$

This TRS will turn out to be very helpful for our purposes, not only in the current subsection but also in the subsequent one. This is due to the following property, which holds for matrix interpretations over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$.

[2] I thank Bertram Felgenhauer for his help in finding these rules.

Lemma 6.4. *Let $\mathcal{M}$ be a matrix interpretation (not necessarily monotone) with carrier set M such that $\mathcal{S} \subseteq \geqslant_{\mathcal{M}}$. Then $+_{\mathcal{M}}(\vec{x}, \vec{y}) = \vec{x} + \vec{y} + \vec{v}$, $\vec{v} \in M$.*

Proof. Without loss of generality, let $+_{\mathcal{M}}(\vec{x}, \vec{y}) = A_1\vec{x} + A_2\vec{y} + \vec{v}$, $\vec{v} \in M$. As $\mathcal{M}$ is weakly compatible with the rule (6.1), we obtain $A_1 \geqslant I$; hence, $A_1^2 \geqslant A_1$ due to non-negativity of A_1. Similarly, by weak compatibility with (6.2), we infer $A_1 \geqslant A_1^2$, which implies $A_1^2 = A_1 \geqslant I$ together with the previous result. Yet this means that A_1 must in fact be equal to I. To this end, we observe that $A_1 \geqslant I$ implies $(A_1 - I)^2 \geqslant 0$, which simplifies to $I \geqslant 2A_1 - A_1^2 = A_1$; hence, $A_1 = I$. In the same way, we obtain $A_2 = I$ from the compatibility constraints associated with (6.3) and (6.4). □

So in any matrix interpretation that is weakly compatible with the TRS $\mathcal{S}$ the symbol $+$ must be interpreted by a function $+_{\mathcal{M}}(\vec{x}, \vec{y}) = \vec{x} + \vec{y} + \vec{v}$ that models addition of two elements of the underlying carrier set (modulo adding a constant). The inherent possibility to count objects can be exploited to show that matrix interpretations over $\mathbb{Q}$ are indeed more powerful than matrix interpretations over $\mathbb{N}$. For this purpose, we extend the TRS $\mathcal{S}$ by the rules (6.5) and (6.6), calling the resulting system $\mathcal{T}_{\mathbb{Q}}$:

$$((x + x) + x) + \mathsf{a} \to \mathsf{g}(x + x) \tag{6.5}$$

$$\mathsf{g}(x + x) \to (x + x) + x \tag{6.6}$$

By construction, this TRS is not compatible, not even weakly compatible, with any matrix interpretation over $\mathbb{N}$.

Lemma 6.5. *Let $\mathcal{M}$ be an n-dimensional matrix interpretation (not necessarily monotone) with carrier set M such that $\mathcal{T}_{\mathbb{Q}} \subseteq \geqslant_{\mathcal{M}}$. Then $M \neq \mathbb{N}^n$.*

Proof. As $\mathcal{M}$ is weakly compatible with $\mathcal{T}_{\mathbb{Q}}$, it is also weakly compatible with the TRS $\mathcal{S}$. So, by Lemma 6.4, the function symbol $+$ must be interpreted by $+_{\mathcal{M}}(\vec{x}, \vec{y}) = \vec{x} + \vec{y} + \vec{v}$, $\vec{v} \in M$. Assuming $\mathsf{g}_{\mathcal{M}}(\vec{x}) = G\vec{x} + \vec{g}$ without loss of generality, we obtain $3I \geqslant 2G$ from weak compatibility of $\mathcal{M}$ with (6.5) and $2G \geqslant 3I$ from weak compatibility with (6.6); hence, $G = \frac{3}{2}I \notin \mathbb{N}^{n \times n}$. Therefore, $\mathcal{M}$ cannot be a matrix interpretation over $\mathbb{N}$. □

The previous lemma, together with the observation that the TRS $\mathcal{T}_{\mathbb{Q}}$ admits a compatible matrix interpretation over $\mathbb{Q}$, directly leads to the main result of this subsection.

Theorem 6.6.

1. *The TRS $\mathcal{T}_{\mathbb{Q}}$ is terminating. In particular, $\mathcal{T}_{\mathbb{Q}}$ is compatible with a monotone matrix interpretation over $\mathbb{Q}$.*

2. *There cannot be an (incremental) termination proof of $\mathcal{T}_{\mathbb{Q}}$ using only monotone matrix interpretations over $\mathbb{N}$.*

3. *No matrix interpretation over $\mathbb{N}$ succeeds on the DP problem $(\_, \mathcal{T}_{\mathbb{Q}})$.*

Proof. The last two statements are immediate consequences of Lemma 6.5. As to the first claim, the following monotone 1-dimensional matrix interpretation (i.e., linear polynomial interpretation) over $\mathbb{Q}$ is compatible with $\mathcal{T}_{\mathbb{Q}}$: $\delta = 1$, $\mathsf{a}_{\mathcal{M}} = 2$, $\mathsf{b}_{\mathcal{M}} = 0$, $\mathsf{g}_{\mathcal{M}}(x) = \frac{3}{2}x + 1$ and $+_{\mathcal{M}}(x, y) = x + y$. □

6.1.2 Matrix Interpretations over the Real Numbers

Next we show that matrix interpretations over $\mathbb{R}$ are more powerful than matrix interpretations over $\mathbb{Q}$. For this purpose, we extend the TRS $\mathcal{S}$ of the previous subsection by the rules (6.7) – (6.9) and call the resulting system $\mathcal{T}_{\mathbb{R}}$:

$$(x + x) + \mathsf{a} \to \mathsf{k}(\mathsf{k}(x)) \tag{6.7}$$
$$\mathsf{k}(\mathsf{k}(x)) \to x + x \tag{6.8}$$
$$\mathsf{k}(x) \to x \tag{6.9}$$

By construction, this TRS admits only matrix interpretations over $\mathbb{R}$.

Lemma 6.7. *Let $\mathcal{M}$ be an n-dimensional matrix interpretation (not necessarily monotone) with carrier set M such that $\mathcal{T}_{\mathbb{R}} \subseteq \geqslant_{\mathcal{M}}$. Then $M \neq \mathbb{N}^n$ and $M \neq \mathbb{Q}_0^n$.*

Proof. As the TRS $\mathcal{S}$ is a subsystem of $\mathcal{T}_{\mathbb{R}}$, $\mathcal{T}_{\mathbb{R}} \subseteq \geqslant_{\mathcal{M}}$ implies $\mathcal{S} \subseteq \geqslant_{\mathcal{M}}$. Hence, by Lemma 6.4, the function symbol $+$ must be interpreted by $+_{\mathcal{M}}(\vec{x}, \vec{y}) = \vec{x} + \vec{y} + \vec{v}$, where $\vec{v} \in M$. Assuming $\mathsf{k}_{\mathcal{M}}(\vec{x}) = K\vec{x} + \vec{k}$ without loss of generality, the (weak) compatibility constraint associated with rule (6.7) implies $2I \geqslant K^2$. We also have $K^2 \geqslant 2I$ by weak compatibility with (6.8) and $K \geqslant I$ due to (6.9). Hence, the $n \times n$ square matrix K must satisfy the following conditions:

$$K^2 = 2I \qquad \text{and} \qquad K \geqslant I \tag{6.10}$$

Clearly, for dimension $n = 1$, the unique solution is $K = \sqrt{2}$; in particular, K is not a rational number. In fact, for any dimension $n \geqslant 1$, the unique solution turns out to be $K = \sqrt{2}I$. To this end, let us first show that the conditions given in (6.10) imply that K is a diagonal matrix. Because of $K \geqslant I$, we can write $K = I + N$ for some non-negative matrix N. Then $K^2 = 2I$ if and only if $N^2 + 2N = I$. Now non-negativity of N implies $I \geqslant N$. Hence, N is a diagonal matrix and therefore also K. So all entries of K^2 are zero except its diagonal entries: $(K^2)_{ii} = K_{ii}^2$ for $i = 1, \ldots, n$. But then K_{ii} must be $\sqrt{2}$ in order to satisfy $K^2 = 2I$ and $K \geqslant I$. In other words, $K = \sqrt{2}I \notin \mathbb{Q}_0^{n \times n}$. Therefore, $\mathcal{M}$ cannot be a matrix interpretation over $\mathbb{N}$ or $\mathbb{Q}$. □

Remark 6.8. Rule (6.9) is essential for the statement of Lemma 6.7. Without it, the conditions given in (6.10) would turn into $K^2 = 2I$ and $K \geqslant 0$, the conjunction of which is satisfiable over $\mathbb{N}^{n \times n}$; for example, by choosing

$$\mathsf{a}_{\mathcal{M}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \mathsf{b}_{\mathcal{M}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \mathsf{k}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad +_{\mathcal{M}}(\vec{x}, \vec{y}) = \vec{x} + \vec{y}$$

we obtain a non-monotone 2-dimensional matrix interpretation over $\mathbb{N}$ that is compatible with the TRS $\mathcal{T}_{\mathbb{R}} \setminus \{(6.9)\}$. However, in case monotonicity of the matrix interpretation in Lemma 6.7 is explicitly required, rule (6.9) becomes superfluous because $K_{11} \geqslant 1$ and $K^2 = 2I$ imply that all entries of the first row and the first column of K are zero except K_{11} (as K must be non-negative). This means that $(K^2)_{11} = K_{11}^2$, so K_{11} must be equal to $\sqrt{2}$, hence irrational, in order to satisfy $K^2 = 2I$.

Lemma 6.7 shows that no matrix interpretation over $\mathbb{N}$ or $\mathbb{Q}$ is weakly compatible with the TRS $\mathcal{T}_\mathbb{R}$. However, $\mathcal{T}_\mathbb{R}$ can be shown terminating by a compatible matrix interpretation over $\mathbb{R}$.

Theorem 6.9.

1. *The TRS $\mathcal{T}_\mathbb{R}$ is terminating. In particular, $\mathcal{T}_\mathbb{R}$ is compatible with a monotone matrix interpretation over $\mathbb{R}$.*

2. *There cannot be an (incremental) termination proof of $\mathcal{T}_\mathbb{R}$ using only monotone matrix interpretations over $\mathbb{N}$ or $\mathbb{Q}$.*

3. *No matrix interpretation over $\mathbb{N}$ or $\mathbb{Q}$ succeeds on the DP problem $(\_, \mathcal{T}_\mathbb{R})$.*

Proof. The last two claims are immediate consequences of Lemma 6.7. Finally, the first claim holds by the following monotone 1-dimensional matrix interpretation over $\mathbb{R}$ that is compatible with $\mathcal{T}_\mathbb{R}$: $\delta = 1$, $\mathsf{a}_\mathcal{M} = 4$, $\mathsf{b}_\mathcal{M} = 0$, $\mathsf{k}_\mathcal{M}(x) = \sqrt{2}x + 1$ and $+_\mathcal{M}(x, y) = x + y$. $\square$

6.2 Dimension Hierarchy

Unlike the previous section, where we have established a hierarchy of matrix interpretations regarding the domain of the matrix entries, the purpose of this section is to examine matrix interpretations with respect to their dimension. That is, we fix $D \in \{\mathbb{N}, \mathbb{Q}_0, \mathbb{R}_0\}$ and consider matrix interpretations over the family of carrier sets $(D^n)_{n \geqslant 1}$. The main result is that the inherent termination hierarchy is infinite with respect to the dimension n, with each level of the hierarchy properly subsuming its predecessor. In other words, $(n+1)$-dimensional matrix interpretations are strictly more powerful for proving termination than n-dimensional matrix interpretations (for any $n \geqslant 1$). We show this by constructing a family of TRSs $(\mathcal{T}_k)_{k \geqslant 2}$ having the property that any of its members $\mathcal{T}_k$ can only be handled with matrix interpretations of dimension at least k. The construction is based on the idea of encoding (i.e., specifying) the degree of the minimal polynomial $\mathrm{m}_A(x)$ of some matrix A occurring in a matrix interpretation in terms of rewrite rules. Thus, if $\mathcal{M}$ is an n-dimensional matrix interpretation such that the degree of the minimal polynomial of some matrix is fixed to a value of k, then the degree of the characteristic polynomial of this matrix must be at least k, i.e., $n \geqslant k$ (since the minimal polynomial divides the characteristic polynomial whose degree is n). In other words, the dimension n of $\mathcal{M}$ must then be at least k. The family of TRSs $(\mathcal{T}_k)_{k \geqslant 2}$ mentioned above is made up as follows. For any natural number $k \geqslant 2$, $\mathcal{T}_k$ denotes the union of the

TRS $\mathcal{S}$ of Section 6.1 and the following rewrite rules:

$$\mathsf{f}^k(x) + \mathsf{d} \to \mathsf{f}^{k-1}(x) + \mathsf{c} \tag{6.11}$$
$$\mathsf{f}^{k-1}(x) + \mathsf{c} \to \mathsf{f}^k(x) \tag{6.12}$$
$$\mathsf{h}(\mathsf{f}^{k-2}(\mathsf{h}(x))) \to \mathsf{h}(\mathsf{f}^{k-1}(\mathsf{h}(x))) + x \tag{6.13}$$
$$\mathsf{h}(\mathsf{f}^{k-1}(\mathsf{h}(x))) \to x \tag{6.14}$$

The intuition is that if $\mathcal{M}$ is an n-dimensional matrix interpretation that is weakly compatible with all rules of $\mathcal{T}_k$, then the minimal polynomial $\mathrm{m}_F(x)$ of the matrix F associated with the interpretation of the unary function symbol f is forced to be equal to the polynomial $p_k(x) = x^k - x^{k-1}$, a monic polynomial of degree k. This is the purpose of the rules (6.11) – (6.14). More precisely, the first two rules ensure that $p_k(x)$ annihilates F, whereas the latter two specify that $p_k(x)$ is the monic polynomial of least degree having this property.

Lemma 6.10. *Let $\mathcal{M}$ be an n-dimensional matrix interpretation (not necessarily monotone), and let $k \geqslant 2$ be a natural number. Then $\mathcal{T}_k \subseteq \geqslant_{\mathcal{M}}$ implies $n \geqslant k$.*

Proof. Let us assume $\mathcal{T}_k \subseteq \geqslant_{\mathcal{M}}$. Then we also have $\mathcal{S} \subseteq \geqslant_{\mathcal{M}}$ because the TRS $\mathcal{S}$ is contained in $\mathcal{T}_k$. Therefore, the function symbol $+$ must be interpreted by $+_{\mathcal{M}}(\vec{x}, \vec{y}) = \vec{x} + \vec{y} + \vec{v}$ according to Lemma 6.4. Assuming, without loss of generality, that $\mathsf{f}_{\mathcal{M}}(\vec{x}) = F\vec{x} + \vec{f}$ and $\mathsf{h}_{\mathcal{M}}(\vec{x}) = H\vec{x} + \vec{h}$, the (weak) compatibility constraint associated with (6.11) implies $F^k \geqslant F^{k-1}$. We also have $F^{k-1} \geqslant F^k$ due to (6.12); hence, $F^k = F^{k-1}$. Next we consider the compatibility constraints associated with (6.13) and (6.14). From the former we infer $HF^{k-2}H \geqslant HF^{k-1}H + I$, which implies $F^{k-2} \neq F^{k-1}$, whereas the latter enforces $HF^{k-1}H \geqslant I$, which implies $F^{k-1} \neq 0$. Thus, the $n \times n$ square matrix F must satisfy the following conditions:

$$F^k = F^{k-1} \qquad F^{k-2} \neq F^{k-1} \qquad F^{k-1} \neq 0 \tag{6.15}$$

These conditions imply that the minimal polynomial of F must be equal to the polynomial $p_k(x) = x^k - x^{k-1}$; i.e., $\mathrm{m}_F(x) = x^k - x^{k-1}$. In order to show this, we first observe that $F^k = F^{k-1}$ means that the polynomial $p_k(x)$ annihilates the matrix F. So $\mathrm{m}_F(x)$ divides $p_k(x)$. Writing $p_k(x) = (x-1)x^{k-1}$ as a product of irreducible factors, we see that if $\mathrm{m}_F(x) \neq p_k(x)$ (i.e., $\mathrm{m}_F(x)$ is a proper divisor of $p_k(x)$ of degree at most $k-1$), then $\mathrm{m}_F(x)$ must divide the polynomial $(x-1)x^{k-2}$ or the polynomial x^{k-1} (depending on whether $(x-1)$ occurs as a factor in $\mathrm{m}_F(x)$ or not). As in both cases the corresponding polynomial annihilates F, we obtain $F^{k-2} = F^{k-1}$ or $F^{k-1} = 0$, contradicting (6.15). Consequently, $p_k(x)$ must indeed be the minimal polynomial of F, and since it divides the characteristic polynomial of F, the degree of the latter must be greater than or equal to the degree of the former, that is, $n \geqslant k$. □

Remark 6.11. If one explicitly requires monotonicity of the matrix interpretation $\mathcal{M}$ in Lemma 6.10, then the condition $F^{k-1} \neq 0$ is automatically satisfied, such that rule (6.14) becomes superfluous in this case.

Lemma 6.10 shows that no matrix interpretation of dimension less than k can be weakly compatible with the TRS $\mathcal{T}_k$. However, $\mathcal{T}_k$ can be shown terminating by a compatible matrix interpretation of dimension k.

Theorem 6.12. *Let $k \geqslant 2$.*

1. *The TRS $\mathcal{T}_k$ is terminating. In particular, $\mathcal{T}_k$ is compatible with a monotone matrix interpretation over $\mathbb{N}$ of dimension k.*
2. *There cannot be an (incremental) termination proof of $\mathcal{T}_k$ using only monotone matrix interpretations of dimension less than k.*
3. *No matrix interpretation of dimension less than k succeeds on the DP problem $(\_, \mathcal{T}_k)$.*

Proof. The last two claims are immediate consequences of Lemma 6.10. The first claim holds by the following monotone k-dimensional matrix interpretation over $\mathbb{N}$ that is compatible with $\mathcal{T}_k$:

$$\mathsf{a}_{\mathcal{M}} = \mathsf{c}_{\mathcal{M}} = (1, 0, \dots, 0)^{\mathrm{T}} \qquad \mathsf{b}_{\mathcal{M}} = 0 \qquad \mathsf{d}_{\mathcal{M}} = 2\,\mathsf{a}_{\mathcal{M}}$$
$$+_{\mathcal{M}}(\vec{x}, \vec{y}) = \vec{x} + \vec{y} \qquad \mathsf{f}_{\mathcal{M}}(\vec{x}) = F\vec{x} \qquad \mathsf{h}_{\mathcal{M}}(\vec{x}) = H\vec{x} + \vec{h}$$

where $\vec{h} = (1, \dots, 1)^{\mathrm{T}}$, all rows of H have the shape $(1, 2, 1, \dots, 1)$ and F is zero everywhere except for the entries F_{11} and $F_{i,i+1}$, $i = 1, \dots, k-1$, which are all set to one:

$$F = \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 \end{pmatrix} \qquad H = \begin{pmatrix} 1 & 2 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 2 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 1 & 1 & \cdots & 1 \end{pmatrix}$$ □

6.3 Conclusion

In this chapter, we have established two hierarchies of matrix interpretations. On the one hand, there is the domain hierarchy stating that matrix interpretations over the real numbers are more powerful with respect to proving termination than matrix interpretations over the rational numbers, which are in turn more powerful than matrix interpretations over the natural numbers (cf. Figure 6.2). On the other hand, we have established a hierarchy of matrix interpretations with respect to matrix dimension, which was shown to be infinite, with each level properly subsuming its predecessor (cf. Figure 6.3).

Both hierarchies hold in the setting of the DP framework as well as for matrix interpretations as a stand-alone termination technique. Concerning the former, we remark that the corresponding results in Theorems 6.6, 6.9 and 6.12 do not only hold for standard reduction pairs but also for reduction pairs incorporating usable rules (as defined in Section 1.4) because all rules of the TRSs $\mathcal{T}_{\mathbb{Q}}$, $\mathcal{T}_{\mathbb{R}}$ and $\mathcal{T}_k$ are usable. It is an easy exercise to make our TRSs also withstand

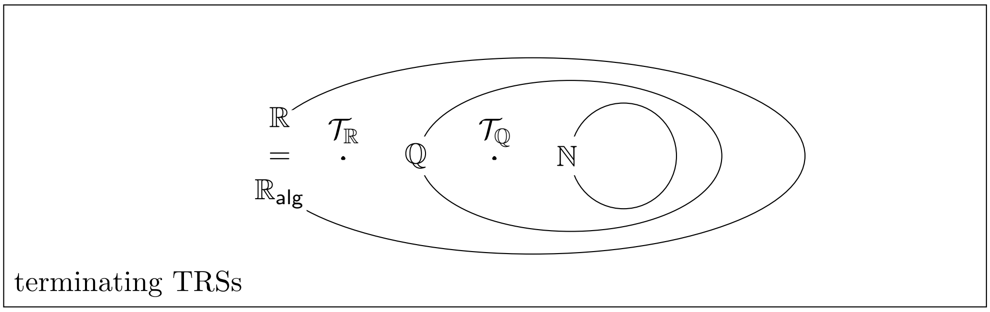

Figure 6.2: The domain hierarchy.

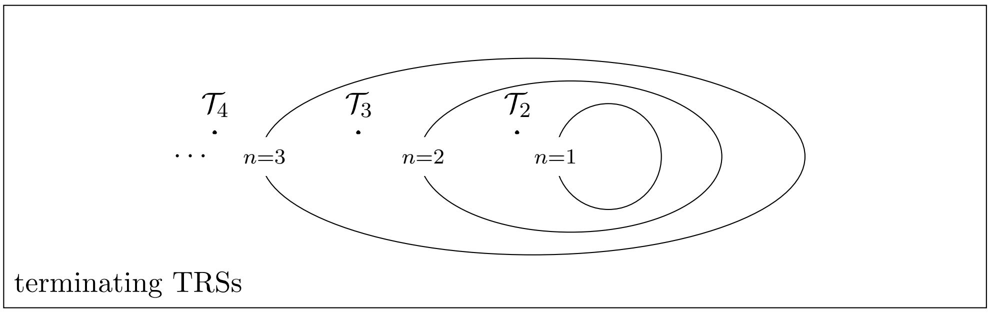

Figure 6.3: The dimension hierarchy.

reduction pairs that incorporate usable rules with (implicit) *argument filters* [26] (induced by matrix interpretations).

Our results concerning the domain hierarchy provide a definitive answer to a question raised in [48] whether *rational numbers are somehow unnecessary when dealing with matrix interpretations*. The answer is in the negative, so the attempt of [48] to simulate matrix interpretations over $\mathbb{Q}$ with matrix interpretations over $\mathbb{N}$ (of higher dimension) must necessarily remain incomplete.

Moreover, we remark that the results of this chapter do not only apply to the standard variant of matrix interpretations of Endrullis *et al.* [20] (though the technical part of the chapter refers to it) but also to the kinds of matrix interpretations introduced in Appendix B and [55] (which are based on various different well-founded orders on vectors of natural numbers) and extensions thereof to vectors of non-negative rational and real numbers. On the technical level, this is due to the fact that our main Lemmata 6.5, 6.7 and 6.10 only require weak compatibility (rather than strict) and do not demand monotonicity of the respective matrix interpretations. Also note that the interpretations given in the proofs of Theorems 6.6, 6.9 and 6.12 can be conceived as matrix interpretations over the base order $>^{w}_{\Sigma}$, which relates two vectors $\vec{x}$ and $\vec{y}$ if and only if there is a weak decrease in every single component of the vectors and a strict decrease with respect to the sum of the components of $\vec{x}$ and $\vec{y}$ (cf. Appendix B or [55]). We also expect our results to carry over to the matrix interpretations of [17]. For linear interpretations, this should be possible without further ado, whereas non-linear interpretations conceivably require the addition of new rules enforcing linearity of the interpretations of *some* function symbols (e.g. by using techniques from Chapter 4).

Next we consider the relationship between traditional matrix interpretations over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$ and arctic matrix interpretations [38], a variant of the matrix interpretation method where the usual operations of addition and multiplication on $\mathbb{N}$ ($\mathbb{Q}$, $\mathbb{R}$) are generalized to the arctic semi-ring, the natural numbers extended with $-\infty$, where semi-ring addition $\oplus$ is the "max"-operation and semi-ring multiplication $\otimes$ is the standard addition operation. In [38] the authors provide evidence that the arctic matrix method does not subsume the traditional one. The converse statement is left as an open issue, but experiments performed by the authors suggest that neither method subsumes the other. We fill this gap by presenting a TRS that admits arctic matrix interpretations while precluding traditional ones.

Lemma 6.13. *There is no matrix interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$ that is weakly compatible with the following TRS:*

$$x + \mathsf{a} \to x \tag{6.16}$$
$$\mathsf{a} + x \to x \tag{6.17}$$
$$\mathsf{f}(x) \to x \tag{6.18}$$
$$\mathsf{f}(x) + \mathsf{b} \to \mathsf{f}(x + x) \tag{6.19}$$

Proof. Assume to the contrary that $\mathcal{M}$ is a weakly compatible matrix interpretation (not necessarily monotone) over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$. Without loss of generality, let $+_{\mathcal{M}}(\vec{x}, \vec{y}) = A_1\vec{x} + A_2\vec{y} + \vec{v}$ and $\mathsf{f}_{\mathcal{M}}(\vec{x}) = F\vec{x} + \vec{f}$. As $\mathcal{M}$ is weakly compatible with (6.16) and (6.17), we obtain $A_1 \geqslant I$ and $A_2 \geqslant I$. Similarly, weak compatibility with (6.18) and (6.19) yields $F \geqslant I$ and $A_1F \geqslant F(A_1 + A_2)$. Denoting by $\mathrm{tr}(M)$ the *trace* of a square matrix M, i.e., the sum of its diagonal entries, the latter inequality implies $\mathrm{tr}(A_1F) \geqslant \mathrm{tr}(FA_1 + FA_2) = \mathrm{tr}(FA_1) + \mathrm{tr}(FA_2)$. Due to the well-known fact that $\mathrm{tr}(A_1F) = \mathrm{tr}(FA_1)$, this inequality can only hold if $\mathrm{tr}(FA_2) = 0$, which contradicts the fact that $FA_2 \geqslant I$ (since $F \geqslant I$ and $A_2 \geqslant I$). $\square$

The following one-dimensional arctic matrix interpretation $\mathcal{M}$, inducing a weakly monotone algebra $(\mathcal{M}, >_{\mathbb{N}}, \geqslant_{\mathbb{N}})$ with carrier $\mathbb{N}$ according to the proof of [38, Theorem 12], is weakly compatible with all rules and strictly compatible with the first two rules of the above TRS:

$$\mathsf{a}_{\mathcal{M}} = \mathsf{b}_{\mathcal{M}} = 0 \qquad \mathsf{f}_{\mathcal{M}}(x) = x \oplus 0 \qquad +_{\mathcal{M}}(x, y) = (1 \otimes x) \oplus (1 \otimes y) \oplus 0$$

(Note that strict monotonicity cannot hold as there are no strictly monotone linear arctic functions of more than one argument [38].)

We conclude with a remark on future work and related work. For future work, we mention the extension of the results of this chapter to more restrictive classes of TRSs like left-linear ones and SRSs. In this context, we also note that the partial result of [23] showing that the dimension hierarchy is infinite applies without further ado since the underlying construction is based on SRSs in contrast to our approach of Section 6.2.

Chapter 7

Derivational Complexity

This chapter is concerned with automated complexity analysis of TRSs via matrix interpretations. Given a terminating TRS, the aim is to obtain (quantitative) information about the maximal length of rewrite sequences (or derivations) in terms of the size of their initial term. This is known as derivational complexity. Developing (automatic) methods for bounding the derivational complexity of TRSs has become an active and competitive[1] research area in the past few years (e.g. [24,38,51–53,59,73,75]), and matrix interpretations play an important role in this context, especially for establishing polynomial bounds.

7.1 Introduction

Many powerful techniques for termination analysis of TRSs have been developed in the course of time, most of which have been automated successfully, as is evident in the results of the (annual) international competition for termination and complexity tools.[2] Moreover, Hofbauer and Lautemann observe in [31] that "proving termination with one of these specific techniques in general proves more than just the absence of infinite derivations. It turns out that in many cases such a proof implies an upper bound on the maximal length of derivations", which they consider as a natural measure for the complexity of (terminating) TRSs. More precisely, the resulting notion of derivational complexity relates the length of a longest derivation to the size of its initial term. While Hofbauer and Lautemann allow arbitrary terms as initial terms of derivations, Hirokawa and Moser [29, 51] consider a variant of complexity, called runtime complexity, that relates the length of a longest derivation to the size of the arguments of the initial term, where the arguments are required to be in normal form. This notion of complexity is more reminiscent of term rewriting being the foundation of much of declarative programming. In the remainder of this chapter, we adopt the notion of derivational complexity as our central definition of the complexity of a TRS. (But our results can easily be adapted to runtime complexity.)

As mentioned above, for a given terminating TRS, one can often estimate its derivational complexity by examining the proof method that established its termination. For example, polynomial interpretations imply a double exponential upper bound on the derivational complexity [31]. However, as term rewriting is a model of computation and algorithms of polynomial complexity are widely

[1] http://www.termination-portal.org/wiki/Complexity/

[2] http://termcomp.uibk.ac.at

accepted as *feasible*, one is especially interested in polynomial derivational complexity. But currently only few techniques for establishing feasible upper complexity bounds are known. Commonly, they are stripped-down variants of existing termination techniques. For example, if a TRS can be shown terminating by means of a compatible *strongly linear* polynomial interpretation (i.e., a polynomial interpretation, where each n-ary function symbol f is interpreted by a function of the shape $f(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i + c$), then its derivational complexity is at most linear (cf. e.g. [29, 51]). In addition, linear derivational complexity can be inferred using the match-bounds technique [24] or arctic matrix interpretations [38]. Also the standard matrix interpretations (over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$) of Chapter 5 can readily be used to establish upper bounds on the derivational complexity of compatible TRSs. In fact, they are the most important technique for obtaining non-linear polynomial bounds. However, in general, the induced bounds are exponential [20,33]. In order to obtain polynomial upper bounds, the matrices used in such interpretations must satisfy certain (additional) restrictions, the study of which is the the main objective of the present chapter.

So what are the conditions ensuring polynomial boundedness of matrix interpretations over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$? Historically, the first approach appearing in the literature [53] was developed for matrix interpretations over $\mathbb{N}$, achieving its goal by restricting the shape of all matrices to upper triangular form. In this way, one obtains a method for establishing polynomial derivational complexity, where the degree of the polynomial bound corresponds to the dimension of the matrices. In [73] this method of triangular matrix interpretations was subsumed by an automata-based approach, where matrices are viewed as weighted (word) automata computing a weight function, which is required to be polynomially bounded. The result is a complete characterization (i.e., necessary and sufficient conditions) of polynomially bounded matrix interpretations over $\mathbb{N}$. In parallel to, but independent of [73], we developed an algebraic approach for polynomial boundedness of matrix interpretations, which will be presented in the remainder of this chapter. Using techniques from linear algebra, we provide a generalization of the method of triangular matrix interpretations that does not restrict the shape of the matrices but nevertheless induces polynomial upper bounds on the derivational complexity of compatible TRSs. In particular, we show how one can infer tighter bounds from triangular matrix interpretations by examining the diagonal structure of upper triangular (complexity) matrices. In contrast to [53, 73], our approach can handle matrix interpretations over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$. Moreover, using joint spectral radius theory [36, 37], a branch of mathematics dedicated to studying the growth rate of products of matrices taken from a set, it naturally extends to a complete characterization of polynomially bounded matrix interpretations over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$.

The remainder of this chapter is organized as follows. Section 7.2 introduces basic notions of term rewriting related to derivational complexity as well as some mathematical prerequisites. In Section 7.3, we review matrix interpretations in the context of complexity analysis of term rewriting and define the notion of polynomial boundedness. Then, in Section 7.4, we present an algebraic approach for achieving polynomial boundedness of an interpretation based on the spectral radius of its maximum matrix. This approach is subsequently extended in Section 7.5 to a complete characterization of polynomially bounded matrix interpretations over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$ using joint spectral radius theory. In Section 7.6, we give details on implementation-specific issues, before concluding in Section 7.7.

The main results of this chapter originate from [59] (the spectral radius approach of Section 7.4) and [49] (the joint spectral radius approach of Section 7.5). We improve upon the results of [59] by considering the minimal polynomial (instead of the characteristic polynomial) associated with the componentwise maximum matrix of a matrix interpretation. Moreover, a comprehensive journal article encompassing the content of [49, 59, 73] is in preparation.

7.2 Preliminaries

Complementing the background material of Chapter 5, we introduce the following additional preliminaries.

Term Rewriting. We define the notion of derivational complexity [31] of TRSs as follows. For a finite and terminating TRS $\mathcal{R}$ over a finite signature, the *derivation height* of a term t with respect to $\mathcal{R}$ is defined as

$$\mathsf{dh}(t, \to_{\mathcal{R}}) = \max\{n \mid t \to_{\mathcal{R}}^{n} u \text{ for some term } u\}$$

The *derivational complexity* of $\mathcal{R}$ is the function

$$\mathsf{dc}_{\mathcal{R}} \colon \mathbb{N} \setminus \{0\} \to \mathbb{N}, k \mapsto \max\{\mathsf{dh}(t, \to_{\mathcal{R}}) \mid |t| \leqslant k\}$$

It computes the maximal derivation height of all terms up to a given size. We say that the derivational complexity of $\mathcal{R}$ is linear (quadratic, cubic) if $\mathsf{dc}_{\mathcal{R}}(k)$ can be bounded by a linear (quadratic, cubic) polynomial in k, i.e., $\mathsf{dc}_{\mathcal{R}}(k) \in O(k^d)$ for $d = 1$ $(2, 3)$. Here, for functions $f, g \colon \mathbb{N} \to \mathbb{R}_0$, we write $f(n) \in O(g(n))$ if there are positive constants c and N such that $f(n) \leqslant c \cdot g(n)$ for all $n \geqslant N$. We call $f(n)$ *asymptotically bounded by a polynomial in* n *of degree* d if $f(n) \in O(n^d)$. Furthermore, $f(n) \in \Theta(g(n))$ if $f(n) \in O(g(n))$ and $g(n) \in O(f(n))$. In this case, we say that f and g are *asymptotically equivalent*.

Let $\mathcal{M}$ be a matrix interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$, t an arbitrary term and α_0 the assignment that maps every variable to 0. In the sequel, we abbreviate $[\alpha_0]_{\mathcal{M}}(t)$ by $[t]_{\mathcal{M}}$ (or just $[t]$ if $\mathcal{M}$ is clear from the context).

Lemma 7.1. *In the situation above,* $[\alpha]_{\mathcal{M}}(t) \geqslant [t]_{\mathcal{M}}$ *for all assignments* α.

Proof. By induction on the structure of t. □

Linear Algebra. Let $\mathbb{C}$ denote the field of complex numbers, A a square matrix from $\mathbb{R}^{n \times n}$ and p a polynomial from $\mathbb{R}[x]$. We denote the *multiplicity of a root* $\lambda \in \mathbb{C}$ of p by $\#p(\lambda)$. The *multiplicity of an eigenvalue* of A is the multiplicity of the corresponding root of the characteristic polynomial, and the *spectral radius* $\rho(A)$ of A is the maximum of the absolute values of its eigenvalues, i.e., $\rho(A) = \max\{|\lambda| \mid \lambda \text{ is an eigenvalue of } A\}$. The matrix A is *upper triangular* if all entries below its main diagonal are zero. One easily verifies that the diagonal entries of such a matrix give the multiset of its eigenvalues. An *upper triangular complexity matrix* is a non-negative upper triangular matrix whose diagonal entries are at most one. We call a matrix interpretation (over $\mathbb{N}$, $\mathbb{Q}$ or $\mathbb{R}$) *triangular* if all matrices occurring in it are upper triangular complexity matrices.

Following [35], we call a function $\|\cdot\| \colon \mathbb{R}^{n \times n} \to \mathbb{R}$ a *matrix norm* if for all matrices $A, B \in \mathbb{R}^{n \times n}$ it satisfies:

1. $\|A\| \geqslant 0$, with $\|A\| = 0$ if and only if $A = 0$
2. $\|cA\| = |c| \cdot \|A\|$ for all $c \in \mathbb{R}$
3. $\|A + B\| \leqslant \|A\| + \|B\|$
4. $\|AB\| \leqslant \|A\| \cdot \|B\|$

The (entrywise) l_1 *norm* of a matrix $A \in \mathbb{R}^{n \times n}$ is given by $\|A\|_1 = \sum_{1 \leqslant i,j \leqslant n} |A_{ij}|$.

With any matrix $A \in \mathbb{R}_0^{n \times n}$ we associate a directed (weighted) graph $G(A)$ on n vertices, numbered from 1 to n, such that there is a directed edge (of weight A_{ij}) in $G(A)$ from i to j if and only if $A_{ij} > 0$. In this context, A is said to be the *adjacency matrix* of the graph $G(A)$. The *weight of a path* in $G(A)$ is the product of the weights of its edges. With a (non-empty) finite set of matrices $S \subseteq \mathbb{R}_0^{n \times n}$ we associate the directed (weighted) graph $G(S) := G(M)$, where M denotes the componentwise maximum of the matrices in S, i.e., $M_{ij} = \max \{A_{ij} \mid A \in S\}$ for all $1 \leqslant i, j \leqslant n$. Following [36], we define a directed graph $G^k(S)$ for $k \geqslant 2$ on n^k vertices, representing ordered tuples of vertices of $G(S)$, such that there is an edge from vertex $(i_1, \ldots, i_k)$ to $(j_1, \ldots, j_k)$ if and only if there is a matrix $A \in S$ with $A_{i_\ell j_\ell} > 0$ for all $\ell = 1, \ldots, k$.

Recurrence Relations. Informally, a recurrence relation is an equation that recursively defines a sequence. Each element of the sequence is defined as a function of the preceding elements. For example, the Fibonacci numbers are defined by $F_n = F_{n-1} + F_{n-2}$, $n \geqslant 2$, with the initial conditions $F_0 = 0$ and $F_1 = 1$. Solving a recurrence relation means obtaining a closed-form solution; in this example, a non-recursive function of n.

A *linear homogeneous recurrence relation with constant coefficients* is an equation of the form $a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_d a_{n-d}$, where the $d \geqslant 1$ *coefficients* $c_1, \ldots, c_d$ are constants with $c_d \neq 0$. The integer d is called the *order* of the recurrence relation. *Initial conditions* for this relation specify particular values for d of the a_i (typically $a_0, a_1, \ldots, a_{d-1}$). The coefficients $c_1, \ldots, c_d$ define the *characteristic polynomial* $\chi(\lambda) := \lambda^d - c_1 \lambda^{d-1} - c_2 \lambda^{d-2} - \cdots - c_d$ whose d roots play a key role in the solution of a recurrence relation (cf. [12, 13]). To be precise, if $\lambda_1, \lambda_2, \ldots, \lambda_r$ ($1 \leqslant r \leqslant d$) are the distinct (possibly complex) roots of the characteristic polynomial such that λ_i is of multiplicity $m_i := \#\chi(\lambda_i)$ ($i = 1, 2, \ldots, r$), then the general solution of the recurrence relation is given by

$$a_n = \sum_{i=1}^{r} (c_{i1} + c_{i2} n + \cdots + c_{i m_i} n^{m_i - 1}) \lambda_i^n$$

where the c_{ik}'s are (complex) constants. Any real solution is of this form as well (with the imaginary part zero). Moreover, if the coefficients of $\chi(\lambda)$ are real numbers, its non-real roots always come in conjugate pairs; i.e., if $\lambda_j := r_j(\cos(\phi_j) + i\sin(\phi_j))$ is a root of $\chi(\lambda)$, then so is its complex conjugate $\lambda_j^* := r_j(\cos(\phi_j) - i\sin(\phi_j))$. In this case, avoiding the use of complex numbers, the most general real solution can be written as

$$\begin{aligned} a_n = {} & \sum_i (c_{i1} + c_{i2}n + \cdots + c_{im_i}n^{m_i-1})\lambda_i^n \\ & + \sum_j (d_{j1} + d_{j2}n + \cdots + d_{jm_j}n^{m_j-1}) r_j^n \cos(n\phi_j) \\ & + \sum_j (d'_{j1} + d'_{j2}n + \cdots + d'_{jm_j}n^{m_j-1}) r_j^n \sin(n\phi_j) \end{aligned}$$

where the c_{ik}'s, d_{jk}'s and d'_{jk}'s are real constants, the λ_i's the distinct real roots of $\chi(\lambda)$ and the λ_j's, $\lambda_j := r_j(\cos(\phi_j) + i\sin(\phi_j))$, the distinct complex roots (modulo conjugates).

7.3 Polynomially Bounded Matrix Interpretations

In this section, we examine the method of matrix interpretations in the context of complexity analysis of term rewriting, with the focus on its ability to establish polynomial upper bounds on the derivational complexity of compatible TRSs. To this end, we assume all TRSs and signatures to be finite in the sequel. Following [73], we introduce the notion of *growth* of a matrix interpretation (extended from $\mathbb{N}$ to $\mathbb{Q}$ and $\mathbb{R}$) and argue, as in [49], that matrix interpretations of polynomial growth can be completely characterized by the polynomial growth (of the entries) of products of matrices taken from a finite set of matrices.

We start by showing how matrix interpretations can be used to bound the derivational complexity of compatible TRSs. Let $\mathcal{M}$ be a monotone matrix interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$ that is compatible with some TRS $\mathcal{R}$ (thus establishing termination of $\mathcal{R}$ according to the results of Section 5.2), and let α be some variable assignment. Then any rewrite sequence

$$t = t_0 \to_{\mathcal{R}} t_1 \to_{\mathcal{R}} t_2 \to_{\mathcal{R}} t_3 \to_{\mathcal{R}} t_4 \to_{\mathcal{R}} \cdots$$

gives rise to a strictly decreasing sequence of vectors (of numbers in D_0)

$$[\alpha]_{\mathcal{M}}(t) >_\delta [\alpha]_{\mathcal{M}}(t_1) >_\delta [\alpha]_{\mathcal{M}}(t_2) >_\delta [\alpha]_{\mathcal{M}}(t_3) >_\delta [\alpha]_{\mathcal{M}}(t_4) >_\delta \cdots$$

such that, by definition of $>_\delta$ (in Section 5.2), the first components of these vectors form a sequence of non-negative numbers in D_0 that is strictly decreasing with respect to the order $>_{D,\delta}$, each rewrite step causing a decrease of at least δ (here, $\delta = 1$ if $D = \mathbb{N}$ such that $>_{D,\delta} = >_{\mathbb{N}}$). Hence, the first component of the vector $\frac{1}{\delta} \cdot [\alpha]_{\mathcal{M}}(t)$ gives an upper bound on $\mathsf{dh}(t, \to_{\mathcal{R}})$, the maximal length of a rewrite sequence emanating from t. According to Lemma 7.1, one obtains the best approximation for $\alpha = \alpha_0$, that is, by considering the vector $\frac{1}{\delta} \cdot [\alpha_0]_{\mathcal{M}}(t) = \frac{1}{\delta} \cdot [t]$. So if we manage to bound the first component of this vector for all terms t up to a given (but arbitrary) size k, then we have actually established an upper bound on the derivational complexity of $\mathcal{R}$. Moreover, as we are only interested in the asymptotic growth of $\frac{1}{\delta} \cdot [t]$ with respect to the size of t, we may neglect the constant factor $\frac{1}{\delta}$. This motivates the following definition (given in [73] for matrix interpretations over $\mathbb{N}$).

Definition 7.2. For $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$, the *growth function* of a matrix interpretation $\mathcal{M}$ over D is defined as $\mathsf{growth}_{\mathcal{M}}(k) = \max\{[t]_1 \mid t \text{ is a term and } |t| \leqslant k\}$.

By the reasoning given above, we have $\mathsf{dh}(t, \rightarrow_{\mathcal{R}}) \leqslant \frac{1}{\delta} \cdot [t]_1$, and therefore $\mathsf{dc}_{\mathcal{R}}(k) \leqslant \frac{1}{\delta} \cdot \mathsf{growth}_{\mathcal{M}}(k)$, from which we obtain the following result.

Lemma 7.3. *Let $\mathcal{R}$ be a TRS and $\mathcal{M}$ a monotone matrix interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$ that is compatible with $\mathcal{R}$. Then $\mathsf{growth}_{\mathcal{M}}(k) \in O(f(k))$ implies $\mathsf{dc}_{\mathcal{R}}(k) \in O(f(k))$.* ☐

In particular, $\mathsf{growth}_{\mathcal{M}}(k) \in O(k^d)$ implies $\mathsf{dc}_{\mathcal{R}}(k) \in O(k^d)$ for all $d \in \mathbb{N}$. As the growth of $\mathcal{M}$ is at most exponential (in the worst case), the derivational complexity of the TRSs one can handle in this way can at most be exponential. This was shown in [20] for matrix interpretations over $\mathbb{N}$, but the result obviously extends to matrix interpretations over $\mathbb{Q}$ and $\mathbb{R}$ (cf. Remark 7.9 below for a formal account). In order to establish polynomial derivational complexity, the matrices occurring in a (compatible) matrix interpretation must satisfy certain additional properties.

Definition 7.4. For a matrix interpretation $\mathcal{M}$ of dimension n, we denote by $S_{\mathcal{M}}$ the set of matrices occurring in (the interpretation functions of) $\mathcal{M}$. We set $S_{\mathcal{M}} = \{0_n\}$ in the special case when $\mathcal{M}$ contains no matrices. Further, we denote by

$$S_{\mathcal{M}}^k = \{A_1 \cdots A_k \mid A_i \in S_{\mathcal{M}}, 1 \leqslant i \leqslant k\}$$

the set of all products of length k of matrices taken from $S_{\mathcal{M}}$. For $k = 0$, this yields the singleton set $S_{\mathcal{M}}^0 = \{I\}$ containing only the identity matrix. Finally, $S_{\mathcal{M}}^*$ denotes the (matrix) monoid generated by $S_{\mathcal{M}}$, i.e., $S_{\mathcal{M}}^* = \bigcup_{k=0}^{\infty} S_{\mathcal{M}}^k$. We often drop the subscript $\mathcal{M}$ if it is clear from the context.

Remark 7.5. Note that there are no matrices in $\mathcal{M}$ if and only if the underlying (term) signature contains only constant symbols. We avoid having to treat this pathological case explicitly by defining $S_{\mathcal{M}} = \{0_n\}$, which makes $S_{\mathcal{M}}$ non-empty in all cases. Moreover, $S_{\mathcal{M}}$ is always finite as all TRSs and signatures are assumed to be finite.

For $\mathcal{M}$ and $\mathcal{R}$ as in Lemma 7.3, an easy sufficient condition for polynomial boundedness of $\mathsf{growth}_{\mathcal{M}}(k)$ (resp. $\mathsf{dc}_{\mathcal{R}}(k)$) is obtained when the growth of the entries of all matrix products in $S_{\mathcal{M}}^*$ is asymptotically bounded by a polynomial in the length of such products (cf. Lemma 7.8 below). Moreover, as shown in [49], this condition is also necessary in the following sense. If $\mathsf{growth}_{\mathcal{M}}(k)$ is polynomially bounded, then there exists a monotone matrix interpretation $\mathcal{N}$ compatible with $\mathcal{R}$ such that $\mathsf{growth}_{\mathcal{N}}(k) = \mathsf{growth}_{\mathcal{M}}(k)$ and the entries of all matrix products in $S_{\mathcal{N}}^*$ are of polynomial growth (cf. Lemma 7.10 below). The proof given in [49] leverages the connection between matrix interpretations and weighted automata. This is possible since matrix interpretations correspond to a rather restricted form of tree automata, called *path-separated* [39]. The idea is to transform a matrix interpretation into the corresponding automaton, trim this automaton by removing useless states and then transform the resulting automaton back into a (compatible) matrix interpretation. Thus, the interpretation $\mathcal{N}$ can be obtained from $\mathcal{M}$ by simply dropping some rows and columns (the ones whose indices correspond to the useless states) in the matrices and vectors occurring in the interpretation functions of $\mathcal{M}$. So, in some sense, $\mathcal{N}$ can be viewed as a condensed version of $\mathcal{M}$ containing no *junk*.

The above discussion motivates the following definition (cf. [49, Definition 1]).

Definition 7.6. A matrix interpretation $\mathcal{M}$ is *polynomially bounded with degree* $d \in \mathbb{N}$ if $\max\{M_{ij} \mid M \in S^k_{\mathcal{M}}\} \in O(k^d)$ for all $1 \leqslant i, j \leqslant n$, where n is the dimension of $\mathcal{M}$. It is said to be *polynomially bounded* if it is polynomially bounded with degree d for some $d \in \mathbb{N}$.

Example 7.7. Consider the TRS consisting of the single rule $\mathsf{f}(x) \to x$ and the following compatible monotone matrix interpretation $\mathcal{M}$ over $\mathbb{N}$:

$$\mathsf{f}_{\mathcal{M}}(\vec{x}) = F\vec{x} + \vec{f} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

By definition, we have $S_{\mathcal{M}} = \{F\}$ and $S^k_{\mathcal{M}} = \{F^k\}$, and it is easy to see that the growth of the entries in the second column of the matrix F^k is exponential in k. So not all entries of the powers of the matrix F are polynomially bounded. Nevertheless, $[t]_1 < |t|$ for any term t and thus $\mathsf{growth}_{\mathcal{M}}(k)$ is polynomially bounded. However, after dropping the second row and column of F and $\vec{f}$, we obtain a one-dimensional matrix interpretation $\mathcal{N}$ over $\mathbb{N}$ with $\mathsf{f}_{\mathcal{N}}(x) = x + 1$ that is still compatible with the rule $\mathsf{f}(x) \to x$, and that conforms to Definition 7.6 in contrast to the interpretation $\mathcal{M}$.

Lemma 7.8. *Let $\mathcal{M}$ be a matrix interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$. If $\mathcal{M}$ is polynomially bounded with degree d, then $\mathsf{growth}_{\mathcal{M}}(k) \in O(k^{d+1})$.*

Proof. As matrix multiplication distributes over addition, the value of the interpretation of a term t in $\mathcal{M}$ can be written as

$$[t] = \sum_i M_i \vec{v}_i$$

where $M_i \in S^{k_i}_{\mathcal{M}}$ is a matrix product of length $0 \leqslant k_i \leqslant \mathrm{depth}(t) \leqslant |t|$ and $\vec{v}_i = \alpha_0(x) = 0$ for some variable x or $\vec{v}_i = \vec{f}$ for some m-ary function symbol f with $f_{\mathcal{M}}(\vec{x}_1, \dots, \vec{x}_m) = \sum_j F_j \vec{x}_j + \vec{f}$. In the sequel, we call $\vec{f}$ the absolute vector of the interpretation of f. The number of summands in $[t]$ equals the number of subterms of t, which is equal to $|t|$. (All of this can easily be shown by a straightforward induction on t.) Denoting by C the maximum entry of the absolute vectors of the interpretations in $\mathcal{M}$, we obtain

$$[t] = \sum_i M_i \vec{v}_i \leqslant \sum_i M_i \cdot (C, \dots, C)^{\mathrm{T}} = \sum_i C \cdot M_i \cdot (1, \dots, 1)^{\mathrm{T}}$$

Hence, the first component of the i-th summand satisfies

$$C \cdot \sum_{j=1}^{n} (M_i)_{1j} \leqslant C \cdot \sum_{j=1}^{n} \max\{M_{1j} \mid M \in S_{\mathcal{M}}^{k_i}\} \in O(k_i^d)$$

where n is the dimension of $\mathcal{M}$. Since $k_i \leqslant |t|$, it is also in $O(|t|^d)$. As the number of summands in $[t]_1$ is bounded by $|t|$, we obtain $[t]_1 \in O(|t|^{d+1})$, which proves the claim. □

Remark 7.9. An easy modification of the above proof shows that the growth of any matrix interpretation $\mathcal{M}$ (over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$) is at most exponential. To see this, let M denote the componentwise maximum of all matrices occurring in $S_{\mathcal{M}}$. Then any matrix product in $S_{\mathcal{M}}^k$ is dominated by M^k (due to non-negativity of the matrices in $S_{\mathcal{M}}$), and the first component of the i-th summand in the above proof satisfies

$$C \cdot \sum_{j=1}^{n} (M_i)_{1j} \leqslant C \cdot \|M\|_1^{k_i} \tag{7.1}$$

This is easy to see for $k_i = 0$ (where $M_i = I$). For $k_i > 0$, we reason as follows:

$$C \cdot \sum_{j=1}^{n} (M_i)_{1j} \leqslant C \cdot \sum_{j=1}^{n} (M^{k_i})_{1j} \leqslant C \cdot \|M^{k_i}\|_1 \leqslant C \cdot \|M\|_1^{k_i}$$

Observing that $\|M\|_1$ is a constant (not depending on i), we infer from (7.1) and $k_i \leqslant |t|$ that every summand is bounded by $O(c^{|t|})$ for some constant $c > 1$. Together with the fact that the number of summands in $[t]_1$ is bounded by $|t|$, this implies $[t]_1 \in O(|t| \cdot c^{|t|})$, which shows that $\mathsf{growth}_{\mathcal{M}}(k)$ is at most exponential in k.

Lemma 7.10. *Let $\mathcal{R}$ be a TRS and $\mathcal{M}$ a monotone matrix interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$ that is compatible with $\mathcal{R}$. If $\mathsf{growth}_{\mathcal{M}}(k)$ is polynomially bounded (in k), then there exists a compatible monotone matrix interpretation $\mathcal{N}$ over D with $\mathsf{growth}_{\mathcal{N}}(k) = \mathsf{growth}_{\mathcal{M}}(k)$ that is polynomially bounded.*

Proof. By [49, Lemma 29 and Corollary 31]. □

Remark 7.11. The proof of [49, Corollary 31] makes use of [49, Theorem 21], which is intended to be a restatement of [73, Theorem 3.3] and the subsequent remark. Unfortunately, the formulation of [49, Theorem 21] is not quite correct. It says that the growth function of a matrix interpretation is polynomially bounded (with degree $d+1$) if and only if the growth function of the corresponding automaton is polynomially bounded (with degree d). Indeed, this statement is correct if one disregards the degrees of the polynomial bounds. Taking them into account, the "if" direction, which is [73, Theorem 3.3], remains correct, whereas the "only if" direction becomes incorrect. However, for the proof of Lemma 7.10, we only need the fact that the growth function of a matrix interpretation is polynomially bounded if and only if the growth function of the corresponding automaton is polynomially bounded, which is correct.

The above lemmata show that polynomial boundedness of the growth function of a matrix interpretation can be completely characterized by the polynomial growth of the entries of all matrix products in the associated matrix monoid (modulo junk in the interpretation), thus justifying Definition 7.6. Alternatively, one could define polynomially bounded matrix interpretations directly via polynomial boundedness of their growth functions (as in [73]). However, we prefer the characterization of Definition 7.6 because it reduces polynomial boundedness of matrix interpretations to a property of a set of matrices, thus unveiling the underlying cause of polynomial boundedness by casting aside the irrelevant. Most notably, it yields a notion of polynomially bounded matrix interpretations which is independent of the underlying well-founded order they are based on. That is to say, while Definition 7.2 is tailored for the traditional variant of matrix interpretations of Endrullis *et al.* [20] (as introduced in Chapter 5), Definition 7.6 also applies to other kinds of matrix interpretations, like the ones of [17] or the ones introduced in Appendix B (resp. [55]). So it is both a more general and a more natural definition.

As an immediate consequence of Lemmata 7.3 and 7.8, the relationship between polynomially bounded matrix interpretations and the derivational complexity of compatible TRSs is as follows.

Corollary 7.12. *Let $\mathcal{R}$ be a TRS and $\mathcal{M}$ a monotone matrix interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$ that is compatible with $\mathcal{R}$. If $\mathcal{M}$ is polynomially bounded with degree d, then $\mathsf{dc}_{\mathcal{R}}(k) \in O(k^{d+1})$.* □

The next example shows that the increase in the degree is necessary in general.

Example 7.13 (continued from Example 7.7)**.** The matrix interpretation $\mathcal{N}$ of Example 7.7 is obviously polynomially bounded with degree 0 as all (matrix) products in $S_{\mathcal{N}}^*$ are bounded by the constant 1. Hence, by Corollary 7.12, the derivational complexity of the rule $\mathsf{f}(x) \to x$ is at most linear. In fact, it is exactly linear as witnessed by the family of terms $(\mathsf{f}^k(x))_{k \in \mathbb{N}}$.

In the following sections, we employ algebraic methods in order to ensure polynomial boundedness of matrix interpretations. We also show that the converse of Corollary 7.12 (resp. Lemma 7.3 with $f(k)$ set to k^d, $d \in \mathbb{N}$) does not hold. That is, there are TRSs of polynomial derivational complexity which are compatible with a matrix interpretation but not with a polynomially bounded one.

7.4 Spectral Radius

In this section, we use results from linear algebra to derive sufficient conditions for polynomial boundedness of matrix interpretations. In contrast to all other approaches [53, 73], our approach can handle matrix interpretations over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$ (rather than just $\mathbb{N}$). Conceptually, it is based on over-approximating the entries of finite matrix products of the form $A_1 \cdots A_k \in S^k$ by the powers of a single matrix M where $M_{ij} = \max\{A_{ij} \mid A \in S\}$, for all $1 \leqslant i, j \leqslant n$, and

$S \subseteq \mathbb{R}_0^{n\times n}$ is a finite set of matrices associated with some matrix interpretation. Thus, by non-negativity of the matrices in S, we have

$$(A_1 \cdots A_k)_{ij} \leqslant (M^k)_{ij} \tag{7.2}$$

for all $A_1, \ldots, A_k \in S$, $k \in \mathbb{N}$ and $1 \leqslant i, j \leqslant n$, such that polynomial boundedness of the entries of $A_1 \cdots A_k$ follows from polynomial boundedness of the entries of M^k. In [53], for example, this is achieved by restricting the shape of the matrices to upper triangular form.

Lemma 7.14 ([53, Lemma 5])**.** *Let $A \in \mathbb{N}^{n\times n}$ be an upper triangular complexity matrix and $k \in \mathbb{N}$. Then $(A^k)_{ij} \in O(k^{n-1})$ for all $1 \leqslant i, j \leqslant n$.* □

Now if all matrices occurring in a matrix interpretation over $\mathbb{N}$ are upper triangular complexity matrices, then so is the componentwise maximum matrix of these matrices, such that, by Lemma 7.14 together with (7.2), the matrix interpretation is polynomially bounded. From this and from Corollary 7.12 we obtain the following result.

Theorem 7.15 ([53, Theorem 6])**.** *If a TRS $\mathcal{R}$ is compatible with a monotone triangular matrix interpretation over $\mathbb{N}$, then $\mathsf{dc}_{\mathcal{R}}(k) \in O(k^n)$, where n is the dimension of the interpretation.* □

We illustrate the use of this theorem on an example TRS taken from TPDB [72].

Example 7.16. The TRS $\mathcal{R}$ consisting of the rewrite rules[3]

$$\mathsf{a}(\mathsf{b}(\mathsf{a}(x))) \to \mathsf{a}(\mathsf{b}(\mathsf{b}(\mathsf{a}(x)))) \qquad \mathsf{b}(\mathsf{b}(\mathsf{b}(x))) \to \mathsf{b}(\mathsf{b}(x))$$

is compatible with the monotone triangular matrix interpretation

$$\mathsf{a}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \qquad \mathsf{b}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

As the dimension of the interpretation is three, the derivational complexity of $\mathcal{R}$ is (at most) cubic according to Theorem 7.15. It is easy to show that there are no triangular matrix interpretations of dimension one and two compatible with $\mathcal{R}$.

However, Lemma 7.14 only gives a rough estimate of the growth of the entries of the matrix A^k, i.e., the degree of the polynomial bound can be lowered in many cases (e.g. in Example 7.16). To this end, we provide a more concise analysis of the growth of A^k in the remainder of this section, obtaining a replacement for Lemma 7.14, which allows us to tighten the bounds established by Theorem 7.15 (for the same interpretation, cf. Theorem 7.24 and Example 7.25). In particular, our approach applies to matrix interpretations over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$ alike. Moreover, it does not enforce any restriction on the shape of the matrices. Clearly, there are non-triangular matrices that exhibit polynomial growth, but in general

[3] TPDB problem TRS/Zantema_04/z126.xml

non-triangular matrix interpretations do not induce polynomial (but rather exponential) upper bounds on the derivational complexity of compatible TRSs. So in order to be useful in (automated) complexity analysis of term rewriting, a characterization of polynomially bounded matrices is required such that, when searching for a compatible matrix interpretation for a given TRS, it is guaranteed beforehand that the search process only considers such matrices. For this purpose, we leverage the Cayley-Hamilton theorem and the theory of linear homogeneous recurrence relations to completely characterize the growth of the powers of real square matrices (independently of the shape of the matrices). In particular, we show that the key point for polynomial boundedness of such matrices is the nature of their eigenvalues, or, more precisely, their spectral radius.

Lemma 7.17. *Let $A \in \mathbb{R}_0^{n \times n}$ and let $p \in \mathbb{R}[x]$ be a monic polynomial that annihilates A and whose roots have absolute value at most one. Then $\rho(A) \leqslant 1$ if and only if all entries of A^k ($k \in \mathbb{N}$) are asymptotically bounded by a polynomial in k of degree d, where $d = \max_\lambda(0, \#p(\lambda) - 1)$ and λ ranges over all roots of p with absolute value exactly one.*

Proof. First, let us assume that $\rho(A) > 1$, i.e., A has an eigenvalue λ of absolute value strictly greater than one. For any eigenvector x associated to λ, we have $Ax = \lambda x$ and hence $A^k x = \lambda^k x$. Since x is non-zero by definition and $|\lambda| > 1$, there is at least one component of $\lambda^k x$ whose absolute value exhibits exponential growth in k. But this can only be the case if at least one entry of A^k has exponential growth in k as well. Conversely, if $\rho(A) \leqslant 1$, we have to show that the entries of A^k are polynomially bounded in k. To this end, let b denote the degree of the polynomial p. Without loss of generality, p can be written as $p(x) = x^t \cdot q(x)$, $0 \leqslant t \leqslant b$, where t is maximal and q is a monic polynomial of degree $b - t$. By assumption, we have $p(A) = 0$. Clearly, if $t = b$, then $A^k = 0$ for all $k \geqslant b$ and $d = 0$, such that the claim follows trivially. If $t < b$ we rearrange the equation $p(A) = 0$ into the form $A^b = c_1 A^{b-1} + c_2 A^{b-2} + \cdots + c_{b-t} A^t$ with coefficients $c_1, \ldots, c_{b-t}$, readily obtaining a recursive equation for the powers of A, namely, for all $k \geqslant b \in \mathbb{N}$, $A^k = c_1 A^{k-1} + c_2 A^{k-2} + \cdots + c_{b-t} A^{k-(b-t)}$, where $c_{b-t} \neq 0$ due to the maximality of t. Thus, we establish the following recurrence relation

$$A_k = c_1 A_{k-1} + c_2 A_{k-2} + \cdots + c_{b-t} A_{k-(b-t)} \tag{7.3}$$

and note that the sequence $(A_j)_{j \geqslant t}$ where $A_j := A^j$ satisfies it by construction. This is a linear homogeneous recurrence relation with constant coefficients and characteristic polynomial $\chi(\lambda) = q(\lambda)$. Since the coefficients of $\chi(\lambda)$ are real numbers, the non-real roots always come in conjugate pairs; that is, if $\lambda_j := r_j(\cos(\phi_j) + i \sin(\phi_j))$ is a root of $\chi(\lambda)$, then so is its complex conjugate $\lambda_j^* := r_j(\cos(\phi_j) - i \sin(\phi_j))$. Thus, the general solution of (7.3) can be written as

$$
\begin{aligned}
A_k = {} & \sum_i (C_{i1} + C_{i2}k + \cdots + C_{im_i}k^{m_i-1})\lambda_i^k \\
& + \sum_j (D_{j1} + D_{j2}k + \cdots + D_{jm_j}k^{m_j-1}) r_j^k \cos(k\phi_j) \\
& + \sum_j (D'_{j1} + D'_{j2}k + \cdots + D'_{jm_j}k^{m_j-1}) r_j^k \sin(k\phi_j)
\end{aligned}
\tag{7.4}
$$

where the λ_i's are the distinct real roots of $\chi(\lambda)$, each having multiplicity m_i (i.e., $m_i = \#\chi(\lambda_i)$), and the λ_j's, $\lambda_j := r_j(\cos(\phi_j) + i\sin(\phi_j))$, the distinct complex roots (modulo conjugates), each having multiplicity m_j. By assumption, the absolute values of all roots of p are at most one; hence, $|\lambda_i| \leqslant 1$ and $r_j \leqslant 1$ in (7.4), such that the asymptotic growth of the entries of the matrix A^k is polynomial rather than exponential. In particular, the degree d of the polynomial bound is at most $m - 1$, where m is the largest of the multiplicities of the roots with absolute value exactly one. If there are no such roots, then $\lim_{k\to\infty} A^k = 0$, such that $d = 0$. □

According to the Cayley-Hamilton theorem (cf. Section 5.1), the characteristic polynomial χ_A of the matrix A in Lemma 7.17 always annihilates A, and its roots have absolute value at most one if and only if $\rho(A) \leqslant 1$. Thus, we obtain the following corollary.

Corollary 7.18. *Let $A \in \mathbb{R}_0^{n\times n}$. Then $\rho(A) \leqslant 1$ if and only if all entries of A^k ($k \in \mathbb{N}$) are asymptotically bounded by a polynomial in k of degree d, where $d = \max_\lambda(0, \#\chi_A(\lambda) - 1)$ and λ ranges over all eigenvalues of A with absolute value exactly one.* □

Example 7.19. Consider the 4×4 matrix

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

It has one real eigenvalue $\lambda_1 = 1$ of multiplicity two and a pair of complex conjugate eigenvalues $\lambda_2 = \frac{1}{2}(-1 + i\sqrt{3})$ and $\lambda_2^* = \frac{1}{2}(-1 - i\sqrt{3})$ of multiplicity one, all of which have absolute value exactly one. Hence, the spectral radius $\rho(A)$ of A is also one. According to Corollary 7.18, the entries of the matrix A^k are bounded by a linear polynomial in k. The actual bound, however, is even lower since $A^4 = A$, such that the powers of A are trivially bounded by a constant, and we can use the method outlined in the proof of Lemma 7.17 to show this. First, we note that the characteristic polynomial of A is $\chi_A(\lambda) = \lambda^4 - \lambda^3 - \lambda + 1$. Thus, by the Cayley-Hamilton theorem, we obtain the recursive equation $A^k = A^{k-1} + A^{k-3} - A^{k-4}$ for all $k \geqslant 4 \in \mathbb{N}$, the general solution of which can be written as

$$
A^k = (C_0 + C_1 k)\lambda_1^k + D\, r^k \cos(k\phi) + D'\, r^k \sin(k\phi) \tag{7.5}
$$

where $r(\cos(\phi) + i\sin(\phi)) = \lambda_2$, that is, $r = 1$ and $\phi = \frac{2\pi}{3}$. In the next step, the exact values of the four constants C_0, C_1, D and D' can be determined, for example, by letting $k = 4, 5, 6, 7$ in (7.5) and solving the resulting systems of linear equations. In doing so, one learns that C_1 is zero, which means that the linear summand in (7.5) vanishes. Further, we obtain $A^k = C_0 + D\cos(k\phi) + D'\sin(k\phi)$, where

$$C_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \quad D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \quad D' = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ 0 & \frac{\sqrt{3}}{3} & 0 & -\frac{\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & 0 \end{pmatrix}$$

which explains why the powers of A are bounded by a constant. In particular, the periodic nature of the sequence $(A^k)_{k\in\mathbb{N}}$ becomes evident.

The reason why the degree of the polynomial bound established by Corollary 7.18 in the above example is too high is due to the fact that in Lemma 7.17 the characteristic polynomial is not (always) the best possible choice for the annihilating polynomial. Obviously, the set of (monic) polynomials that annihilate a matrix A is infinite. However, from linear algebra we know that this set is generated by a unique monic polynomial of minimum degree that annihilates A, namely, the minimal polynomial $\mathrm{m}_A(x)$ of A. Any other annihilating polynomial is a (polynomial) multiple of $\mathrm{m}_A(x)$. Hence, the minimal polynomial divides the characteristic polynomial $\chi_A(x)$. Moreover, $\mathrm{m}_A(\lambda) = 0$ if and only if λ is an eigenvalue of A, so every root of $\mathrm{m}_A(x)$ is a root of $\chi_A(x)$, and vice versa. However, in case $\mathrm{m}_A(x) \neq \chi_A(x)$, the multiplicity of a root in $\mathrm{m}_A(x)$ may be lower than its multiplicity in $\chi_A(x)$ (cf. [35]). In light of these facts, we conclude that in Lemma 7.17 the minimal polynomial $\mathrm{m}_A(x)$ is the best choice (in theory; in an implementation a constraint solver might find other annihilating polynomials faster).

Corollary 7.20. *Let $A \in \mathbb{R}_0^{n\times n}$. Then $\rho(A) \leqslant 1$ if and only if all entries of A^k ($k \in \mathbb{N}$) are asymptotically bounded by a polynomial in k of degree d, where $d = \max_\lambda(0, \#\mathrm{m}_A(\lambda) - 1)$ and λ ranges over all eigenvalues of A with absolute value exactly one.* □

Indeed, using Corollary 7.20 in Example 7.19 yields the exact degree of the polynomial bound because the eigenvalue $\lambda_1 = 1$ only occurs with multiplicity one in the minimal polynomial.

Based on Corollary 7.20 (resp. Corollary 7.18), one obtains as a direct consequence of Corollary 7.12 (together with (7.2)) the following theorem concerning complexity analysis via matrix interpretations.

Theorem 7.21. *Let $\mathcal{R}$ be a TRS and $\mathcal{M}$ a compatible monotone matrix interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$ of dimension n. Further, let A denote the componentwise maximum of all matrices occurring in $S_\mathcal{M}$. If the spectral radius of A is at most one, then $\mathsf{dc}_\mathcal{R}(k) \in O(k^{d+1})$, where $d = \max_\lambda(0, \#\mathrm{m}_A(\lambda) - 1)$ (resp. $d = \max_\lambda(0, \#\chi_A(\lambda) - 1)$) and λ ranges over all eigenvalues of A with absolute value exactly one.* □

Remark 7.22. Actually, d can be strengthened to $\max_\lambda(0, \#\chi_A(\lambda)) - 1$ and $\max_\lambda(0, \#\mathrm{m}_A(\lambda)) - 1$, respectively, because the pathological case $\rho(A) < 1$, where $\lim_{k\to\infty} A^k = 0$, implies $\mathsf{dc}_{\mathcal{R}}(k) \in O(1)$. As $\lim_{k\to\infty} A^k = 0$ cannot happen if $A_{11} \geqslant 1$, this means that there are no matrices in $\mathcal{M}$ (i.e., $S_{\mathcal{M}} = \{0\}$), which can only be the case if the underlying signature contains only constant symbols. But then the inherent simple shape of the rules of $\mathcal{R}$ results in a derivational complexity that is trivially bounded by a constant (as all TRSs are assumed to be finite).

Next we illustrate the use of Theorem 7.21 on an example.

Example 7.23. Consider the TRS $\mathcal{R}$ consisting of the following rewrite rules:[4]

$$\begin{aligned} \mathsf{h}(x, \mathsf{c}(y, z)) &\to \mathsf{h}(\mathsf{c}(\mathsf{s}(y), x), z) \\ \mathsf{h}(\mathsf{c}(\mathsf{s}(x), \mathsf{c}(\mathsf{s}(0), y)), z) &\to \mathsf{h}(y, \mathsf{c}(\mathsf{s}(0), \mathsf{c}(x, z))) \end{aligned}$$

This system is compatible with the monotone matrix interpretation

$$\mathsf{0}_{\mathcal{M}} = \begin{pmatrix} 0 \\ 3 \\ 1 \\ 0 \end{pmatrix} \qquad \mathsf{c}_{\mathcal{M}}(\vec{x}, \vec{y}) = \begin{pmatrix} 1&1&1&0 \\ 0&0&0&0 \\ 0&0&1&0 \\ 0&1&0&0 \end{pmatrix} \vec{x} + \begin{pmatrix} 1&0&0&0 \\ 0&0&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix} \vec{y} + \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

$$\mathsf{s}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1&0&1&0 \\ 0&0&0&0 \\ 0&1&0&0 \\ 0&0&0&0 \end{pmatrix} \vec{x} \qquad \mathsf{h}_{\mathcal{M}}(\vec{x}, \vec{y}) = \begin{pmatrix} 1&0&1&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix} \vec{x} + \begin{pmatrix} 1&0&0&1 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix} \vec{y} + \begin{pmatrix} 0 \\ 2 \\ 1 \\ 2 \end{pmatrix}$$

The maximum matrix A has the characteristic polynomial $\chi_A(x) = x(x-1)^3$ and the minimal polynomial $\mathrm{m}_A(x) = x(x-1)^2$. Thus, the upper bound for the derivational complexity of $\mathcal{R}$ derived from Theorem 7.21 is quadratic since $\#\mathrm{m}_A(1) = 2$ in the minimal polynomial (the bound is worse when using the characteristic polynomial as $\#\chi_A(1) = 3$).

Next we specialize Theorem 7.21 to triangular matrix interpretations. In such interpretations all matrices are upper triangular complexity matrices whose diagonal entries are restricted to the closed interval $[0, 1]$. Hence, this is also true for the componentwise maximum matrix A. Since the diagonal entries of a triangular matrix give the multiset of its eigenvalues, the matrix A is therefore guaranteed to have spectral radius at most one.

Theorem 7.24. *If a TRS $\mathcal{R}$ is compatible with a monotone triangular matrix interpretation $\mathcal{M}$ over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$ of dimension n, then $\mathsf{dc}_{\mathcal{R}}(k) \in O(k^d)$, where d is the number of ones in the diagonal of the componentwise maximum of all matrices occurring in $S_{\mathcal{M}}$.*[5]

Proof. By Theorem 7.21 and Remark 7.22 (note that $d = \#\chi_A(1)$ if $d > 0$). $\square$

[4] TPDB problem TRS/Endrullis_06/direct.xml

[5] Independently, in [73, Proposition 7.6] the same result has been established for $\mathbb{N}$ using techniques from automata theory.

Note that the bound established by Theorem 7.24 is at least as tight as the one of Theorem 7.15 since $d \leqslant n$. In practice, however, one can often infer tighter bounds from the same interpretation, just by inspecting the diagonal structure of its matrices. This makes Theorem 7.24 very easy to implement (as there is no explicit reference to the characteristic and/or minimal polynomial).

Example 7.25 (continued from Example 7.16)**.** The diagonal of the componentwise maximum of the two matrices given in Example 7.16 has the shape $(1, 0, 0)$. Hence, $\mathcal{R}$ has (at most) linear derivational complexity by Theorem 7.24, whereas the bound established by Theorem 7.15 is cubic. Incidentally, the bound inferred from Theorem 7.24 is even optimal as it is easy to see that the derivational complexity of $\mathcal{R}$ is at least linear.

In the example below, we demonstrate why triangular matrices may fail. Similar (but larger) systems are contained in TPDB, e.g., TRS/Cime_04/dpqs.xml.

Example 7.26. Consider the TRS $\mathcal{R}$ consisting of the rules

$$\mathsf{f}(\mathsf{f}(x)) \to \mathsf{f}(\mathsf{g}(\mathsf{f}(x))) \qquad \mathsf{g}(\mathsf{g}(x)) \to x$$

which is compatible with the following monotone matrix interpretation over $\mathbb{N}$:

$$\mathsf{f}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \quad \mathsf{g}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

The eigenvalues of the componentwise maximum matrix are -1, 1 and 1, and its minimal polynomial coincides with its characteristic polynomial. Hence, Theorem 7.21 deduces a quadratic upper bound on the derivational complexity of $\mathcal{R}$. However, there cannot exist a triangular matrix interpretation compatible with $\mathcal{R}$.

Lemma 7.27. *There is no triangular matrix interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$ that is compatible with the TRS $\mathcal{R}$ of Example 7.26.*

Proof. Assume to the contrary that $\mathcal{M}$ is a compatible triangular matrix interpretation (not necessarily a monotone one) of dimension n such that $\mathsf{g}_{\mathcal{M}}(\vec{x}) = G\vec{x} + \vec{g}$. Then compatibility with the second rule demands $G^2 \geqslant I$. As G is a triangular matrix, we must have $(G^2)_{ii} = (G_{ii})^2 \geqslant 1$ for all $1 \leqslant i \leqslant n$, hence, $G \geqslant I$. But then $\mathsf{g}_{\mathcal{M}}(\vec{x}) \geqslant \vec{x}$, such that $\mathcal{M}$ cannot be compatible with the first rule. □

The next example shows the benefit of matrix interpretations over $\mathbb{Q}$ ($\mathbb{R}$).

Example 7.28. Consider the TRS $\mathcal{R}$ consisting of the following rewrite rules:[6]

$$\begin{aligned} \mathsf{t}(\mathsf{o}(x)) &\to \mathsf{m}(\mathsf{a}(x)) & \mathsf{t}(\mathsf{e}(x)) &\to \mathsf{n}(\mathsf{s}(x)) \\ \mathsf{a}(\mathsf{l}(x)) &\to \mathsf{a}(\mathsf{t}(x)) & \mathsf{o}(\mathsf{m}(\mathsf{a}(x))) &\to \mathsf{t}(\mathsf{e}(\mathsf{n}(x))) \\ \mathsf{s}(\mathsf{a}(x)) &\to \mathsf{l}(\mathsf{a}(\mathsf{t}(\mathsf{o}(\mathsf{m}(\mathsf{a}(\mathsf{t}(\mathsf{e}(x)))))))) & \mathsf{n}(\mathsf{s}(x)) &\to \mathsf{a}(\mathsf{l}(\mathsf{a}(\mathsf{t}(x)))) \end{aligned}$$

[6]TPDB problem TRS/Secret_05_SRS/matchbox2.xml

The monotone triangular matrix interpretation $\mathcal{M}$ over $\mathbb{Q}$ ($\mathbb{R}$) given below is compatible with $\mathcal{R}$ (for $\delta \leqslant \frac{1}{2}$):

$$\mathsf{t}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \mathsf{o}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} \frac{3}{2} \\ \frac{7}{2} \\ 0 \end{pmatrix}$$

$$\mathsf{m}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} \frac{1}{2} \\ 0 \\ 0 \end{pmatrix} \qquad \mathsf{a}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$$

$$\mathsf{e}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} \frac{3}{2} \\ 0 \\ 3 \end{pmatrix} \qquad \mathsf{n}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\mathsf{s}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 \\ 3 \\ 3 \end{pmatrix} \qquad \mathsf{l}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ 0 \\ \frac{1}{2} \end{pmatrix}$$

The diagonal of the componentwise maximum matrix has the shape $(1, \frac{1}{2}, 0)$. Hence, by Theorem 7.24, the derivational complexity of $\mathcal{R}$ is at most linear. Our implementation (which is described in Section 7.6) could find a compatible triangular matrix interpretation of the same dimension over $\mathbb{N}$ establishing a quadratic but not a linear bound.

With the help of the techniques introduced in Chapter 6 one can even create TRSs of polynomial derivational complexity where matrix interpretations over $\mathbb{N}$ fail (and matrix interpretations over $\mathbb{Q}$ succeed).

Example 7.29. Consider the TRS $\mathcal{R}$ consisting of the rules (6.1) – (6.4) from Section 6.1 and the following rules:

$$((\mathsf{g}(\mathsf{g}(x)) + \mathsf{g}(\mathsf{g}(x))) + x) + \mathsf{a} \to (\mathsf{g}(x) + \mathsf{g}(x)) + \mathsf{g}(x) \tag{7.6}$$
$$((\mathsf{g}(x) + \mathsf{g}(x)) + \mathsf{g}(x)) + \mathsf{a} \to (\mathsf{g}(\mathsf{g}(x)) + \mathsf{g}(\mathsf{g}(x))) + x \tag{7.7}$$
$$\mathsf{f}(\mathsf{a}) \to \mathsf{f}(\mathsf{g}(\mathsf{a})) \tag{7.8}$$

This system is compatible with the following monotone interpretation over $\mathbb{Q}$:

$$\delta = \tfrac{1}{2} \quad \mathsf{a}_{\mathcal{M}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \mathsf{b}_{\mathcal{M}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \mathsf{g}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \vec{x}$$
$$\mathsf{f}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad +_{\mathcal{M}}(\vec{x}, \vec{y}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \vec{y}$$

Being a triangular interpretation whose maximum matrix has a diagonal of the shape $(1, 1)$, we infer from Theorem 7.24 that the derivational complexity of $\mathcal{R}$ is at most quadratic. However, there is no compatible monotone matrix interpretation $\mathcal{M}$ over $\mathbb{N}$ (of any dimension). This can be seen as follows. To begin with, we know from Lemma 6.4 that the function symbol $+$ must be interpreted by $+_{\mathcal{M}}(\vec{x},\vec{y}) = \vec{x}+\vec{y}+\vec{v}$. Then, assuming $\mathsf{g}_{\mathcal{M}}(\vec{x}) = G\vec{x}+\vec{g}$ without loss of generality, the compatibility constraint associated with (7.6) implies $2G^2 + I \geqslant 3G$. We also have $3G \geqslant 2G^2 + I$ due to (7.7); hence, $2G^2 + I = 3G$. This means that the monic polynomial $p(x) = x^2 - \frac{3}{2}x + \frac{1}{2} = (x-1)(x-\frac{1}{2})$ annihilates the matrix G. So the minimal polynomial $\mathrm{m}_G(x)$ divides $p(x)$; there are three possibilities: $\mathrm{m}_G(x) = x - \frac{1}{2}$, $\mathrm{m}_G(x) = x - 1$ and $\mathrm{m}_G(x) = p(x)$. The first case where $G = \frac{1}{2}I$ is not permissible since $G_{11} \geqslant 1$ is required for monotonicity, whereas the second case where $G = I$ violates compatibility with (7.8) (since $\mathsf{g}_{\mathcal{M}}(\vec{x}) \geqslant \vec{x}$ then). So we conclude that $\mathrm{m}_G(x) = p(x)$. As $\mathrm{m}_G(x)$ divides the characteristic polynomial $\chi_G(x)$ and $\mathrm{m}_G(x) = 0$ if and only if x is an eigenvalue of G, $\mathrm{m}_G(x)$ and $\chi_G(x)$ have the same irreducible factors, but with possibly different multiplicities (cf. e.g. [35]); hence, $\chi_G(x) = (x-1)^r(x-\frac{1}{2})^s$ with $r, s \geqslant 1$. Now if G were an integer matrix, then all coefficients of $\chi_G(x)$ would be integers as well (cf. Section 5.1). However, the constant coefficient of $\chi_G(x)$ above is not an integer (for any value of r and s). Therefore, G cannot be an integer matrix, and hence $\mathcal{M}$ cannot be a matrix interpretation over $\mathbb{N}$.

We conclude this section with an example that demonstrates the conceptual limitations arising from the over-approximation of matrix products by the powers of the corresponding maximum matrix (cf. (7.2)). For this purpose, we construct a TRS that is compatible with a polynomially bounded matrix interpretation but not with one whose maximum matrix is polynomially bounded. The idea is as follows. If a TRS $\mathcal{R}$ admits a polynomially bounded matrix interpretation but not a triangular one, then the weighted graph $G(A)$ associated with its maximum matrix A must have a cycle of length $\ell \geqslant 2$. Denoting by C the adjacency matrix of *one* such cycle and assuming, without loss of generality, that all its vertices are distinct, we observe that all diagonal entries of C are zero, and that C^ℓ is a (diagonal) matrix where the diagonal entries corresponding to the vertices of the cycle are positive. By adding the rule $\mathsf{b}(x) \to x$ to $\mathcal{R}$, where b is a fresh symbol not occurring in $\mathcal{R}$, we obtain a TRS that is still compatible with a polynomially bounded matrix interpretation $\mathcal{M}$; e.g. by adding $\mathsf{b}_{\mathcal{M}}(\vec{x}) = \vec{x} + (1,0,\dots,0)^{\mathrm{T}}$ to the interpretation for $\mathcal{R}$. In particular, the matrix B associated with the interpretation of the symbol b must satisfy $B \geqslant I$. Thus, the maximum matrix M of $\mathcal{M}$ must satisfy $M \geqslant \max(I, C) = I + C$ such that $M^\ell \geqslant (I+C)^\ell \geqslant I + C^\ell$. As a consequence, $(M^\ell)_{jj} > 1$ for some index j, which shows that M^k exhibits exponential growth in k. In the concrete example below, we take the TRS $\mathcal{R}$ of Example 7.26 and denote by $\mathcal{R}_{\mathrm{jsr}}$ the union of $\mathcal{R}$ and $\{\mathsf{b}(x) \to x\}$.

Lemma 7.30. *There is a monotone matrix interpretation compatible with $\mathcal{R}_{\mathrm{jsr}}$ that is polynomially bounded (in the sense of Definition 7.6), but there is no compatible monotone matrix interpretation where all entries in the componentwise maximum matrix are polynomially bounded.*

Proof. For the first claim, we extend the interpretation given in Example 7.26, which is polynomially bounded, by $\mathsf{b}_{\mathcal{M}}(\vec{x}) = \vec{x} + (1,0,0)^{\mathrm{T}}$, thus obtaining a compatible monotone matrix interpretation $\mathcal{M}$ over $\mathbb{N}$ that is still polynomially bounded as the matrix associated with the interpretation of b is the identity matrix (cf. also Theorem 7.39 below). For the second claim, let $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$ and assume that $\mathcal{M}$ is a compatible matrix interpretation over D (of dimension n) with $\mathsf{b}_{\mathcal{M}}(\vec{x}) = B\vec{x} + \vec{b}$, $\mathsf{f}_{\mathcal{M}}(\vec{x}) = F\vec{x} + \vec{f}$ and $\mathsf{g}_{\mathcal{M}}(\vec{x}) = G\vec{x} + \vec{g}$. The proof of Lemma 7.27 shows that G must satisfy $G^2 \geqslant I$ and $G \ngeqslant I$, i.e., there exists an index l such that $G_{ll} < 1$. So we must have $(G^2)_{ll} = \sum_j G_{lj}G_{jl} \geqslant 1$ and consequently $\sum_{j \neq l} G_{lj}G_{jl} > 0$. As the rule $\mathsf{b}(x) \to x$ demands $B \geqslant I$, we conclude that the maximum matrix M must satisfy $M \geqslant \max(I, G)$. From this and $\sum_{j \neq l} G_{lj}G_{jl} > 0$ we obtain $(M^2)_{ll} > 1$, which shows that $(M^k)_{ll}$ exhibits exponential growth in k (as M is non-negative). □

7.5 Joint Spectral Radius

Instead of using a single maximum matrix to over-approximate the growth of finite matrix products taken from a set of matrices S, in this section we provide a concise analysis using joint spectral radius theory [36, 37]. In particular, we shall see that the joint spectral radius of S completely characterizes when the growth (of the entries) of all such products, i.e., the products in the associated matrix monoid S^*, is polynomially bounded, just like the spectral radius of a single matrix characterizes polynomial growth of the powers of this matrix.

Definition 7.31. Let S be a finite set of matrices from $\mathbb{R}^{n \times n}$, and let $\|\cdot\|$ denote a matrix norm. The *growth function* associated with S is defined as $\mathsf{growth}_S(k, \|\cdot\|) = \max\{\, \|A_1 \cdots A_k\| \mid A_i \in S, 1 \leqslant i \leqslant k \,\}$.

Using the norm $\|\cdot\|_1$ given by the sum of the absolute values of all matrix entries, we observe that a matrix interpretation $\mathcal{M}$ is polynomially bounded (with degree d) if and only if $\mathsf{growth}_{S_{\mathcal{M}}}(k, \|\cdot\|_1)$ is polynomially bounded (with degree d) in k. (The formal proof of this fact is given in Lemma 7.37 below.) As all matrix norms are equivalent [35], i.e., given any two matrix norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$, there exist positive constants c_1 and c_2 such that $c_1\|A\|_\alpha \leqslant \|A\|_\beta \leqslant c_2\|A\|_\alpha$ for any matrix A, we conclude that $\mathsf{growth}_{S_{\mathcal{M}}}(k, \|\cdot\|_1)$ and $\mathsf{growth}_{S_{\mathcal{M}}}(k, \|\cdot\|)$ are asymptotically equivalent for any matrix norm $\|\cdot\|$. The asymptotic behaviour can be characterized by the joint spectral radius of $S_{\mathcal{M}}$.

Definition 7.32. Let $S \subseteq \mathbb{R}^{n \times n}$ be finite, and let $\|\cdot\|$ denote a matrix norm. The *joint spectral radius* $\rho(S)$ of S is defined by the limit

$$\rho(S) = \lim_{k \to \infty} \max\{\, \|A_1 \cdots A_k\|^{1/k} \mid A_i \in S, 1 \leqslant i \leqslant k \,\}$$

It is well-known that this limit always exists and that it does not depend on the chosen norm, which follows from the equivalence of all matrix norms. If $S = \{A\}$ is a singleton set, the joint spectral radius $\rho(S)$ of S and the spectral radius $\rho(A)$ of A coincide:

$$\rho(S) = \lim_{k \to \infty} \|A^k\|^{1/k} = \max\{\, |\lambda| \mid \lambda \text{ is an eigenvalue of } A \,\} = \rho(A)$$

Since Definition 7.32 is independent of the actual norm, and since we are only interested in the asymptotic behaviour of $\mathsf{growth}_S(k, \|\cdot\|)$, from now on we simply write $\mathsf{growth}_S(k)$. The following theorem provides a characterization of polynomial boundedness of $\mathsf{growth}_S(k)$ by the joint spectral radius of S. It is essentially [7, Theorem 1.2] stated for real rather than complex matrices.

Theorem 7.33. *For a finite set of matrices $S \subseteq \mathbb{R}^{n\times n}$, $\mathsf{growth}_S(k) \in O(k^d)$ for some $d \in \mathbb{N}$ if and only if $\rho(S) \leqslant 1$. In particular, $d \leqslant n-1$.* ☐

Hence, polynomial boundedness of $\mathsf{growth}_S(k)$ is decidable if $\rho(S) \leqslant 1$ is decidable. But it is well-known that the latter is undecidable in general, even if S consists of finitely many non-negative rational matrices (cf. [36, Theorem 2.6]). However, in case S is a finite set of non-negative integer matrices, then $\rho(S) \leqslant 1$ is decidable. In particular, there exists a polynomial-time algorithm that decides it (cf. [36, Theorem 3.1]). This algorithm is based on the following lemma.

Lemma 7.34 ([36, Lemma 3.3])**.** *Let $S \subseteq \mathbb{R}_0^{n\times n}$ be a finite set of non-negative real square matrices. Then there is a product $A \in S^*$ such that $A_{ii} > 1$ for some index $i \in \{1, \dots, n\}$ if and only if $\rho(S) > 1$.* ☐

According to [36], for a finite set $S \subseteq \mathbb{N}^{n\times n}$, the existence of such a product can be characterized in terms of the graphs $G(S)$ and $G^2(S)$ one can associate with S. More precisely, there is a product $A \in S^*$ with $A_{ii} > 1$ if and only if

1. there is a cycle in $G(S)$ containing at least one edge of weight $w > 1$, or
2. there is a cycle in $G^2(S)$ containing at least one vertex (i,i) and at least one vertex (p,q) with $p \neq q$.

Hence, we have $\rho(S) \leqslant 1$ if and only if neither of the two conditions holds, which can be checked in polynomial time according to [36]. Furthermore, as already mentioned in [36, Chapter 3], this graph-theoretic characterization does not only hold for non-negative integer matrices, but for any finite set of matrices where all matrix entries are either zero or at least one (because then all paths in $G(S)$ have weight at least one).

Theorem 7.35. *For a finite set of matrices $S \subseteq \mathbb{R}_0^{n\times n}$ where all matrix entries are either zero or at least one, $\rho(S) \leqslant 1$ is decidable in polynomial time.* ☐

So, in the situation of Theorem 7.35, polynomial boundedness of $\mathsf{growth}_S(k)$ is decidable in polynomial time. In addition, the exact degree of growth can be computed in polynomial time (cf. [36, Theorem 3.3 and Proposition 3.3]).

Theorem 7.36. *Let $S \subseteq \mathbb{R}_0^{n\times n}$ be a finite set of matrices such that $\rho(S) \leqslant 1$ and all matrix entries are either zero or at least one, and let $d \geqslant 0$ be the largest integer possessing the following property: there exist d different pairs of indices (i_1, j_1), …, (i_d, j_d) such that for every pair (i_s, j_s) the indices i_s, j_s are different and there is a product $A \in S^*$ for which $A_{i_s i_s}, A_{i_s j_s}, A_{j_s j_s} \geqslant 1$, and for each $1 \leqslant s \leqslant d-1$, there exists $B \in S^*$ with $B_{j_s i_{s+1}} \geqslant 1$. Then*

$$\mathsf{growth}_S(k) \in \begin{cases} \Theta(k^d) & \text{if } d \geqslant 1 \\ O(k^d) & \text{if } d = 0 \end{cases}$$

Moreover, the growth rate *d is computable in polynomial time and $d \leqslant n-1$.*

Proof. The case $\rho(S) < 1$ is a consequence of [36, Lemma 3.2], whereas the case $\rho(S) = 1$ follows from [36, Theorem 3.3 and Proposition 3.3]. □

Next we elaborate on the ramifications of joint spectral radius theory on complexity analysis of TRSs via polynomially bounded matrix interpretations. To begin with, we show that the characterization of polynomially bounded matrix interpretations given in Definition 7.6 can be rephrased as follows.

Lemma 7.37. *For any matrix interpretation $\mathcal{M}$ and matrix norm $\|\cdot\|$,*

1. *$\mathcal{M}$ is polynomially bounded if and only if $\rho(S_{\mathcal{M}}) \leqslant 1$, and*
2. *$\mathcal{M}$ is polynomially bounded with degree d if and only if $\mathsf{growth}_{S_{\mathcal{M}}}(k, \|\cdot\|)$ is polynomially bounded with degree d (in k).*

Proof. By Theorem 7.33, the condition $\rho(S_{\mathcal{M}}) \leqslant 1$ is equivalent to polynomial boundedness of $\mathsf{growth}_{S_{\mathcal{M}}}(k, \|\cdot\|)$. So it remains to prove the second claim. As all matrix norms are equivalent (see above), it is sufficient to prove this claim for a single norm of our choice. Choosing the norm $\|\cdot\|_1$ amounts to showing the equivalence

$$\forall\, i, j \;\; \max\{\, M_{ij} \mid M \in S^k_{\mathcal{M}} \,\} \in O(k^d) \iff \max\{\, \|M\|_1 \mid M \in S^k_{\mathcal{M}} \,\} \in O(k^d)$$

The "if" direction follows directly from the fact that $\|M\|_1 \geqslant M_{ij}$ for all indices i and j, which is obvious from the definition of $\|\cdot\|_1$, whereas the "only if" direction is due to the fact that $\sum_{i,j} \max\{\, M_{ij} \mid M \in S^k_{\mathcal{M}} \,\} \geqslant \max\{\, \|M\|_1 \mid M \in S^k_{\mathcal{M}} \,\}$ for all $k \in \mathbb{N}$. By the equivalence of all matrix norms, $\max\{\, \|M\|_1 \mid M \in S^k_{\mathcal{M}} \,\}$ and $\max\{\, \|M\| \mid M \in S^k_{\mathcal{M}} \,\}$ are asymptotically equivalent. □

The relationship between polynomially bounded matrix interpretations and the derivational complexity of compatible TRSs expressed in Corollary 7.12 yields the following result.

Theorem 7.38. *Let $\mathcal{R}$ be a TRS and $\mathcal{M}$ a monotone matrix interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$ of dimension n that is compatible with $\mathcal{R}$. If $\rho(S_{\mathcal{M}}) \leqslant 1$, then $\mathsf{dc}_{\mathcal{R}}(k) \in O(k^n)$.*

Proof. By Lemma 7.37, Theorem 7.33 and Corollary 7.12. □

As this theorem assumes the worst-case growth rate for $\mathsf{growth}_{S_{\mathcal{M}}}(k)$, the inferred degree of the polynomial bound may generally be too high (and unnecessarily so). Yet with the help of Theorem 7.36, from which we obtain the exact growth rate, Theorem 7.38 can be strengthened (in this respect), at the expense of having to restrict the set of permissible matrices.

Theorem 7.39. *Let $\mathcal{R}$ be a TRS and $\mathcal{M}$ a compatible monotone matrix interpretation over $D \in \{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$ where all matrix entries are either zero or at least one. If $\rho(S_{\mathcal{M}}) \leqslant 1$, then $\mathsf{dc}_{\mathcal{R}}(k) \in O(k^{d+1})$, where d refers to the growth rate obtained from Theorem 7.36.*

Proof. By Lemma 7.37, Theorem 7.36 and Corollary 7.12. □

The following example shows that the joint spectral radius theory approach can handle TRSs where the spectral radius approach of Section 7.4 fails. Besides, it demonstrates the ability of the former to establish tighter complexity bounds for systems where both approaches succeed.

Example 7.40. According to Lemma 7.30, the TRS $\mathcal{R}_{\mathrm{jsr}}$ whose derivational complexity is polynomial cannot be handled by the spectral radius approach. However, as there exists a compatible monotone matrix interpretation (over $\mathbb{N}$) that is polynomially bounded, Theorems 7.38 and 7.39 must apply. Indeed, the interpretation $\mathcal{M}$ given in the proof of Lemma 7.30 and Example 7.26 consisting of the matrices

$$S_{\mathcal{M}} = \left\{ \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$$

satisfies $\rho(S_{\mathcal{M}}) \leqslant 1$. This can be tested, for example, using the procedure described after Lemma 7.34. Taking a closer look, an enumeration of all products of matrices in $S_{\mathcal{M}}$ up to length three yields $S_{\mathcal{M}}^2 = S_{\mathcal{M}}^3$. Therefore, the monoid $S_{\mathcal{M}}^*$ generated by $S_{\mathcal{M}}$ is finite, hence bounded by a constant. So the derivational complexity of $\mathcal{R}_{\mathrm{jsr}}$ is (at most) linear according to Theorem 7.39. This also shows that the derivational complexity of the TRS $\mathcal{R} \subseteq \mathcal{R}_{\mathrm{jsr}}$ of Example 7.26 is (at most) linear, too, thus improving upon the (quadratic) bound established by Theorem 7.21 in Example 7.26.

7.6 Implementation Issues

Next we report on how the spectral radius approach of Section 7.4 can be implemented. For information on how to implement the joint spectral radius approach of Section 7.5, we refer the reader to [49].

In Theorem 7.21, we consider some TRS together with a compatible matrix interpretation and demand that the componentwise maximum matrix A has a spectral radius of at most one. So we have to make sure that the absolute values of all its eigenvalues (real and complex ones) are at most one. However, since A is a non-negative real square matrix, we only have to ensure this condition for all (non-negative) real eigenvalues of A. This follows directly from the Perron-Frobenius theorem ([68], weak form), which states that the spectral radius of a non-negative real square matrix is an eigenvalue of the matrix; i.e., there exists a non-negative real eigenvalue that dominates in absolute value all eigenvalues.

As far as automation is concerned, the main problem that has to be dealt with is the following. Given some square matrix A with unknown entries, all of which are supposed to be non-negative real (or integer) numbers, we need a set of constraints, expressed in terms of the unknown entries, that enforce $\rho(A) \leqslant 1$; e.g., for which non-negative values of a, b, c and d does the matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

have a spectral radius of at most one? In the sequel, we present three different approaches.

(A) The first approach is based on the explicit calculation of the eigenvalues of A, i.e., the explicit calculation of the roots of the characteristic polynomial $\chi_A(\lambda)$. For the 2-dimensional case, we have $\chi_A(\lambda) = \lambda^2 - (a+d)\lambda + ad - bc$. So, by the quadratic formula (cf. Section 3.1), we obtain the roots

$$\lambda_{1,2} = \frac{a+d}{2} \pm \frac{\sqrt{(a-d)^2 + 4bc}}{2}$$

both of which must be real because all matrix entries are non-negative. In particular, λ_1 ($\geqslant \lambda_2$) is non-negative, such that it suffices to require $\lambda_1 \leqslant 1$ according to the Perron-Frobenius theorem (since $\lambda_1 \geqslant |\lambda_2|$). Simplifying this condition as much as possible, we infer that the spectral radius of A is at most one if and only if $a+d \leqslant 2$ and $a+d \leqslant ad-bc+1$. This explicit approach also applies to matrices of dimension three and four as there exist formulas for the solution of arbitrary cubic and quartic polynomial equations with symbolic coefficients (though the respective calculations are tedious). However, for equations of degree five or higher, there are no formulas that express the solutions of such equations in terms of their coefficients using only the four basic arithmetic operations and radicals (n-th roots, for some integer n).

(B) Next we present an alternative and simpler approach for 3-dimensional matrices. To this end, let

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

be a non-negative real square matrix with characteristic polynomial $\chi_A(\lambda)$

$$\lambda^3 - (a+e+i)\lambda^2 + (ei - fh + ai - cg + ae - bd)\lambda - (aei + bfg + cdh - ceg - bdi - afh)$$

which we abbreviate by $\lambda^3 + p\lambda^2 + q\lambda + r$. By the Perron-Frobenius theorem, it suffices to constrain the real roots of $\chi_A(\lambda)$ to the closed interval $[-1, 1]$. To achieve this, we make use of the well-known fact that the cubic polynomial $\chi_A(\lambda)$ either has only one real root (and two complex conjugate roots) if its *discriminant* $D := p^2q^2 - 4q^3 - 4p^3r - 27r^2 + 18pqr$ is negative or three (not necessarily distinct) real roots if $D \geqslant 0$. Visualizing the geometric shape of $\chi_A(\lambda)$, it is not hard to see that in the latter case all three roots are in $[-1, 1]$ if and only if $\chi_A(-1) \leqslant 0$, $\chi_A(1) \geqslant 0$ and $\chi'_A(\lambda) \geqslant 0$ for all $\lambda \in \mathbb{R}$ with $|\lambda| \geqslant 1$ (here, χ'_A denotes the first derivative of χ_A). Thus, we conclude that the matrix A has a spectral radius of at most one if and only if

$$\begin{aligned} &(D < 0 \wedge \chi_A(-1) \leqslant 0 \wedge \chi_A(1) \geqslant 0) \vee \\ &(\chi_A(-1) \leqslant 0 \wedge \chi_A(1) \geqslant 0 \wedge \chi'_A(\lambda) \geqslant 0 \text{ for all } |\lambda| \geqslant 1) \end{aligned}$$

These are polynomial constraints in the entries of A. In particular, the constraint $\chi'_A(\lambda) = 3\lambda^2 + 2p\lambda + q \geqslant 0$ for all $|\lambda| \geqslant 1$ can be shown to be equivalent to

$$(p^2 - 3q \leqslant 0) \vee (-3 \leqslant p \leqslant 3 \wedge -(q+3) \leqslant 2p \leqslant q+3)$$

by means of the quadratic formula. Here, the term $p^2 - 3q$ is essentially the discriminant of $\chi'_A(\lambda)$; if it is negative, then $\chi'_A(\lambda)$ has no real root, such that the constraint trivially holds, otherwise it has two real roots λ_1 and λ_2. In case $\lambda_1 = \lambda_2$, the constraint also holds because then $\chi'_A(\lambda) = 3 \cdot (\lambda - \lambda_1)^2$. Finally, if $\lambda_1 \neq \lambda_2$, then both must necessarily lie in the closed interval $[-1, 1]$ for the constraint to hold, which is ensured by the second disjunct in the above formula.

(C) Last but not least, with reference to Lemma 7.17 (resp. Corollaries 7.18 and 7.20), we present a generic method that works for matrices with unknown entries of any dimension. Let A be an n-dimensional square matrix whose entries are supposed to be real numbers (not necessarily non-negative for the moment). Then its characteristic polynomial is a monic polynomial of degree n, which can be written as $\chi_A(\lambda) = \lambda^n + \sum_{i=0}^{n-1} c_i \lambda^i$, where the coefficients c_i, $0 \leqslant i \leqslant n-1$, are polynomial expressions in the entries of A. Since all coefficients are supposed to be real numbers, $\chi_A(\lambda)$ can always be factored as

$$\chi_A(\lambda) = (\lambda - r)^b \cdot \prod_j (\lambda^2 + p_j \lambda + q_j)^{m_j} \tag{7.9}$$

where $b = 0$ if n is even, $b = 1$ otherwise, $m_j \geqslant 1$ ($m_j \in \mathbb{N}$) is the multiplicity of the quadratic factor $\lambda^2 + p_j\lambda + q_j$, and $r, p_j, q_j \in \mathbb{R}$. Thus, the absolute values of all roots (real and complex ones) of $\chi_A(\lambda)$ are at most one if and only if $|r| \leqslant 1$ (in case $b = 1$) and the absolute values of the roots of all quadratic factors are at most one. But when does the latter condition hold for a given quadratic factor $\lambda^2 + p_j\lambda + q_j$? By the quadratic formula, we obtain the roots

$$\lambda_{1,2} = -\frac{p_j}{2} \pm \frac{\sqrt{p_j^2 - 4q_j}}{2}$$

If the discriminant $p_j^2 - 4q_j$ is negative, then both roots are complex, i.e.,

$$\lambda_{1,2} = -\frac{p_j}{2} \pm i \frac{\sqrt{4q_j - p_j^2}}{2}$$

and have absolute value $|\lambda_1| = |\lambda_2| = \sqrt{q_j}$. Hence, we demand $\sqrt{q_j} \leqslant 1$, or equivalently, $q_j \leqslant 1$. In the other case, if $p_j^2 - 4q_j \geqslant 0$, both roots are real, and the constraints $|\lambda_1| \leqslant 1$ and $|\lambda_2| \leqslant 1$ simplify to

$$-2 \leqslant p_j \leqslant 2 \text{ and } -(q_j + 1) \leqslant p_j \leqslant q_j + 1$$

As a consequence, the spectral radius of the matrix $A \in \mathbb{R}^{n \times n}$ with characteristic polynomial (7.9) is at most one if and only if $b = 1$ implies $-1 \leqslant r \leqslant 1$ and for all quadratic factors $\lambda^2 + p_j\lambda + q_j$ in (7.9),

$$(p_j^2 - 4q_j < 0 \wedge q_j \leqslant 1) \vee (p_j^2 - 4q_j \geqslant 0 \wedge -2 \leqslant p_j \leqslant 2 \wedge -(q_j+1) \leqslant p_j \leqslant q_j+1)$$

If A is non-negative, the constraints can be simplified according to the Perron-Frobenius theorem (by disregarding the complex roots of $\chi_A(\lambda)$).

Next we sketch how one can adapt the above factorization approach to encode the minimal polynomial m_A of A (cf. [49] for a detailed description due to H. Zankl). In accordance with Lemma 7.17, instead of demanding certain properties of the characteristic polynomial, we encode a monic polynomial p that annihilates A and whose roots have absolute value at most one. Based on (7.9), we introduce variables $C, C_j \in \{0,1\}$ to cancel some factors, such that

$$p(\lambda) = (C\lambda - Cr + 1 - C)^b \cdot \prod_j (C_j\lambda^2 + C_j p_j \lambda + C_j q_j + 1 - C_j)^{m_j}$$

For $C = 0$, the corresponding factor simplifies to one and hence has no effect, whereas the factor contributes to p if $C = 1$. The same holds for the variables C_j. To ensure that all roots of p have absolute value at most one, we use the same approach as above, whereas we have to explicitly add the constraint $p(A) = 0$ to guarantee that p annihilates A. Note that p is always monic by construction. Finding the minimal polynomial m_A is then an optimization problem, i.e., it amounts to minimizing the degree of p.

Non-negative Integer Matrices. If all matrix entries are non-negative integers, one can also apply a different approach. It is based on graph theory and the following lemma, which is an immediate consequence of Lemma 7.34.

Lemma 7.41. *Let $A \in \mathbb{N}^{n\times n}$. Then $\rho(A) > 1$ if and only if $(A^k)_{ii} > 1$ for some $k \in \mathbb{N}$ and $i \in \{1, \ldots, n\}$.* □

Viewing $A \in \mathbb{N}^{n\times n}$ as the adjacency matrix of the associated directed weighted graph $G(A)$ of n vertices, numbered from 1 to n, the condition $(A^k)_{ii} > 1$ for some $k \in \mathbb{N}$ and $i \in \{1, \ldots, n\}$ mentioned in the previous lemma holds if and only if

1. there is a cycle in $G(A)$ containing at least one edge of weight $w > 1$, or
2. there are (at least) two different paths (cycles) from some vertex to itself.

This is due to the well-known fact that the entry $(A^k)_{ij}$ equals the sum of the weights of all distinct paths in $G(A)$ of length k from vertex i to vertex j (note that the weight of any path is at least one since $A \in \mathbb{N}^{n\times n}$). Hence, we conclude that $\rho(A) \leqslant 1$ if and only if neither of the two conditions holds. As every cycle of $G(A)$ is composed of simple cycles, that is, cycles with no repeated vertices (aside from the necessary repetition of the start and end vertex), we may restrict to simple cycles for both conditions.

Next we make two important observations. First, for $A \in \mathbb{N}^{n\times n}$, $G(A)$ cannot have a simple cycle containing an edge of weight greater than one if every matrix in the set $\{A, A^2, \ldots, A^n\}$ has diagonal entries less than or equal to one. Concerning the second condition, let us assume that there are two different simple cycles C_1 and C_2 of length l_1 and l_2, $1 \leqslant l_1, l_2 \leqslant n$, from some vertex i to itself. Considering all paths of length $\mathsf{lcm}(l_1, l_2)$, the least common multiple of l_1 and l_2, we clearly have $(A^{\mathsf{lcm}(l_1,l_2)})_{ii} > 1$. In addition, we also have $(A^{l_1+l_2})_{ii} > 1$ because there are two different cycles, each of weight at least one, from vertex i to itself of length l_1+l_2, namely, the concatenation of C_1 and C_2 as well as the concatenation of C_2 and C_1. Hence, we can detect the existence of the cycles C_1 and C_2 by examining the diagonal entries of all matrices in the set $\{A, A^2, \ldots, A^m\}$, where $m = \min(l_1 + l_2, \mathsf{lcm}(l_1, l_2))$. More generally, we can detect any pair of cycles satisfying condition 2 by examining the diagonal entries of the matrices in the set $\{A, A^2, \ldots, A^{p(n)}\}$, where $p(n) = \max\{\min(l_1 + l_2, \mathsf{lcm}(l_1, l_2)) \mid 1 \leqslant l_1, l_2 \leqslant n\}$. The left part of the table below shows the values of $p(n)$ for various values of n.

n	1	2	3	4	5	6
$p(n)$	1	2	5	7	9	11

n	1	2	3	4	5	6
$q(n)$	1	2	3	5	7	9

The following lemma provides a more intelligible characterization of this function. Its proof can be found in Appendix A.

Lemma 7.42. *For $n \in \mathbb{N}$, $n \geqslant 1$, we have*

$$p(n) = \begin{cases} n & \text{if } n = 1 \text{ or } n = 2 \\ 2n - 1 & \text{if } n \geqslant 3 \end{cases} \qquad \square$$

In particular, we observe that $p(n) \geqslant n$ for all $n \geqslant 1$, from which we draw the following conclusion. If every matrix in the set $\{A, A^2, \ldots, A^{p(n)}\}$, the cardinality of which is linear in n, has diagonal entries less than or equal to one, then neither condition 1 nor condition 2 can hold, which implies $\rho(A) \leqslant 1$. The converse is obvious. So we obtain the following result.

Lemma 7.43. *For $A \in \mathbb{N}^{n\times n}$, $\rho(A) \leqslant 1$ holds if and only if every matrix in the set $\{A, A^2, \ldots, A^{p(n)}\}$ has diagonal entries less than or equal to one.* $\square$

Now let us apply this result to matrix interpretations. By definition, all matrices of a monotone matrix interpretation $\mathcal{M}$ must have a top-left entry of at least one. Hence, this is also true for the maximum matrix A of $S_{\mathcal{M}}$ (except for the pathological case when there are no matrices in $\mathcal{M}$, which we ignore here). In other words, in $G(A)$, vertex 1 has a loop (of length one) to itself. This corresponds to a dimension reduction by one for precluding all instances of condition 2. More precisely, we do not have to consider the cases $l_1 = n$ or $l_2 = n$ because then not only C_1 and C_2 but also C_1 (resp. C_2) and the loop of vertex 1 satisfy condition 2 (for $n > 1$), and we can detect this by examining the diagonal entries of the matrix A^n, which has to be considered anyway for precluding all instances of condition 1. Therefore, if $A_{11} > 0$, we have $\rho(A) \leqslant 1$ if and only if every matrix in the set $\{A, A^2, \ldots, A^{q(n)}\}$ has diagonal entries less than or equal to one, where $q(n) = \max(n, p(n-1))$ for $n > 1$ and $q(1) = p(1) = 1$. As a consequence of Lemma 7.42, the function $q(n)$ can be characterized as follows.

Corollary 7.44. *For $n \in \mathbb{N}$, $n \geqslant 1$, we have*

$$q(n) = \begin{cases} n & \text{if } n = 1 \text{ or } n = 2 \\ 2n - 3 & \text{if } n \geqslant 3 \end{cases} \qquad \square$$

Some values for $q(n)$ are displayed in the right part of the above table.

7.7 Conclusion

In this chapter, we studied matrix interpretations inducing polynomial upper bounds on the derivational complexity of compatible TRSs. Using results from linear algebra, we identified in Section 7.4 a simple sufficient condition for polynomial boundedness of a matrix interpretation based on the spectral radius of its maximum matrix. In contrast to the conditions given in related work [53, 73], it also applies to interpretations over $\mathbb{Q}$ and $\mathbb{R}$. In Section 7.5, we then showed that the joint spectral radius of the matrices of the interpretation provides a better characterization of polynomial boundedness and gave conditions for the decidability of the latter. In this way, we obtain a complete characterization of polynomially bounded matrix interpretations over $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$, complementing the result of [73], which gives a necessary and sufficient condition for interpretations over $\mathbb{N}$. In fact, a thorough examination of the automata-based approach of [73] and the joint spectral radius theory approach of Section 7.5 reveals that the latter subsumes the former. We give the proof of this claim in [49]. Here, we sketch it as follows. In [73] the growth of a matrix interpretation $\mathcal{M}$ over $\mathbb{N}$ is bounded by the growth of an associated (trim) weighted automaton $\mathcal{A}$ with weights in $\mathbb{N}$ whose transitions are given by the matrices occurring in $\mathcal{M}$, such that polynomial boundedness of $\mathcal{M}$ follows from polynomial boundedness of the weight function computed by $\mathcal{A}$ (and vice versa). To this end, the growth of the weight function of a weighted automaton is studied by relating it to the growth of the ambiguity of a non-weighted automaton, which facilitates the use of results from Weber and Seidl who provide in [74] a complete characterization of the degree of growth of the ambiguity of nondeterministic finite automata (NFAs). However, this detour via classical automata is not absolutely necessary as one can lift the concepts of [74] directly to weighted automata. This was done, for example, by Kuich in [42]. Then the growth of (the weight function of) the automaton $\mathcal{A}$ above is bounded by a polynomial if and only if $\mathcal{A}$ does not satisfy a property called EDA (as in [74], but lifted from NFAs to weighted automata), which is, by definition, equivalent to the condition of Lemma 7.34 (for trim automata). Therefore, $\mathcal{A}$ does not satisfy EDA if and only if $\rho(S_{\mathcal{M}}) \leqslant 1$, which is equivalent to polynomial boundedness of $\mathcal{M}$ (resp. $\mathsf{growth}_{S_{\mathcal{M}}}$) according to Lemma 7.37 (note that $S_{\mathcal{M}}$ is the set of transition matrices of $\mathcal{A}$). Moreover, if the growth of $\mathcal{A}$ is polynomial, then the exact degree of growth is given by the largest $d \in \mathbb{N}$ for which $\mathcal{A}$ satisfies a property called IDA_d (also lifted from NFAs to weighted automata). Again, by definition, the value of d is exactly the growth rate of $\mathsf{growth}_{S_{\mathcal{M}}}$ mentioned in Theorem 7.36. As a consequence, for matrix interpretations over $\mathbb{N}$, the joint spectral radius theory approach and the automata-based approach of [73] actually coincide, despite their, at first sight, seemingly different presentations.

The approaches described in this chapter to obtain polynomial upper bounds on the derivational complexity of rewrite systems from compatible matrix interpretations have been implemented in the complexity tool CaT [75] (by the tool developers). Experimental results can be found in [59] and [49]. Summarizing these results, we note that the joint spectral radius approach of Section 7.5 is not only more powerful in theory but also in practice. However, this increase in power comes along with a considerably more elaborate implementation compared to the spectral radius approach of Section 7.4, which is accordingly easier to implement, especially Theorem 7.24, but already gives good results, in particular in comparison with the method of triangular matrix interpretations of [53], as it often succeeds in inferring tighter bounds from such interpretations; e.g., the number of linear (quadratic) upper bounds increases by 84% (16%) if one compares Theorems 7.15 and $7.24_{\mathbb{N}}$. It remains to be seen whether joint spectral radius theory will advance implementations for matrix interpretations over $\mathbb{Q}$ and $\mathbb{R}$. For matrices with rational or real entries the joint spectral radius is not computable in general. In the future we will investigate whether good approximations can be obtained to improve the complexity bounds inferred from matrix interpretations over $\mathbb{Q}$ and $\mathbb{R}$.

We conclude this chapter with three additional remarks. First, we mention that our results are not limited to the study of *derivational* complexity, but can also be employed in the context of *runtime* complexity of rewrite systems [29]. Next we observe that matrix interpretations are incomplete when it comes to establishing polynomial derivational complexity. This shows that new ideas are necessary to obtain a complete characterization of TRSs with polynomial derivational complexity (cf. RTA open problem #107).[7]

Lemma 7.45. *There exists a TRS with linear derivational complexity that is compatible with a monotone matrix interpretation but not with a polynomially bounded one.*

Proof. In order to prove this result, we extend the TRS $\mathcal{R}_{\text{jsr}}$ of Section 7.4 (before Lemma 7.30) by the rule $\mathsf{b}(x) \to \mathsf{g}(x)$. The resulting TRS $\mathcal{S}$ has linear derivational complexity as it is match-bounded by 3 [24], and we obtain a compatible monotone matrix interpretation by changing the interpretation of b in the proof of Lemma 7.30 to $\mathsf{b}_{\mathcal{M}}(\vec{x}) = B\vec{x} + \vec{b}$, where

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \qquad \vec{b} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$$

That there cannot exist a polynomially bounded matrix interpretation (in the sense of Definition 7.6) for $\mathcal{S}$ follows from the proof of Lemma 7.30 (replacing M by B) since $B \geqslant \max(I, G)$, and hence some entries in B^k exhibit exponential growth in k. According to Lemma 7.10, this implies that there cannot exist a compatible monotone matrix interpretation $\mathcal{M}$ for which $\mathsf{growth}_{\mathcal{M}}$ is polynomially bounded. ☐

Corollary 7.46. *The converse of Corollary 7.12 (resp. Lemma 7.3 with $f(k)$ set to k^d, $d \in \mathbb{N}$) does not hold.* ☐

Finally, we report on a contribution of our results to the theory of Lindenmayer systems (L systems, [67]). Such systems use iterated morphisms to generate languages over some alphabet Σ. The simplest type of L system, called D0L system, is a triple $G = (\Sigma, h, w_0)$, where Σ is a finite non-empty alphabet, h an

[7] `http://rtaloop.mancoosi.univ-paris-diderot.fr/problems/107.html`

endomorphism of Σ^* and w_0 an element of Σ^+ (i.e., w_0 is not the empty word). The endomorphism h is typically specified by the images of the elements of Σ, given as a set of *production rules* $\{a \to h(a) \mid a \in \Sigma\}$. A D0L system G generates the sequence $S(G)$ of words $(w_k)_{k\in\mathbb{N}}$, where $w_{k+1} = h(w_k)$ for all $k \in \mathbb{N}$. The *language* of G is the set of words in $S(G)$. If every symbol of Σ occurs in some word of $S(G)$, then G is said to be *reduced*. For each D0L system G, one can effectively compute an *equivalent* reduced D0L system G_{red} generating the same sequence of words $S(G_{\text{red}}) = S(G)$. In what follows, we assume all considered D0L systems to be reduced. The *growth function* of a D0L system G is defined as $f_G(k) = |w_k|$, where $|w_k|$ denotes the length of the word w_k. It is well-known that it can be written as

$$f_G(k) = \pi \cdot M^k \cdot (1, \ldots, 1)^{\mathrm{T}}$$

for some row vector $\pi \in \mathbb{N}^{1\times n}$, where n is the cardinality of the alphabet Σ, and some matrix $M \in \mathbb{N}^{n\times n}$, referred to as the *growth matrix* of G. Every D0L growth function is either exponential or polynomially bounded, the former being characterized by $\rho(M) > 1$ and the latter by $\rho(M) \leqslant 1$. As the growth matrix M is a non-negative integer matrix, we can use Lemma 7.43 for deciding $\rho(M) \leqslant 1$, thereby obtaining the following result.

Corollary 7.47. *Let G be a reduced D0L system with growth matrix $M \in \mathbb{N}^{n\times n}$. The growth function of G is polynomially bounded if and only if every matrix in the set $\{M, M^2, \ldots, M^{p(n)}\}$, where $p(n) = n$ if $n \leqslant 2$ and $p(n) = 2n - 1$ otherwise, has diagonal entries less than or equal to one.* □

A similar result is mentioned in [67] but with $p(n) = 2^n + n - 1$. So Corollary 7.47 considerably reduces the number of matrix powers one has to inspect before one can conclude $\rho(M) \leqslant 1$ or $\rho(M) > 1$. According to Professor Salomaa, one of the authors of [67] and leading expert in the field, this result is novel. In light of this, it might be fruitful to revisit more advanced types of L systems containing several morphisms, corresponding to a set of several matrices, using joint spectral radius theory.

Bibliography

[1] B. Alarcón, S. Lucas, and R. Navarro-Marset. Proving termination with matrix interpretations over the reals. In *Proceedings of the 10th International Workshop on Termination (WST 2009)*, pages 12–15, 2009.

[2] A. S. Amitsur and J. Levitzki. Minimal identities for algebras. *Proceedings of the American Mathematical Society*, 1(4):449–463, 1950.

[3] T. Arts and J. Giesl. Termination of term rewriting using dependency pairs. *Theoretical Computer Science*, 236(1-2):133–178, 2000.

[4] F. Baader and T. Nipkow. *Term Rewriting and All That.* Cambridge University Press, 1998.

[5] R. Bagby. *Introductory Analysis: A Deeper View of Calculus.* Academic Press, 2001.

[6] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in Real Algebraic Geometry.* Springer, second edition, 2006.

[7] J. P. Bell. A gap result for the norms of semigroups of matrices. *Linear Algebra and its Applications*, 402:101–110, 2005.

[8] E. D. Bloch. *The Real Numbers and Real Analysis.* Springer, 2011.

[9] J. Bochnak, M. Coste, and M.-F. Roy. *Real Algebraic Geometry.* Springer, 1998.

[10] G. Bonfante and F. Deloup. Complexity invariance of real interpretations. In *Proceedings of the 7th Annual Conference on Theory and Applications of Models of Computation (TAMC 2010)*, volume 6108 of *Lecture Notes in Computer Science*, pages 139–150, 2010.

[11] C. W. Brown. An overview of QEPCAD B: A tool for real quantifier elimination and formula simplification. *Journal of Japan Society for Symbolic and Algebraic Computation*, 10(1):13–22, 2003.

[12] C. A. Charalambides. *Enumerative Combinatorics.* Chapman & Hall/CRC, 2002.

[13] C. Chuan-Chong and K. Khee-Meng. *Principles and Techniques in Combinatorics.* World Scientific Publishing Company, 1992.

[14] G. E. Collins. Quantifier elimination for real closed fields by cylindrical algebraic decomposition. In *Proceedings of the 2nd GI Conference on Automata Theory and Formal Languages*, volume 33 of *Lecture Notes in Computer Science*, pages 134–183, 1975.

[15] G. E. Collins and H. Hong. Partial cylindrical algebraic decomposition for quantifier elimination. *Journal of Symbolic Computation*, 12(3):299–328, 1991.

[16] E. Contejean, C. Marché, A.-P. Tomás, and X. Urbain. Mechanically proving termination using polynomial interpretations. *Journal of Automated Reasoning*, 34(4):325–363, 2005.

[17] P. Courtieu, G. Gbedo, and O. Pons. Improved matrix interpretation. In *Proceedings of the 36th International Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM 2010)*, volume 5901 of *Lecture Notes in Computer Science*, pages 283–295, 2010.

[18] N. Dershowitz. A note on simplification orderings. *Information Processing Letters*, 9(5):212–215, 1979.

[19] J. Dick, J. Kalmus, and U. Martin. Automating the Knuth Bendix ordering. *Acta Informatica*, 28(2):95–119, 1990.

[20] J. Endrullis, J. Waldmann, and H. Zantema. Matrix interpretations for proving termination of term rewriting. *Journal of Automated Reasoning*, 40(2–3):195–220, 2008.

[21] C. Fuhs, J. Giesl, A. Middeldorp, P. Schneider-Kamp, R. Thiemann, and H. Zankl. Maximal termination. In *Proceedings of the 19th International Conference on Rewriting Techniques and Applications (RTA 2008)*, volume 5117 of *Lecture Notes in Computer Science*, pages 110–125, 2008.

[22] A. Gebhardt, D. Hofbauer, and J. Waldmann. Matrix evolutions. In *Proceedings of the 9th International Workshop on Termination (WST 2007)*, pages 4–8, 2007.

[23] A. Gebhardt and J. Waldmann. Weighted automata define a hierarchy of terminating string rewriting systems. *Acta Cybernetica*, 19(2):295–312, 2009.

[24] A. Geser, D. Hofbauer, J. Waldmann, and H. Zantema. On tree automata that certify termination of left-linear term rewriting systems. *Information and Computation*, 205(4):512–534, 2007.

[25] J. Giesl, R. Thiemann, and P. Schneider-Kamp. The dependency pair framework: Combining techniques for automated termination proofs. In *Proceedings of the 11th International Conference on Logic for Programming, Artificial Intelligence and Reasoning (LPAR 11)*, volume 3452 of *Lecture Notes in Artificial Intelligence*, pages 301–331, 2005.

[26] J. Giesl, R. Thiemann, P. Schneider-Kamp, and S. Falke. Mechanizing and improving dependency pairs. *Journal of Automated Reasoning*, 37(3):155–203, 2006.

[27] N. Hirokawa and A. Middeldorp. Automating the dependency pair method. *Information and Computation*, 199(1-2):172–199, 2005.

[28] N. Hirokawa and A. Middeldorp. Tyrolean Termination Tool: Techniques and features. *Information and Computation*, 205(4):474–511, 2007.

[29] N. Hirokawa and G. Moser. Automated complexity analysis based on the dependency pair method. In *Proceedings of the 4th International Joint Conference on Automated Reasoning (IJCAR 2008)*, volume 5195 of *Lecture Notes in Artificial Intelligence*, pages 364–379, 2008.

[30] D. Hofbauer. Termination proofs by context-dependent interpretations. In *Proceedings of the 12th International Conference on Rewriting Techniques and Applications (RTA 2001)*, volume 2051 of *Lecture Notes in Computer Science*, pages 108–121, 2001.

[31] D. Hofbauer and C. Lautemann. Termination proofs and the length of derivations (preliminary version). In *Proceedings of the 3rd International Conference on Rewriting Techniques and Applications (RTA 1989)*, volume 355 of *Lecture Notes in Computer Science*, pages 167–177, 1989.

[32] D. Hofbauer and J. Waldmann. Termination of $\{aa \to bc, bb \to ac, cc \to ab\}$. *Information Processing Letters*, 98(4):156–158, 2006.

[33] D. Hofbauer and J. Waldmann. Termination of string rewriting with matrix interpretations. In *Proceedings of the 17th International Conference on Rewriting Techniques and Applications (RTA 2006)*, volume 4098 of *Lecture Notes in Computer Science*, pages 328–342, 2006.

[34] H. Hong and D. Jakuš. Testing positiveness of polynomials. *Journal of Automated Reasoning*, 21(1):23–38, 1998.

[35] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

[36] R. M. Jungers. *The Joint Spectral Radius: Theory and Applications*. Springer, 2009.

[37] R. M. Jungers, V. Protasov, and V. D. Blondel. Efficient algorithms for deciding the type of growth of products of integer matrices. *Linear Algebra and its Applications*, 428(10):2296–2311, 2008.

[38] A. Koprowski and J. Waldmann. Arctic termination ... below zero. In *Proceedings of the 19th International Conference on Rewriting Techniques and Applications (RTA 2008)*, volume 5117 of *Lecture Notes in Computer Science*, pages 202–216, 2008.

[39] A. Koprowski and J. Waldmann. Max/plus tree automata for termination of term rewriting. *Acta Cybernetica*, 19(2):357–392, 2009.

[40] K. Korovin and A. Voronkov. Orienting rewrite rules with the Knuth-Bendix order. *Information and Computation*, 183(2):165–186, 2003.

[41] M. Korp, C. Sternagel, H. Zankl, and A. Middeldorp. Tyrolean Termination Tool 2. In *Proceedings of the 20th International Conference on Rewriting Techniques and Applications (RTA 2009)*, volume 5595 of *Lecture Notes in Computer Science*, pages 295–304, 2009.

[42] W. Kuich. Finite automata and ambiguity. Technical Report 253, Institute für Informationsverarbeitung, Technische Universität Graz und ÖCG, 1988.

[43] D. Lankford. On proving term rewrite systems are noetherian. Technical Report MTP-3, Louisiana Technical University, Ruston, LA, USA, 1979.

[44] I. Lepper. Derivation lengths and order types of Knuth-Bendix orders. *Theoretical Computer Science*, 269(1-2):433–450, 2001.

[45] S. Lucas. Polynomials over the reals in proofs of termination: From theory to practice. *Theoretical Informatics and Applications*, 39(3):547–586, 2005.

[46] S. Lucas. On the relative power of polynomials with real, rational, and integer coefficients in proofs of termination of rewriting. *Applicable Algebra in Engineering, Communication and Computing*, 17(1):49–73, 2006.

[47] S. Lucas. Practical use of polynomials over the reals in proofs of termination. In *Proceedings of the 9th International Conference on Principles and Practice of Declarative Programming (PPDP 2007)*, pages 39–50. Association of the Computing Machinery, 2007.

[48] S. Lucas. From matrix interpretations over the rationals to matrix interpretations over the naturals. In *Proceedings of the 10th International Conference on Artificial Intelligence and Symbolic Computation (AISC 2010)*, volume 6167 of *Lecture Notes in Artificial Intelligence*, pages 116–131, 2010.

[49] A. Middeldorp, G. Moser, F. Neurauter, J. Waldmann, and H. Zankl. Joint spectral radius theory for automated complexity analysis of rewrite systems. In *Proceedings of the 4th International Conference on Algebraic Informatics (CAI 2011)*, volume 6742 of *Lecture Notes in Computer Science*, pages 1–20, 2011.

[50] A. Middeldorp and H. Zantema. Simple termination of rewrite systems. *Theoretical Computer Science*, 175(1):127–158, 1997.

[51] G. Moser. *Proof Theory at Work: Complexity Analysis of Term Rewrite Systems.* Habilitation thesis, University of Innsbruck, 2009.

[52] G. Moser and A. Schnabl. The derivational complexity induced by the dependency pair method. *Logical Methods in Computer Science*, 7(3:1):1 – 38, 2011.

[53] G. Moser, A. Schnabl, and J. Waldmann. Complexity analysis of term rewriting based on matrix and context dependent interpretations. In *Proceedings of the 28th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2008)*,

volume 2 of *Leibniz International Proceedings in Informatics*, pages 304–315, 2008.

[54] F. Neurauter and A. Middeldorp. Polynomial interpretations over the reals do not subsume polynomial interpretations over the integers. In *Proceedings of the 21st International Conference on Rewriting Techniques and Applications (RTA 2010)*, volume 6 of *Leibniz International Proceedings in Informatics*, pages 243–258, 2010.

[55] F. Neurauter and A. Middeldorp. Revisiting matrix interpretations for proving termination of term rewriting. In *Proceedings of the 22nd International Conference on Rewriting Techniques and Applications (RTA 2011)*, volume 10 of *Leibniz International Proceedings in Informatics*, pages 251–266, 2011.

[56] F. Neurauter and A. Middeldorp. On the domain and dimension hierarchy of matrix interpretations. In *Proceedings of the 18th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR 18)*, volume 7180 of *Lecture Notes in Computer Science*, pages 320–334, 2012.

[57] F. Neurauter and A. Middeldorp. Polynomial interpretations over the natural, rational and real numbers revisited. *Logical Methods in Computer Science*, submitted.

[58] F. Neurauter, A. Middeldorp, and H. Zankl. Monotonicity criteria for polynomial interpretations over the naturals. In *Proceedings of the 5th International Joint Conference on Automated Reasoning (IJCAR 2010)*, volume 6173 of *Lecture Notes in Artificial Intelligence*, pages 502–517, 2010.

[59] F. Neurauter, H. Zankl, and A. Middeldorp. Revisiting matrix interpretations for polynomial derivational complexity of term rewriting. In *Proceedings of the 17th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR 17)*, volume 6397 of *Lecture Notes in Computer Science*, pages 550–564, 2010.

[60] M. Oberguggenberger and A. Ostermann. *Analysis for Computer Scientists.* Springer, 2011.

[61] P. A. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization.* PhD thesis, California Institute of Technology, 2000.

[62] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming*, 96(2):293–320, 2003.

[63] T. Pheidas. Extensions of Hilbert's tenth problem. *Journal of Symbolic Logic*, 59(2):372–397, 1994.

[64] B. Reznick. Some concrete aspects of Hilbert's 17th problem. In *Real Algebraic Geometry and Ordered Structures*, volume 253 of *Contemporary Mathematics*, pages 251–272, 2000.

[65] J. Robinson. Definability and decision problems in arithmetic. *Journal of Symbolic Logic*, 14(2):98–114, 1949.

[66] H. E. Rose. *Linear Algebra: A Pure Mathematical Approach*. Birkhäuser, 2002.

[67] G. Rozenberg and A. Salomaa. *The Mathematical Theory of L Systems*. Academic Press, Inc., 1980.

[68] D. Serre. *Matrices: Theory and Applications*. Springer, 2002.

[69] A. Tarski. *A Decision Method for Elementary Algebra and Geometry*. University of California Press, 1951.

[70] Terese. *Term Rewriting Systems*, volume 55 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2003.

[71] R. Thiemann. *The DP Framework for Proving Termination of Term Rewriting*. PhD thesis, RWTH Aachen, 2007. Available as Technical Report AIB-2007-17.

[72] Termination problems data base, version 7.0.2, 2010. Available from http://termcomp.uibk.ac.at/status/downloads/tpdb-7.0.2.tar.gz.

[73] J. Waldmann. Polynomially bounded matrix interpretations. In *Proceedings of the 21st International Conference on Rewriting Techniques and Applications (RTA 2010)*, volume 6 of *Leibniz International Proceedings in Informatics*, pages 357–372, 2010.

[74] A. Weber and H. Seidl. On the degree of ambiguity of finite automata. *Theoretical Computer Science*, 88(2):325–349, 1991.

[75] H. Zankl and M. Korp. Modular complexity analysis via relative complexity. In *Proceedings of the 21st International Conference on Rewriting Techniques and Applications (RTA 2010)*, volume 6 of *Leibniz International Proceedings in Informatics*, pages 385–400, 2010.

[76] H. Zankl and A. Middeldorp. Satisfiability of non-linear (ir)rational arithmetic. In *Proceedings of the 16th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR 16)*, volume 6355 of *Lecture Notes in Artificial Intelligence*, pages 481–500, 2010.

[77] H. Zantema. Termination of term rewriting: Interpretation and type elimination. *Journal of Symbolic Computation*, 17(1):23–50, 1994.

Appendix A

Supplementary Proofs

A.1 Proofs of Chapter 4

Lemma A.1. *Let $\mathcal{P}$ be a weakly (strictly) monotone polynomial interpretation over $\mathbb{R}$ such that $\mathcal{S}_1 \subseteq >_{\mathcal{P}}$ and $\mathcal{S}_2 \subseteq \geqslant_{\mathcal{P}}$ for some finite sets of rewrite rules $\mathcal{S}_1$ and $\mathcal{S}_2$. Then there exists a weakly (strictly) monotone polynomial interpretation $\mathcal{N}$ over $\mathbb{R}$ such that $\mathcal{S}_1 \subseteq >_{\mathcal{N}}$, $\mathcal{S}_2 \subseteq \geqslant_{\mathcal{N}}$ and all coefficients of the polynomials occurring in $\mathcal{N}$ are in $\mathbb{R}_{\mathsf{alg}}$.*

Proof. Let $\mathcal{F}$ denote the signature associated with $\mathcal{S}_1 \cup \mathcal{S}_2$, and let δ be a positive real number and $\mathcal{P} = (\{f_{\mathbb{R}}\}_{f \in \mathcal{F}}, \delta)$ a weakly (strictly) monotone polynomial interpretation over $\mathbb{R}$ such that $\mathcal{S}_1 \subseteq >_{\mathcal{P}}$ and $\mathcal{S}_2 \subseteq \geqslant_{\mathcal{P}}$. That is, $\mathcal{P}$ satisfies the following conditions:

(a) for each n-ary symbol $f \in \mathcal{F}$, $f_{\mathbb{R}}(x_1, \ldots, x_n) \geqslant 0$ for all $x_1, \ldots, x_n \in \mathbb{R}_0$,

(b) for each symbol $f \in \mathcal{F}$, $f_{\mathbb{R}}$ is monotone with respect to $\geqslant_{\mathbb{R}_0}$ ($>_{\mathbb{R}_0,\delta}$),

(c) for each rule $\ell \to r \in \mathcal{S}_1$, $P_\ell >_{\mathbb{R}_0,\delta} P_r$ for all $x_1, \ldots, x_m \in \mathbb{R}_0$, and

(d) for each rule $s \to t \in \mathcal{S}_2$, $P_s \geqslant_{\mathbb{R}_0} P_t$ for all $y_1, \ldots, y_k \in \mathbb{R}_0$.

Next we treat δ as a variable and replace all coefficients of the polynomials occurring in $\mathcal{P}$ by distinct variables $c_1, \ldots, c_j$. Thus, for each n-ary function symbol $f \in \mathcal{F}$, its interpretation function is a parametric polynomial $f_{\mathbb{R}} \in \mathbb{Z}[x_1, \ldots, x_n, c_1, \ldots, c_j] \subseteq \mathbb{Z}[x_1, \ldots, x_n, c_1, \ldots, c_j, \delta]$, where all non-zero coefficients are 1. As a consequence, we claim that all four of the conditions listed above can be expressed as (conjunctions of) quantified polynomial inequalities of the shape

$$p(x_1, \ldots, x_n, c_1, \ldots, c_j, \delta) \geqslant 0 \text{ for all } x_1, \ldots, x_n \in \mathbb{R}_0 \tag{A.1}$$

for some polynomial $p \in \mathbb{Z}[x_1, \ldots, x_n, c_1, \ldots, c_j, \delta]$. This is easy to see for the first condition. For the last two conditions, and by Lemma 3.15 (resp. Corollary 3.17) also for the second condition, it is a direct consequence of the nature of the interpretation functions and the usual closure properties of polynomials. Now any of the quantified inequalities (A.1) can readily be expressed as a formula in the language of ordered fields with coefficients in $\mathbb{Z}$, where $c_1, \ldots, c_j$ and δ are the only free variables. By taking the conjunction of all these formulas, existentially quantifying δ and adding the conjunct $\delta > 0$, we obtain a formula Φ in the language of ordered fields with free variables $c_1, \ldots, c_j$

and coefficients in $\mathbb{Z}$ (as $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{F}$ are finite). By assumption (on $\mathcal{P}$), we know that there are coefficients $C_1, \dots, C_j \in \mathbb{R}$ such that $\Phi(C_1, \dots, C_j)$ is true in $\mathbb{R}$, i.e., there exists a satisfying assignment for Φ in $\mathbb{R}$ mapping its free variables $c_1, \dots, c_j$ to $C_1, \dots, C_j \in \mathbb{R}$. In order to prove the lemma, we have to show that there also exists a satisfying assignment mapping each free variable to a real algebraic number. We reason as follows. By [6, Theorem 2.77], there is a quantifier-free formula Ψ with free variables $c_1, \dots, c_j$ and coefficients in $\mathbb{Z}$ that is $\mathbb{R}$-equivalent to Φ, i.e., for all $y_1, \dots, y_j \in \mathbb{R}$, $\Phi(y_1, \dots, y_j)$ is true in $\mathbb{R}$ if and only if $\Psi(y_1, \dots, y_j)$ is true in $\mathbb{R}$. Hence, by assumption (on $\mathcal{P}$), we have that $\Psi(C_1, \dots, C_j)$ is true in $\mathbb{R}$. Therefore, the sentence $\exists c_1 \cdots \exists c_j \, \Psi$ is true in $\mathbb{R}$ as well. But by Theorem 4.3 (resp. [6, Theorem 2.80]), this sentence is true in $\mathbb{R}$ if and only if it is true in $\mathbb{R}_{\mathsf{alg}}$. So there exists an assignment for Ψ in $\mathbb{R}_{\mathsf{alg}}$ mapping its free variables $c_1, \dots, c_j$ to $C'_1, \dots, C'_j \in \mathbb{R}_{\mathsf{alg}}$ such that $\Psi(C'_1, \dots, C'_j)$ is true in $\mathbb{R}_{\mathsf{alg}}$, and hence also in $\mathbb{R}$ as Ψ is a boolean combination of atomic formulas in the variables $c_1, \dots, c_j$ with coefficients in $\mathbb{Z}$. But then $\Phi(C'_1, \dots, C'_j)$ is true in $\mathbb{R}$ as well because of the $\mathbb{R}$-equivalence of Φ and Ψ. $\square$

A.2 Proofs of Chapter 6

Termination of the TRS $\mathcal{R}_{\mathbb{R}}$ of Chapter 6 consisting of the rules

$$\begin{aligned}
\mathsf{f}(\mathsf{f}(x)) &\to \mathsf{g}(x) & \mathsf{k}(x, x, \mathsf{b}_1) &\to \mathsf{k}(\mathsf{g}(x), \mathsf{b}_2, \mathsf{b}_2) \\
\mathsf{g}(\mathsf{c}(x)) &\to \mathsf{f}(\mathsf{c}(\mathsf{f}(x))) & \mathsf{k}(x, \mathsf{a}_2, \mathsf{b}_1) &\to \mathsf{k}(\mathsf{a}_1, x, \mathsf{b}_1) \\
\mathsf{f}(\mathsf{f}(\mathsf{f}(\mathsf{f}(x)))) &\to \mathsf{k}(x, x, x) & \mathsf{k}(\mathsf{a}'_1, x, \mathsf{b}_1) &\to \mathsf{k}(x, \mathsf{a}'_2, \mathsf{b}_1) \\
\mathsf{k}(\mathsf{g}(x), \mathsf{b}_3, \mathsf{b}_3) &\to \mathsf{k}(x, x, \mathsf{b}_4)
\end{aligned}$$

can be shown by a monotone matrix interpretation $\mathcal{M}$ over $\mathbb{N}$ of dimension two that is compatible with all rules of $\mathcal{R}_{\mathbb{R}}$:

$$\begin{aligned}
\mathsf{f}_{\mathcal{M}}(\vec{x}) &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \mathsf{a_1}_{\mathcal{M}} &= \begin{pmatrix} 2 \\ 0 \end{pmatrix} & \mathsf{b_1}_{\mathcal{M}} &= \begin{pmatrix} 0 \\ 2 \end{pmatrix} \\
\mathsf{g}_{\mathcal{M}}(\vec{x}) &= \begin{pmatrix} 2 & 0 \\ 2 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \mathsf{a_2}_{\mathcal{M}} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \mathsf{b_2}_{\mathcal{M}} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\mathsf{k}_{\mathcal{M}}(\vec{x}, \vec{y}, \vec{z}) &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \vec{y} + \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix} \vec{z} & \mathsf{a'_1}_{\mathcal{M}} &= \begin{pmatrix} 2 \\ 0 \end{pmatrix} & \mathsf{b_3}_{\mathcal{M}} &= \begin{pmatrix} 2 \\ 2 \end{pmatrix} \\
\mathsf{c}_{\mathcal{M}}(\vec{x}) &= \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 2 \\ 0 \end{pmatrix} & \mathsf{a'_2}_{\mathcal{M}} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \mathsf{b_4}_{\mathcal{M}} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\end{aligned}$$

Likewise, the TRS $\mathcal{R}_{\mathbb{Q}}$ consisting of the rules

$$\begin{aligned}
\mathsf{h}(\mathsf{f}(x)) &\to \mathsf{g}(x) & \mathsf{k}_2(x,x,\mathsf{b}_1) &\to \mathsf{k}_2(\mathsf{h}(x),\mathsf{b}_2,\mathsf{b}_2)\\
\mathsf{g}(\mathsf{c}(x)) &\to \mathsf{h}(\mathsf{c}(\mathsf{f}(x))) & \mathsf{k}_2(\mathsf{h}(x),\mathsf{b}_3,\mathsf{b}_3) &\to \mathsf{k}_2(x,x,\mathsf{b}_4)\\
\mathsf{h}(\mathsf{h}(x)) &\to \mathsf{k}_2(x,x,x) & \mathsf{k}_3(x,\mathsf{a}_2,\mathsf{a}_3,\mathsf{b}_1) &\to \mathsf{k}_3(\mathsf{a}_1,x,\mathsf{a}_3,\mathsf{b}_1)\\
\mathsf{h}(\mathsf{h}(x)) &\to \mathsf{k}_3(x,x,x,x) & \mathsf{k}_3(\mathsf{a}_1',x,\mathsf{a}_3',\mathsf{b}_1) &\to \mathsf{k}_3(x,\mathsf{a}_2',\mathsf{a}_3',\mathsf{b}_1)\\
\mathsf{f}(\mathsf{f}(\mathsf{f}(x))) &\to \mathsf{k}_2(x,x,x) & \mathsf{k}_3(\mathsf{a}_1,x,\mathsf{a}_3,\mathsf{b}_1) &\to \mathsf{k}_3(\mathsf{a}_1,\mathsf{a}_2,x,\mathsf{b}_1)\\
\mathsf{f}(\mathsf{f}(\mathsf{f}(\mathsf{f}(x)))) &\to \mathsf{k}_3(x,x,x,x) & \mathsf{k}_3(\mathsf{a}_1',\mathsf{a}_2',x,\mathsf{b}_1) &\to \mathsf{k}_3(\mathsf{a}_1',x,\mathsf{a}_3',\mathsf{b}_1)\\
\mathsf{k}_2(x,\mathsf{a}_2,\mathsf{b}_1) &\to \mathsf{k}_2(\mathsf{a}_1,x,\mathsf{b}_1) & \mathsf{k}_3(x,x,x,\mathsf{b}_1) &\to \mathsf{k}_3(\mathsf{g}(x),\mathsf{b}_2,\mathsf{b}_2,\mathsf{b}_2)\\
\mathsf{k}_2(\mathsf{a}_1',x,\mathsf{b}_1) &\to \mathsf{k}_2(x,\mathsf{a}_2',\mathsf{b}_1) & \mathsf{k}_3(\mathsf{g}(x),\mathsf{b}_3,\mathsf{b}_3,\mathsf{b}_3) &\to \mathsf{k}_3(x,x,x,\mathsf{b}_4)
\end{aligned}$$

can be shown terminating by a compatible, monotone matrix interpretation over $\mathbb{N}$ of dimension two:

$$\mathsf{c}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix}1 & 0\\1 & 2\end{pmatrix}\vec{x} + \begin{pmatrix}3\\1\end{pmatrix} \qquad \mathsf{f}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix}1 & 3\\1 & 0\end{pmatrix}\vec{x} + \begin{pmatrix}1\\0\end{pmatrix}$$

$$\mathsf{g}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix}3 & 6\\0 & 0\end{pmatrix}\vec{x} + \begin{pmatrix}0\\2\end{pmatrix} \qquad \mathsf{h}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix}2 & 2\\0 & 0\end{pmatrix}\vec{x} + \begin{pmatrix}0\\2\end{pmatrix}$$

$$\mathsf{k}_{2\mathcal{M}}(\vec{x}_1,\vec{x}_2,\vec{x}_3) = \begin{pmatrix}1 & 1\\0 & 0\end{pmatrix}(\vec{x}_1+\vec{x}_2) + \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}\vec{x}_3 + \begin{pmatrix}0\\2\end{pmatrix}$$

$$\mathsf{k}_{3\mathcal{M}}(\vec{x}_1,\vec{x}_2,\vec{x}_3,\vec{x}_4) = \begin{pmatrix}1 & 2\\0 & 0\end{pmatrix}(\vec{x}_1+\vec{x}_2+\vec{x}_3+\vec{x}_4)$$

$$\mathsf{a}_{1\mathcal{M}} = \begin{pmatrix}1\\2\end{pmatrix} \quad \mathsf{a}'_{1\mathcal{M}} = \begin{pmatrix}7\\4\end{pmatrix} \quad \mathsf{a}_{2\mathcal{M}} = \begin{pmatrix}2\\2\end{pmatrix} \quad \mathsf{a}'_{2\mathcal{M}} = \begin{pmatrix}4\\1\end{pmatrix} \quad \mathsf{a}_{3\mathcal{M}} = \begin{pmatrix}1\\7\end{pmatrix}$$

$$\mathsf{a}'_{3\mathcal{M}} = \begin{pmatrix}1\\0\end{pmatrix} \quad \mathsf{b}_{1\mathcal{M}} = \begin{pmatrix}3\\3\end{pmatrix} \quad \mathsf{b}_{2\mathcal{M}} = \begin{pmatrix}0\\0\end{pmatrix} \quad \mathsf{b}_{3\mathcal{M}} = \begin{pmatrix}3\\0\end{pmatrix} \quad \mathsf{b}_{4\mathcal{M}} = \begin{pmatrix}5\\0\end{pmatrix}$$

Both interpretations were found by $\mathsf{T_TT_2}$ [41].

A.3 Proofs of Chapter 7

In the proof of Lemma 7.42 below, we use the following notation for the division predicate on integer numbers. We write $a \mid b$ to indicate that a non-zero integer a divides some integer b and $a \nmid b$ whenever a does not divide b.

Proof of Lemma 7.42. For $n \in \mathbb{N}$, $n \geqslant 1$, we have

$$\begin{aligned}
p(n) &= \max\{\min(l_1+l_2,\mathsf{lcm}(l_1,l_2)) \mid 1 \leqslant l_1,l_2 \leqslant n\}\\
&= \max\{\min(l_1+l_2,\mathsf{lcm}(l_1,l_2)) \mid 1 \leqslant l_1 \leqslant l_2 \leqslant n\}\\
&= \max(\{\min(l_1+l_2,\mathsf{lcm}(l_1,l_2)) \mid 1 \leqslant l_1 \leqslant l_2 \leqslant n, l_1 \nmid l_2\} \cup\\
&\qquad\quad \{\min(l_1+l_2,\mathsf{lcm}(l_1,l_2)) \mid 1 \leqslant l_1 \leqslant l_2 \leqslant n, l_1 \mid l_2\})\\
&= \max(\{\min(l_1+l_2,\mathsf{lcm}(l_1,l_2)) \mid 1 \leqslant l_1 \leqslant l_2 \leqslant n, l_1 \nmid l_2\} \cup\\
&\qquad\quad \{\min(l_1+l_2,l_2) \mid 1 \leqslant l_1 \leqslant l_2 \leqslant n, l_1 \mid l_2\})\\
&= \max(\{\min(l_1+l_2,\mathsf{lcm}(l_1,l_2)) \mid 1 \leqslant l_1 \leqslant l_2 \leqslant n, l_1 \nmid l_2\} \cup \{1,\ldots,n\})\\
&= \max(\{\min(l_1+l_2,\mathsf{lcm}(l_1,l_2)) \mid 2 \leqslant l_1 < l_2 \leqslant n, l_1 \nmid l_2\} \cup \{1,\ldots,n\})
\end{aligned}$$

It is easily checked that $p(n) = n$ if $n = 1$ or $n = 2$. For the remaining case when $n \geqslant 3$, we use the following fact (which will be proved below):

$$\begin{aligned}\forall n \geqslant 3 \quad \{\min(l_1 + l_2, \mathsf{lcm}(l_1, l_2)) \mid 2 \leqslant l_1 < l_2 \leqslant n, l_1 \nmid l_2\} = \\ \{l_1 + l_2 \mid 2 \leqslant l_1 < l_2 \leqslant n, l_1 \nmid l_2\}\end{aligned} \tag{A.2}$$

Thus, we obtain

$$\begin{aligned} p(n) &= \max(\{l_1 + l_2 \mid 2 \leqslant l_1 < l_2 \leqslant n, l_1 \nmid l_2\} \cup \{1, 2, \ldots, n\}) \\ &= \max(n + (n-1), n) \\ &= 2n - 1 \end{aligned}$$

for all $n \geqslant 3$. Finally, we show (A.2) by induction on n. For $n = 3$, the condition $2 \leqslant l_1 < l_2 \leqslant n$ uniquely instantiates l_1 to two and l_2 to three, such that

$$\begin{aligned}\{\min(l_1 + l_2, \mathsf{lcm}(l_1, l_2)) \mid 2 \leqslant l_1 < l_2 \leqslant n, l_1 \nmid l_2\} = \\ \min(5, 6) = 5 = \{l_1 + l_2 \mid 2 \leqslant l_1 < l_2 \leqslant n, l_1 \nmid l_2\}\end{aligned}$$

In the inductive step, we write the left-hand side of (A.2) as

$$\begin{aligned}\{\min(l_1 + l_2, \mathsf{lcm}(l_1, l_2)) \mid 2 \leqslant l_1 < l_2 \leqslant n - 1, l_1 \nmid l_2\} \cup \\ \{\min(l_1 + n, \mathsf{lcm}(l_1, n)) \mid 2 \leqslant l_1 < n, l_1 \nmid n\}\end{aligned}$$

apply the induction hypothesis to the left subexpression and simplify the right subexpression to obtain

$$\{l_1 + l_2 \mid 2 \leqslant l_1 < l_2 \leqslant n - 1, l_1 \nmid l_2\} \cup \{l_1 + n \mid 2 \leqslant l_1 < n, l_1 \nmid n\}$$

which is equivalent to $\{l_1 + l_2 \mid 2 \leqslant l_1 < l_2 \leqslant n, l_1 \nmid l_2\}$, the right-hand side of (A.2). $\square$

Appendix B

Alternative Base Orders for Matrix Interpretations

As mentioned in Chapter 5, matrix interpretations are a powerful technique for proving termination of term rewrite systems, which is based on the well-known paradigm of interpreting terms into a domain equipped with a suitable well-founded (order) relation. Traditionally, in the method of Endrullis *et al.* [20] one uses vectors of natural numbers as domain, with two vectors being in relation if there is a strict decrease in the respective first components and a weak decrease in all other components. In this appendix, we re-examine the basics of the method, especially focusing on the actual role of the well-founded (order) relation it is based on. Obviously, there are many other such relations on vectors of natural numbers, so why the choice of this particular instance? In [20] the justification is as follows (in addition to the convincing fact that the resulting termination method is very powerful):

> *Of course other orders on vectors could have been chosen, too, but many of them are not suitable for our purpose. For instance, choosing a lexicographic order fails because then multiplication by a constant matrix is not monotone in general.*

But still the question remains whether there exist other (order) relations inducing variants of matrix interpretations that are also useful for proving termination of TRSs. For this purpose, we study various alternative well-founded orders on vectors of (natural) numbers based on vector norms. The underlying idea is that every rewrite step is supposed to decrease the "length" of the associated vectors. This leads directly to the notion of normed vector spaces, norms being a suitable measure of the length or magnitude of a vector. Basically, we consider two classes of orders, weakly decreasing orders, where two vectors are comparable only if there is a weak decrease in every single component, and orders without this property. The conclusion is that the latter kind of orders induces only weak forms of matrix interpretations that are no more powerful than linear polynomial interpretations. For weakly decreasing orders (like the order in [20]), however, the situation is different. That is to say that some of them do indeed induce matrix interpretations that are useful for proving termination. In particular, one of these variants subsumes traditional matrix interpretations and has the additional advantage that it gives rise to a more powerful implementation.

The remainder of this appendix is organized as follows. After extending the preliminaries on matrix interpretations of Chapter 5 with additional definitions and terminology (irrelevant to Chapters 6 and 7) in Section B.1, we introduce in Section B.2 the orders on vectors of natural numbers considered in the sequel. Sections B.3 and B.4 are dedicated to matrix interpretations over weakly decreasing orders and the comparison between them, while Section B.5 features matrix interpretations over non-weakly decreasing orders. Finally, in Section B.6, we present a generalization of traditional matrix interpretations before we conclude in Section B.7.

The results presented in this appendix originally appeared in [55].

B.1 Preliminaries

The remainder of this appendix builds upon the background material introduced in Chapters 1 and 5 (some of which is repeated below) and the following additional definitions.

For any commutative ring R, we denote the ring of all *n-dimensional square matrices* over R by $R^{n\times n}$. As usual, the *transpose* of a matrix (vector) M is denoted by M^{T}. For any (column) vector $\vec{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$, $(\vec{x})_i$ denotes its i-th component x_i. Likewise, M_{ij} denotes the entry in the i-th row and j-th column of a matrix M, and M_{j-} (M_{-j}) refers to the j-th row (column). A *zero column* is a column where all entries are zero. For $M \in R^{n\times n}$ and $\mathcal{I} \subseteq \{1, \ldots, n\}$, $(M)_{\mathcal{I}}$ denotes the *submatrix* of M formed by the rows and columns whose indices are in the index set $\mathcal{I}$. A *permutation matrix* is a square matrix whose entries are all 0's and 1's, with exactly one 1 in each row and exactly one 1 in each column. The *cardinality* of a (finite) set S is denoted by $|S|$.

B.2 Well-founded Orders on Vectors of Natural Numbers

In this section, we introduce several well-founded orders on vectors of natural numbers serving as foundation for alternative kinds of matrix interpretations. We consider two classes of orders on $\mathbb{N}^n$, $n \geqslant 1$, weakly decreasing orders and non-weakly decreasing ones.

B.2.1 Weakly Decreasing Orders

A binary relation $>$ on $\mathbb{N}^n$ is *weakly decreasing* if $(x_1, \ldots, x_n)^{\mathrm{T}} > (y_1, \ldots, y_n)^{\mathrm{T}}$ implies $x_i \geqslant_{\mathbb{N}} y_i$ for all $i \in \{1, \ldots, n\}$, i.e., $> \subseteq \geqslant^w$, where $\geqslant^w$ denotes the componentwise partial order on $\mathbb{N}^n$ induced by $\geqslant_{\mathbb{N}}$.

Definition B.1. Let $\mathcal{I} \subseteq \{1, \ldots, n\}$ be a non-empty index set, and let $\vec{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$ and $\vec{y} = (y_1, \ldots, y_n)^{\mathrm{T}}$ be vectors in $\mathbb{N}^n$. We define the binary relations $>^w_1$, $>^w_{\mathcal{I}}$, $>^w_{\Sigma}$, $>^w_{\ell}$ and $>^w_{\mathrm{m}}$ on $\mathbb{N}^n$ as follows:

- Weak decrease + strict decrease in first component:

$$\vec{x} >^w_1 \vec{y} :\iff \vec{x} \geqslant^w \vec{y} \text{ and } x_1 >_{\mathbb{N}} y_1$$

- Weak decrease + strict decrease in some component(s):

$$\vec{x} >^w_{\mathcal{I}} \vec{y} :\iff \vec{x} \geqslant^w \vec{y} \text{ and } \exists\, j \in \mathcal{I}\colon x_j >_{\mathbb{N}} y_j$$

- Weak decrease + strict decrease in sum of components:

$$\vec{x} >^w_{\Sigma} \vec{y} :\iff \vec{x} \geqslant^w \vec{y} \text{ and } \sum_{i=1}^{n} x_i >_{\mathbb{N}} \sum_{i=1}^{n} y_i$$

- Weak decrease + strict decrease in Euclidean length:

$$\vec{x} >^w_{\ell} \vec{y} :\iff \vec{x} \geqslant^w \vec{y} \text{ and } \sum_{i=1}^{n} x_i^2 >_{\mathbb{N}} \sum_{i=1}^{n} y_i^2$$

- Weak decrease + strict decrease in maximum component:

$$\vec{x} >^w_{\mathrm{m}} \vec{y} :\iff \vec{x} \geqslant^w \vec{y} \text{ and } \max_i x_i >_{\mathbb{N}} \max_i y_i$$

It is routine to verify that all these relations are in fact well-founded orders (i.e., transitive and irreflexive relations) on vectors of natural numbers.

Lemma B.2. *The relations $>^w_1$, $>^w_{\mathcal{I}}$, $>^w_{\Sigma}$, $>^w_{\ell}$ and $>^w_{\mathrm{m}}$ are well-founded orders on $\mathbb{N}^n$.* □

The relations listed above are not the only well-founded orders on $\mathbb{N}^n$. Numerous variations exist. Some of these (like parameterizing $>^w_{\Sigma}$ or $>^w_{\ell}$ by an index set $\mathcal{I}$) are implicitly covered because of the lemma below, while others (like demanding a strict decrease in all components specified by an index set $\mathcal{I}$) proved to be impractical.

Intuitively, the order $>^w_{\mathcal{I}}$ is a generalization of $>^w_1$, the order used in [20], where the strict decrease is not necessarily fixed to one specific component; in particular, $>^w_1 = >^w_{\mathcal{I}}$ for $\mathcal{I} = \{1\}$. Moreover, its extension to matrices yields the main order considered in [17]. As to the remaining three orders, two vectors being in relation means that there is a strict decrease in the lengths of the vectors with respect to the Manhattan, Euclidean or maximum norm, respectively [35]. The relationship between these orders is described in the following lemma.

Lemma B.3. *Let $\mathcal{I}$, $\mathcal{J}$ and $\mathcal{K}$ be non-empty index sets such that $\mathcal{I} = \{1, \ldots, n\}$ and $\mathcal{J} \subseteq \mathcal{K} \subseteq \mathcal{I}$. Then the following statements hold:*

1. *$>^w_{\mathcal{J}} \subseteq >^w_{\mathcal{K}}$ and $>^w_1 = >^w_{\{1\}}$,*
2. *$>^w_1$ and $>^w_{\mathrm{m}}$ are incomparable for $n \geqslant 2$, identical otherwise,*
3. *$>^w_{\mathrm{m}} \subset >^w_{\mathcal{I}} = >^w_{\Sigma} = >^w_{\ell}$ for $n \geqslant 2$, all identical otherwise, and*
4. *$>^w_{\mathcal{I}}$ is the strict part of $\geqslant^w$.*

Proof.

1. By definition of the respective orders.

2. Clearly, if $n = 1$, then $>^w_1 \, = \, >^w_\mathrm{m} \, = \, >_\mathbb{N}$. Concerning the first claim, for $\vec{x} = (1, 1, \ldots, 1)^\mathrm{T}$ and $\vec{y} = (0, 1, \ldots, 1)^\mathrm{T}$, we have $\vec{x} >^w_1 \vec{y}$ but not $\vec{x} >^w_\mathrm{m} \vec{y}$. Similarly, for $\vec{x} = (0, 1, \ldots, 1)^\mathrm{T}$ and $\vec{y} = 0$, we have $\vec{x} >^w_\mathrm{m} \vec{y}$ but not $\vec{x} >^w_1 \vec{y}$.

3. For $n = 1$, $>^w_\mathrm{m} \, = \, >^w_\mathcal{I} \, = \, >^w_\Sigma \, = \, >^w_\ell \, = \, >_\mathbb{N}$ is an immediate consequence of the respective definitions. To show $>^w_\mathrm{m} \, \subset \, >^w_\mathcal{I}$ for $n \geqslant 2$, let us assume that $(x_1, \ldots, x_n)^\mathrm{T} >^w_\mathrm{m} (y_1, \ldots, y_n)^\mathrm{T}$, meaning that there is a weak decrease in every single component and an index $j \in \{1, \ldots, n\}$ such that $x_j = \max_i x_i >_\mathbb{N} \max_i y_i$. In particular, we have $x_j >_\mathbb{N} y_j$, which shows that $(x_1, \ldots, x_n)^\mathrm{T} >^w_\mathcal{I} (y_1, \ldots, y_n)^\mathrm{T}$. Further, the inclusion is strict since $(1, 1, \ldots, 1)^\mathrm{T} >^w_\mathcal{I} (0, 1, \ldots, 1)^\mathrm{T}$ but not $(1, 1, \ldots, 1)^\mathrm{T} >^w_\mathrm{m} (0, 1, \ldots, 1)^\mathrm{T}$. Finally, we show $>^w_\mathcal{I} \, = \, >^w_\Sigma \, = \, >^w_\ell$ by showing

$$\exists\, j \in \mathcal{I} \colon x_j >_\mathbb{N} y_j \iff \sum_{i=1}^n x_i >_\mathbb{N} \sum_{i=1}^n y_i \iff \sum_{i=1}^n x_i^2 >_\mathbb{N} \sum_{i=1}^n y_i^2$$

   under the assumption that $x_i \geqslant_\mathbb{N} y_i$ for all $i \in \{1, \ldots, n\}$. Yet this is obvious after rewriting the latter expression to the equivalent expression

$$\exists\, j \in \mathcal{I} \colon x_j >_\mathbb{N} y_j \iff \sum_{i=1}^n x_i - y_i >_\mathbb{N} 0 \iff \sum_{i=1}^n (x_i + y_i)(x_i - y_i) >_\mathbb{N} 0$$

4. Clearly, if $\vec{x} \geqslant^w \vec{y}$ and $\vec{x} \neq \vec{y}$, then there exists a component x_j such that $x_j >_\mathbb{N} y_j$, which shows that $\vec{x} >^w_\mathcal{I} \vec{y}$. The converse is obvious. □

The last item of Lemma B.3 gives rise to the following important corollary.

Corollary B.4. *Let $>$ be a binary relation on $\mathbb{N}^n$ that is weakly decreasing. Then the following statements hold:*

1. If $>$ is irreflexive, then $> \, \subseteq \, >^w_\mathcal{I}$ for $\mathcal{I} = \{1, \ldots, n\}$.

2. The relation $>$ is well-founded if and only if it is irreflexive.

Proof. The first item is a direct consequence of the last item of Lemma B.3. As to the second item, we observe that irreflexivity is obviously necessary for well-foundedness, and that sufficiency follows from the first item and the fact that $>^w_\mathcal{I}$ is well-founded. □

So any irreflexive and weakly decreasing binary relation on $\mathbb{N}^n$, hence any weakly decreasing order, is well-founded, and any well-founded weakly decreasing relation $>$ on $\mathbb{N}^n$ is subsumed by the order $>^w_\mathcal{I}$ for $\mathcal{I} = \{1, \ldots, n\}$, i.e., $> \, \subseteq \, >^w_\mathcal{I}$.

Corollary B.5. *For $\mathcal{I} = \{1, \ldots, n\}$, $>^w_\mathcal{I}$ is the most general of the well-founded weakly decreasing relations on $\mathbb{N}^n$ (in the sense that it subsumes any other such relation).* □

B.2.2 Non-weakly Decreasing Orders

Taking a closer look at Definition B.1, one observes that weak decreasingness is not the essential property for obtaining well-founded relations (resp. orders) on vectors of natural numbers. In fact, the last three orders remain well-founded orders on $\mathbb{N}^n$ even after dropping this property. We denote the corresponding orders by $>_\Sigma$, $>_\ell$ and $>_\mathrm{m}$, respectively. Concerning $>^w_\mathcal{I}$, one must be careful when dropping weak decreasingness because the resulting relation $>_\mathcal{I}$ is well-founded if and only if the index set $\mathcal{I}$ is a singleton set, in which case $>_\mathcal{I}$ is also an order. In the remainder of this appendix, this is implicitly assumed whenever we refer to $>_\mathcal{I}$. Finally, we note that all four orders coincide if $n = 1$, all being equal to $>_\mathbb{N}$. For $n \geqslant 2$, however, all these orders are pairwise incomparable (for all singleton sets $\mathcal{I}$).

B.3 Matrix Interpretations and Weakly Decreasing Orders

In this section, we take the orders introduced in Definition B.1 and build matrix interpretations on top of them. According to Lemma B.3 (items 1 and 3), we only have to consider the family of orders $(>^w_\mathcal{I})_\mathcal{I}$ parameterized by some non-empty index set $\mathcal{I} \subseteq \{1, \ldots, n\}$ and $>^w_\mathrm{m}$, the order induced by the maximum norm. We shall see, however, that the latter kind of matrix interpretation is subsumed by an instance of the former.

Before we can go about formally defining matrix interpretations over $>^w_\mathcal{I}$ and $>^w_\mathrm{m}$, we have to have an understanding of when a linear function is monotone (in the sense of Definition 1.2) with respect to the relations $\geqslant^w$ and $>^w_\mathcal{I}$ ($>^w_\mathrm{m}$). We consider linear functions of the form $f(\vec{x}_1, \ldots, \vec{x}_k) = \sum_{i=1}^k F_i\vec{x}_i + \vec{f}$, where $\vec{f} \in \mathbb{N}^n$ and $F_i \in \mathbb{N}^{n\times n}$ for all $i \in \{1, \ldots, k\}$. Obviously, all such functions are monotone with respect to $\geqslant^w$. Concerning monotonicity with respect to $>^w_\mathcal{I}$, we give necessary and sufficient conditions in the lemma below. A similar lemma, showing sufficiency of the conditions, appeared in [17].

Lemma B.6. *Let $\mathcal{I} \subseteq \{1, \ldots, n\}$ be a non-empty index set. The function $f(\vec{x}_1, \ldots, \vec{x}_k) = \sum_{i=1}^k F_i\vec{x}_i + \vec{f}$ is monotone with respect to $>^w_\mathcal{I}$ if and only if for each $(F_i)_\mathcal{I}$, $i = 1, \ldots, k$, all column sums are at least one.*

Proof. Let $\vec{x}_1, \ldots, \vec{x}_k$ and $\vec{y}$ be arbitrary vectors in $\mathbb{N}^n$ such that $\vec{x}_i >^w_\mathcal{I} \vec{y}$ for some $i \in \{1, \ldots, k\}$. Then there exist a vector $\vec{d} \in \mathbb{N}^n$ and an index $j \in \mathcal{I}$ such that $\vec{x}_i = \vec{y} + \vec{d}$ and $d_j >_\mathbb{N} 0$. Now $f(\ldots, \vec{x}_i, \ldots) >^w_\mathcal{I} f(\ldots, \vec{y}, \ldots)$ holds if and only if $F_i\,\vec{x}_i >^w_\mathcal{I} F_i\,\vec{y}$, which is equivalent to $F_i\,\vec{d} >^w_\mathcal{I} 0$. If all column sums of $(F_i)_\mathcal{I}$ are at least one, then we have $(F_i)_{-j} >^w_\mathcal{I} 0$, which yields $F_i\,\vec{d} >^w_\mathcal{I} 0$ since $F_i\,\vec{d} \geqslant^w (F_i)_{-j} \cdot d_j \geqslant^w (F_i)_{-j}$.

Conversely, if $(F_i)_\mathcal{I}$ has a zero column, then let $j' \in \mathcal{I}$ denote the index of the column of F_i it originates from, and let $\vec{x}_i$ be zero everywhere except for its j'-th component, which we set to one. Then $\vec{x}_i >^w_\mathcal{I} 0$ but $f(\ldots, 0, \vec{x}_i, 0, \ldots) = (F_i)_{-j'} + \vec{f} \not>^w_\mathcal{I} \vec{f} = f(0, \ldots, 0)$. □

We are now ready to formally define matrix interpretations over instances of $>^w_{\mathcal{I}}$ (cf. also the E-compatible matrix interpretations of [17]).

Definition B.7. Let $\mathcal{F}$ denote a signature and $\mathcal{I} \subseteq \{1, \dots, n\}$ a non-empty index set. A *matrix interpretation* $\mathcal{M}_{>^w_{\mathcal{I}}}$ *over* $>^w_{\mathcal{I}}$ of *dimension* $n \in \mathbb{N} \setminus \{0\}$ is an $\mathcal{F}$-algebra with carrier $\mathbb{N}^n$ together with the orders $>^w_{\mathcal{I}}$ and $\geqslant^w$ on $\mathbb{N}^n$, where each k-ary function symbol $f \in \mathcal{F}$ is interpreted by a function of the shape

$$f_{\mathcal{M}} \colon (\mathbb{N}^n)^k \to \mathbb{N}^n, (\vec{x}_1, \dots, \vec{x}_k) \mapsto \sum_{i=1}^{k} F_i \vec{x}_i + \vec{f}$$

with $\vec{f} \in \mathbb{N}^n$ and $F_i \in \mathbb{N}^{n \times n}$ for all $i \in \{1, \dots, k\}$. If all interpretation functions are monotone with respect to $>^w_{\mathcal{I}}$, then $\mathcal{M}_{>^w_{\mathcal{I}}}$ is said to be *monotone*.

One easily verifies that the triple $(\mathcal{M}_{>^w_{\mathcal{I}}}, >^w_{\mathcal{I}}, \geqslant^w)$ constitutes a weakly monotone $\mathcal{F}$-algebra. In case $\mathcal{M}_{>^w_{\mathcal{I}}}$ is monotone, all requirements of an extended monotone $\mathcal{F}$-algebra are satisfied (cf. Definition 1.4). Moreover, the traditional notion of matrix interpretations of Endrullis *et al.* [20] is included in Definition B.7 by choosing the special index set $\mathcal{I} = \{1\}$. Leveraging the results given for monotone algebras in Section 1.3, we note that termination of a TRS can either be shown *directly* by a compatible monotone matrix interpretation (cf. Theorem 1.6 and Corollary 1.7) or *incrementally* by a sequence of monotone matrix interpretations, each of which removes some rewrite rules until eventually all rewrite rules have been removed. Alternatively, matrix interpretations can be applied in the context of the DP framework, where the algebras induced by them are only required to be weakly monotone. In any case, one should be able to check (weak) compatibility of a matrix interpretation with a given set of rewrite rules. The following well-known lemma is helpful for this purpose.

Lemma B.8. *Let $\mathcal{M}$ be an $\mathcal{F}$-algebra with carrier $\mathbb{N}^n$ as in Definition B.7 and t a term with $\mathcal{V}\mathrm{ar}(t) = \{x_1, \dots, x_m\} \subseteq \mathcal{V}$. Then there exist matrices $T_1, \dots, T_m \in \mathbb{N}^{n \times n}$ and a vector $\vec{t} \in \mathbb{N}^n$ such that $[\alpha]_{\mathcal{M}}(t) = T_1 \alpha(x_1) + \cdots + T_m \alpha(x_m) + \vec{t}$ for any variable assignment $\alpha \colon \mathcal{V} \to \mathbb{N}^n$.* □

Hence, the compatibility checks $[\alpha]_{\mathcal{M}}(\ell) >^w_{\mathcal{I}} [\alpha]_{\mathcal{M}}(r)$ and $[\alpha]_{\mathcal{M}}(\ell) \geqslant^w [\alpha]_{\mathcal{M}}(r)$ associated with some rewrite rule $\ell \to r$ boil down to the comparison of such linear functions, which is decidable according to the next lemma. Here, $\geqslant$ denotes the componentwise partial order on $\mathbb{N}^{n \times n}$ induced by $\geqslant_{\mathbb{N}}$.

Lemma B.9. *Let $L_1, \dots, L_m$, $R_1, \dots, R_m$ and $\vec{\ell}$, $\vec{r}$ correspond to a rewrite rule $\ell \to r$ as in Lemma B.8 (with $\mathcal{V}\mathrm{ar}(\ell) = \{x_1, \dots, x_m\} \supseteq \mathcal{V}\mathrm{ar}(r)$). Then, for $\rhd \in \{>^w_{\mathcal{I}}, \geqslant^w\}$, $[\alpha]_{\mathcal{M}}(\ell) \rhd [\alpha]_{\mathcal{M}}(r)$ for all variable assignments $\alpha \colon \mathcal{V} \to \mathbb{N}^n$ if and only if $\vec{\ell} \rhd \vec{r}$ and $L_i \geqslant R_i$ for all $i \in \{1, \dots, m\}$.*

Proof. We prove the lemma for $\rhd = >^w_{\mathcal{I}}$. The proof for $\geqslant^w$ is similar. First, we note that $[\alpha]_{\mathcal{M}}(\ell) >^w_{\mathcal{I}} [\alpha]_{\mathcal{M}}(r)$ holds for all $\alpha \colon \mathcal{V} \to \mathbb{N}^n$ if and only if

$$\sum_{i=1}^{m} (L_i - R_i)\vec{x}_i + (\vec{\ell} - \vec{r}) >^w_{\mathcal{I}} 0 \text{ for all } \vec{x}_1, \dots, \vec{x}_m \in \mathbb{N}^n \tag{B.1}$$

If $\vec{\ell} >^w_{\mathcal{I}} \vec{r}$ and $L_i \geqslant R_i$ for all $i \in \{1, \dots, m\}$, then (B.1) obviously holds because all entries of $L_i - R_i$ are non-negative and $\vec{\ell} - \vec{r} >^w_{\mathcal{I}} 0$. Conversely, (B.1) implies $\vec{\ell} >^w_{\mathcal{I}} \vec{r}$ for $\vec{x}_1 = \cdots = \vec{x}_m = 0$ as well as $L_i \geqslant R_i$ for all $i \in \{1, \dots, m\}$ because if there were some $s \in \{1, \dots, m\}$ with a negative entry in $L_s - R_s$, say $(L_s - R_s)_{jk} < 0$, then choosing $\vec{x}_i = 0$ whenever $i \neq s$ and $\vec{x}_s$ to be zero everywhere except for its k-th component $(\vec{x}_s)_k$ would cause a negative entry in $\sum_{i=1}^{m}(L_i - R_i)\vec{x}_i + (\vec{\ell} - \vec{r})$ for $(\vec{x}_s)_k$ chosen large enough, which would contradict (B.1). □

We close this section with the treatment of matrix interpretations over $>^w_{\mathrm{m}}$. In particular, we show that they are subsumed by the instance of matrix interpretations over $>^w_{\mathcal{I}}$ one obtains by choosing $\mathcal{I} = \{1, \dots, n\}$, which is assumed to be the case in the rest of this section. According to Lemma B.3, we have $>^w_{\mathrm{m}} \subseteq >^w_{\mathcal{I}}$ for all dimensions $n \geqslant 1$. This directly implies that the same inclusion also holds for non-monotone matrix interpretations based on these two orders. For monotone matrix interpretations, the following issue has to be taken into account. If the monotonicity conditions with respect to $>^w_{\mathcal{I}}$ are more strict than the ones for $>^w_{\mathrm{m}}$, then the set of potential interpretation functions is smaller, and it is therefore very well conceivable that the inclusion on the base orders does not propagate to the notions of matrix interpretations built on top of them. However, this is not the case for the two orders considered here.

Lemma B.10. *Let $f(\vec{x}_1, \dots, \vec{x}_k) = \sum_{i=1}^{k} F_i\vec{x}_i + \vec{f}$ with $F_1, \dots, F_k \in \mathbb{N}^{n \times n}$ and $\vec{f} \in \mathbb{N}^n$. Then monotonicity of f with respect to $>^w_{\mathrm{m}}$ implies monotonicity with respect to $>^w_{\mathcal{I}}$.*

Proof. This can be shown using contraposition. Assume that f is not monotone with respect to $>^w_{\mathcal{I}}$. According to Lemma B.6 this means that (at least) one of its matrices has a zero column. Without loss of generality, let the j-th column of some F_i, $i \in \{1, \dots, k\}$, be a zero column and let $\vec{x}_i$ be zero everywhere except for its j-th component. Then $\vec{x}_i >^w_{\mathrm{m}} 0$ but $f(\dots, 0, \vec{x}_i, 0, \dots) = \vec{f} \not>^w_{\mathrm{m}} \vec{f} = f(0, \dots, 0)$, i.e., f is not monotone with respect to $>^w_{\mathrm{m}}$. □

Hence, if $\mathcal{M}_{>^w_{\mathrm{m}}}$ is a matrix interpretation over $>^w_{\mathrm{m}}$, consisting of a set of interpretation functions that are monotone with respect to $\geqslant^w$, then the same functions together with $>^w_{\mathcal{I}}$ constitute a matrix interpretation $\mathcal{M}_{>^w_{\mathcal{I}}}$ over $>^w_{\mathcal{I}}$ that is monotone whenever $\mathcal{M}_{>^w_{\mathrm{m}}}$ is monotone according to Lemma B.10, and that is (weakly) compatible with all rules which are (weakly) compatible with $\mathcal{M}_{>^w_{\mathrm{m}}}$ due the inclusion $>^w_{\mathrm{m}} \subseteq >^w_{\mathcal{I}}$. Thus, matrix interpretations over $>^w_{\mathrm{m}}$ are subsumed by matrix interpretations over $>^w_{\mathcal{I}}$ for $\mathcal{I} = \{1, \dots, n\}$.

B.4 Comparison

After the discussion in Section B.3 the family of orders $(>^w_{\mathcal{I}})_{\mathcal{I}}$ parameterized by some non-empty index set $\mathcal{I} \subseteq \{1, \dots, n\}$ remains as a potentially interesting foundation for matrix interpretations. It includes the traditional order $>^w_1$ as well as $>^w_{\Sigma} = >^w_{\ell}$, the most general of the weakly decreasing orders on $\mathbb{N}^n$ (cf. Lemma B.3 and Corollary B.5). Now the purpose of this section is to compare the resulting variants of matrix interpretations to each other and thus also to the traditional approach.

First, we remark that we do not have to consider all possible index sets since matrix interpretations are invariant under permutations. For example, matrix interpretations over $>^w_{\{1\}}$ are equivalent (with respect to termination proving power) to matrix interpretations over $>^w_{\{j\}}$, where $j \in \{2, \dots, n\}$. The relevant property is that there is a strict decrease in a single fixed vector component, it is not important which component. All that matters is the cardinality of the index set $\mathcal{I}$. Hence, for n-dimensional matrix interpretations, we are left with n different index sets, and, without loss of generality, we can restrict to the sets $\mathcal{I}_d = \{1, \dots, d\}$ for $d = 1, 2, \dots, n$. By definition, the following inclusions hold: $>^w_{\mathcal{I}_1} \subset >^w_{\mathcal{I}_2} \subset \cdots \subset >^w_{\mathcal{I}_n}$. This implies that the same inclusions also hold for non-monotone matrix interpretations based on these orders because if $\mathcal{M}_{>^w_{\mathcal{I}}}$ is a matrix interpretation over $>^w_{\mathcal{I}}$, then the same interpretation functions also form a matrix interpretation $\mathcal{M}_{>^w_{\mathcal{J}}}$ over $>^w_{\mathcal{J}}$, for $\mathcal{I} \subset \mathcal{J}$, that is (weakly) compatible with all rules which are (weakly) compatible with $\mathcal{M}_{>^w_{\mathcal{I}}}$ due the inclusion $>^w_{\mathcal{I}} \subseteq >^w_{\mathcal{J}}$. However, as explained at the end of the previous section, we cannot immediately extend this conclusion to monotone matrix interpretations based on these orders because monotonicity of a function with respect to $>^w_{\mathcal{I}}$ does not imply monotonicity with respect to $>^w_{\mathcal{J}}$ according to Lemma B.6. Moreover, taking the matrix dimension into account, the situation turns out to be a bit more intricate.

Example B.11. Consider the TRS $\mathcal{S}_1 = \{\, \mathsf{f}(\mathsf{a}) \to \mathsf{f}(\mathsf{g}(\mathsf{a})), \mathsf{g}(\mathsf{b}) \to \mathsf{g}(\mathsf{f}(\mathsf{b}))\,\}$. Termination of this system can be established with the following compatible and monotone 2-dimensional matrix interpretation over $>^w_{\{1,2\}}$:

$$\mathsf{f}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \vec{x} \qquad \mathsf{g}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \vec{x} \qquad \mathsf{a}_{\mathcal{M}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \mathsf{b}_{\mathcal{M}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

However, one can show that there is no compatible monotone matrix interpretation over $>^w_{\{1\}}$ of dimension two. Similarly, termination of the TRS $\mathcal{S}_2$

$$\begin{array}{lll} \mathsf{f}(\mathsf{g}(x)) \to \mathsf{f}(\mathsf{a}(\mathsf{g}(\mathsf{g}(\mathsf{f}(x))), \mathsf{g}(\mathsf{g}(\mathsf{f}(x))))) & \mathsf{a}(x, x) \to \mathsf{h}(x) & \mathsf{f}(x) \to x \\ \mathsf{h}(\mathsf{h}(x)) \to \mathsf{c}(\mathsf{h}(x)) & \mathsf{c}(x) \to x & \mathsf{g}(x) \to x \end{array}$$

can be established via the following compatible and monotone 2-dimensional matrix interpretation over $>^w_{\{1\}}$

$$\mathsf{a}_{\mathcal{M}}(\vec{x}, \vec{y}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \vec{y} + \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

$$\mathsf{c}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \mathsf{f}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\mathsf{g}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 \\ 3 \end{pmatrix} \qquad \mathsf{h}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

but there is no compatible monotone matrix interpretation over $>^w_{\{1,2\}}$ of dimension two.

The bottom line of this example is that if we fix the dimension, then matrix interpretations over $>^w_{\{1\}}$ are incomparable to matrix interpretations over $>^w_{\{1,2\}}$. (We are not aware of a general construction that works for any dimension.) However, without this restriction the situation is different. That is to say that for dimension three, for example, there is a compatible monotone matrix interpretation over $>^w_{\{1\}}$ for the TRS $\mathcal{S}_1$:

$$\mathsf{f}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \vec{x} \quad \mathsf{g}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \vec{x} \quad \mathsf{a}_{\mathcal{M}} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \mathsf{b}_{\mathcal{M}} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

Likewise, there is a compatible monotone matrix interpretation over $>^w_{\{1,2\}}$ of dimension three for the TRS $\mathcal{S}_2$:

$$\mathsf{a}_{\mathcal{M}}(\vec{x}, \vec{y}) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{y} + \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

$$\mathsf{c}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathsf{f}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\mathsf{g}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix} \vec{x} + \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \quad \mathsf{h}_{\mathcal{M}}(\vec{x}) = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Indeed, that is no coincidence, as will be shown in the remainder of this section. In particular, we shall see that, when considering monotone interpretations, for any index set $\mathcal{I}$ there exists a larger index set $\mathcal{J}$ such that matrix interpretations over $>^w_{\mathcal{J}}$ are at least as powerful (with respect to proving termination) as matrix interpretations over $>^w_{\mathcal{I}}$ if one does not impose a restriction on the dimension of the matrices. Moreover, under some conditions (e.g. for string rewriting), the various instances of matrix interpretations over $>^w_{\mathcal{I}}$ are all equivalent with respect to termination proving power. For this purpose, we introduce a couple of transformations on matrix interpretations.

As to the first transformation, let $P \in \mathbb{N}^{n \times n}$ be a non-singular matrix and $\mathcal{M}$ some matrix interpretation consisting of a collection of interpretation functions $\{f_{\mathcal{M}}\}_{f \in \mathcal{F}}$ such that each k-ary function symbol f in the signature is interpreted by a function $f_{\mathcal{M}}(\vec{x}_1, \ldots, \vec{x}_k) = \sum_{i=1}^{k} F_i \vec{x}_i + \vec{f}$, where $\vec{f} \in \mathbb{N}^n$ and $F_i \in \mathbb{N}^{n \times n}$ for all $i \in \{1, \ldots, k\}$. Then we associate with $\mathcal{M}$ a matrix interpretation $\Phi_P(\mathcal{M})$, where each k-ary function symbol f is interpreted by a function $f_{\Phi_P(\mathcal{M})}(\vec{x}_1, \ldots, \vec{x}_k) = \sum_{i=1}^{k} P F_i P^{-1} \vec{x}_i + P \vec{f}$.

Remark B.12. Note that in general PF_iP^{-1} is not a non-negative matrix, even if P and F_i are non-negative. As we need this property in our context, we must be careful when applying this transformation, unless P happens to be a (generalized) permutation matrix (for which P^{-1} is always non-negative).

According to Lemma B.8, the interpretation of a term t with respect to $\mathcal{M}$ and a variable assignment α can be written as $[\alpha]_{\mathcal{M}}(t) = T_1\alpha(x_1) + \cdots + T_m\alpha(x_m) + \vec{t}$. By construction of $\Phi_P(\mathcal{M})$, we obtain the following lemma.

Lemma B.13. *Let $T_1, \ldots, T_m$ and $\vec{t}$ correspond to a term t as described in Lemma B.8. Then $[\alpha]_{\Phi_P(\mathcal{M})}(t) = PT_1P^{-1}\alpha(x_1) + \cdots + PT_mP^{-1}\alpha(x_m) + P\vec{t}$ for any assignment α.*

Proof. By induction on the structure of t. □

Corollary B.14. *For every ground term t, $[\alpha]_{\Phi_P(\mathcal{M})}(t) = P \cdot [\alpha]_{\mathcal{M}}(t)$.* □

Our next transformation associates with an n-dimensional matrix interpretation $\mathcal{M}$ (as above) an $(n+1)$-dimensional matrix interpretation $\Psi(\mathcal{M})$, where each k-ary function symbol f is interpreted by a function $f_{\Psi(\mathcal{M})}(\vec{x}_1, \ldots, \vec{x}_k) = \sum_{i=1}^{k} F'_i \vec{x}_i + \vec{f}'$ such that for all $i \in \{1, \ldots, k\}$,

$$\vec{f}' = \begin{pmatrix} 0 \\ \vec{f} \end{pmatrix} \text{ and } F'_i = \begin{pmatrix} f_i & 0 \\ 0 & F_i \end{pmatrix} \text{ for some } f_i \in \mathbb{N} \setminus \{0\}$$

(Here, the numbers f_i are fresh, that is, not related to the components of $\vec{f}$.) Moreover, we associate with the matrix interpretation $\mathcal{M}$ (resp. $\Psi(\mathcal{M})$) a linear polynomial interpretation $\mathcal{P}(\mathcal{M})$, where each k-ary function symbol f is interpreted by a linear polynomial $f_{\mathcal{P}(\mathcal{M})}(x_1, \ldots, x_k) = \sum_{i=1}^{k} f_i x_i$ (with the f_i's of $\Psi(\mathcal{M})$).

Lemma B.15. *Let t be an arbitrary term. Then for all variable assignments $\alpha \colon \mathcal{V} \to \mathbb{N}^n$ and $\beta \colon \mathcal{V} \to \mathbb{N}$,*

$$[\gamma]_{\Psi(\mathcal{M})}(t) = \begin{pmatrix} [\beta]_{\mathcal{P}(\mathcal{M})}(t) \\ [\alpha]_{\mathcal{M}}(t) \end{pmatrix}$$

for the assignment $\gamma \colon \mathcal{V} \to \mathbb{N}^{n+1}$, $x \mapsto \begin{pmatrix} \beta(x) \\ \alpha(x) \end{pmatrix}$.

Proof. By induction on the structure of t. □

Corollary B.16. *For every ground term t, $[\gamma]_{\Psi(\mathcal{M})}(t) = \begin{pmatrix} 0 \\ [\alpha]_{\mathcal{M}}(t) \end{pmatrix}$.* □

Again, by Lemma B.8, we can write $[\alpha]_{\mathcal{M}}(t) = T_1\alpha(x_1) + \cdots + T_m\alpha(x_m) + \vec{t}$. Likewise, the interpretation of t with respect to $\mathcal{P}(\mathcal{M})$ and some variable assignment β can be written as $[\beta]_{\mathcal{P}(\mathcal{M})}(t) = t_1\beta(x_1) + \cdots + t_m\beta(x_m)$, where $t_1, \ldots, t_m \in \mathbb{N}$. Plugging these expressions into Lemma B.15, we obtain the following lemma.

Lemma B.17. *Let $T_1, \ldots, T_m$, $t_1, \ldots, t_m$ and $\vec{t}$ correspond to a term t as described above. Then, in the situation of Lemma B.15,*

$$[\gamma]_{\Psi(\mathcal{M})}(t) = \sum_{i=1}^{m} \begin{pmatrix} t_i & 0 \\ 0 & T_i \end{pmatrix} \gamma(x_i) + \begin{pmatrix} 0 \\ \vec{t} \end{pmatrix}$$

Proof.

$$\sum_{i=1}^{m} \begin{pmatrix} t_i & 0 \\ 0 & T_i \end{pmatrix} \gamma(x_i) + \begin{pmatrix} 0 \\ \vec{t} \end{pmatrix} = \sum_{i=1}^{m} \begin{pmatrix} t_i & 0 \\ 0 & T_i \end{pmatrix} \begin{pmatrix} \beta(x_i) \\ \alpha(x_i) \end{pmatrix} + \begin{pmatrix} 0 \\ \vec{t} \end{pmatrix} =$$

$$\sum_{i=1}^{m} \begin{pmatrix} t_i\beta(x_i) \\ T_i\alpha(x_i) \end{pmatrix} + \begin{pmatrix} 0 \\ \vec{t} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{m} t_i\beta(x_i) \\ \sum_{i=1}^{m} T_i\alpha(x_i) + \vec{t} \end{pmatrix} = \begin{pmatrix} [\beta]_{\mathcal{P}(\mathcal{M})}(t) \\ [\alpha]_{\mathcal{M}}(t) \end{pmatrix} = [\gamma]_{\Psi(\mathcal{M})}(t)$$

□

If all the f_i's introduced by $\Psi(\mathcal{M})$ are one, then each t_i in $[\beta]_{\mathcal{P}(\mathcal{M})}(t)$ corresponds to the number of occurrences of the associated variable x_i.

Lemma B.18. *Let t be an arbitrary term with $\mathcal{V}\mathrm{ar}(t) = \{x_1, \dots, x_m\}$, and let all interpretation functions in $\mathcal{P}(\mathcal{M})$ have the shape $f_{\mathcal{P}(\mathcal{M})}(x_1, \dots, x_k) = \sum_{i=1}^{k} x_i$ (for each k-ary function symbol f). Then $[\beta]_{\mathcal{P}(\mathcal{M})}(t) = \sum_{i=1}^{m} |t|_{x_i}\beta(x_i)$ for any variable assignment β.*

Proof. By induction on the structure of t. □

We are now ready to present the main results of this section comparing monotone matrix interpretations over various instances of $>^w_{\mathcal{I}}$. In what follows, for a given TRS $\mathcal{R}$, $\mathcal{R} \subseteq \geqslant_{\mathcal{P}(\mathcal{M})}$ abbreviates $[\beta]_{\mathcal{P}(\mathcal{M})}(\ell) \geqslant_{\mathbb{N}} [\beta]_{\mathcal{P}(\mathcal{M})}(r)$ for all variable assignments $\beta\colon \mathcal{V} \to \mathbb{N}$ and all rewrite rules $\ell \to r \in \mathcal{R}$.

Lemma B.19. *Let $\mathcal{M}$ be a monotone n-dimensional matrix interpretation over $>^w_{\mathcal{I}}$, where $\mathcal{I} \subseteq \{1, \dots, n\}$, and let $\mathcal{R}$ be a TRS satisfying $\mathcal{R} \subseteq \geqslant_{\mathcal{P}(\mathcal{M})}$. Then (weak) compatibility of $\mathcal{R}$ with $\mathcal{M}$ implies (weak) compatibility with a monotone $(n{+}1)$-dimensional matrix interpretation over $>^w_{\mathcal{J}}$, where $|\mathcal{J}| = |\mathcal{I}|{+}1$, $\mathcal{J} \subseteq \{1, \dots, n+1\}$.*

Proof. Assuming that $\mathcal{M}$ is (weakly) compatible with $\mathcal{R}$, we show that $\Psi(\mathcal{M})$ is (weakly) compatible as well. To this end, we let $\mathcal{J} = \{1\} \cup \{\, x+1 \mid x \in \mathcal{I} \,\}$ and reason as follows. By assumption, all interpretation functions of $\mathcal{M}$ are monotone with respect to $>^w_{\mathcal{I}}$, that is, for each matrix $M \in \mathcal{M}$, all column sums of $(M)_{\mathcal{I}}$ are at least one according to Lemma B.6. By construction of $\Psi(\mathcal{M})$, this implies that for each matrix $M' \in \Psi(\mathcal{M})$, all column sums of $(M')_{\mathcal{J}}$ are also at least one. Hence, all interpretation functions of $\Psi(\mathcal{M})$ are monotone with respect to $>^w_{\mathcal{J}}$. As to compatibility of $\Psi(\mathcal{M})$ with $\mathcal{R}$, for any rule $\ell \to r$, $[\gamma]_{\Psi(\mathcal{M})}(\ell) >^w_{\mathcal{J}} [\gamma]_{\Psi(\mathcal{M})}(r)$ holds for all variable assignments γ if and only if

$$\begin{pmatrix} [\beta]_{\mathcal{P}(\mathcal{M})}(\ell) \\ [\alpha]_{\mathcal{M}}(\ell) \end{pmatrix} >^w_{\mathcal{J}} \begin{pmatrix} [\beta]_{\mathcal{P}(\mathcal{M})}(r) \\ [\alpha]_{\mathcal{M}}(r) \end{pmatrix}$$

holds for all variable assignments α and β (cf. Lemma B.15). By definition of $>^w_{\mathcal{J}}$, it remains to show that there is a weak decrease in every single component and a strict decrease in some component with index $j \in \mathcal{J}$. By compatibility of $\mathcal{M}$ with $\mathcal{R}$, we have $[\alpha]_{\mathcal{M}}(\ell) >^w_{\mathcal{I}} [\alpha]_{\mathcal{M}}(r)$ for all assignments α, which immediately establishes the latter requirement and, together with the assumption $\mathcal{R} \subseteq \geqslant_{\mathcal{P}(\mathcal{M})}$, also the former. Similarly, weak compatibility of $\mathcal{M}$ with $\ell \to r$, asserting that $[\alpha]_\mathcal{M}(\ell) \geqslant^w [\alpha]_\mathcal{M}(r)$ for all assignments α, and the fact that $\mathcal{R} \subseteq \geqslant_{\mathcal{P}(\mathcal{M})}$ imply

$$\begin{pmatrix}[\beta]_{\mathcal{P}(\mathcal{M})}(\ell) \\ [\alpha]_\mathcal{M}(\ell)\end{pmatrix} \geqslant^w \begin{pmatrix}[\beta]_{\mathcal{P}(\mathcal{M})}(r) \\ [\alpha]_\mathcal{M}(r)\end{pmatrix}$$

for all variable assignments α and β, which shows that $\Psi(\mathcal{M})$ is weakly compatible with $\ell \to r$. □

Using Lemma B.18 we can replace the semantic condition $\mathcal{R} \subseteq \geqslant_{\mathcal{P}(\mathcal{M})}$ by a (more familiar) syntactic condition.

Corollary B.20. *Let $\mathcal{R}$ be a non-duplicating TRS. Then (weak) compatibility of $\mathcal{R}$ with a monotone n-dimensional matrix interpretation over $>^w_\mathcal{I}$, where $\mathcal{I} \subseteq \{1,\dots,n\}$, implies (weak) compatibility with a monotone $(n+1)$-dimensional matrix interpretation over $>^w_\mathcal{J}$, where $|\mathcal{J}| = |\mathcal{I}| + 1$, $\mathcal{J} \subseteq \{1,\dots,n+1\}$.*

Proof. Setting all the f_i's introduced by $\Psi(\mathcal{M})$ to one, the condition $\mathcal{R} \subseteq \geqslant_{\mathcal{P}(\mathcal{M})}$ becomes equivalent to $\mathcal{R}$ being non-duplicating according to Lemma B.18. □

Example B.21. Consider the TRS $\mathcal{R} = \{\, \mathsf{f}(x) \to \mathsf{g}(\mathsf{h}(x,x)), \mathsf{g}(\mathsf{a}) \to \mathsf{f}(\mathsf{a}) \,\}$. Termination can be established via the following compatible and monotone 2-dimensional matrix interpretation over $>^w_{\{1\}}$:

$$\mathsf{f}_\mathcal{M}(\vec{x}) = \begin{pmatrix}3 & 0\\ 2 & 0\end{pmatrix}\vec{x} + \begin{pmatrix}2\\0\end{pmatrix} \qquad \mathsf{g}_\mathcal{M}(\vec{x}) = \begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}\vec{x} + \begin{pmatrix}1\\0\end{pmatrix}$$
$$\mathsf{h}_\mathcal{M}(\vec{x},\vec{y}) = \begin{pmatrix}2 & 0\\ 0 & 0\end{pmatrix}\vec{x} + \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\vec{y} \qquad \mathsf{a}_\mathcal{M} = \begin{pmatrix}0\\3\end{pmatrix}$$

Moreover, the following linear polynomial interpretation orients all rules weakly:

$$\mathsf{f}_{\mathcal{P}(\mathcal{M})}(x) = 2x \qquad \mathsf{g}_{\mathcal{P}(\mathcal{M})}(x) = x \qquad \mathsf{h}_{\mathcal{P}(\mathcal{M})}(x,y) = x + y \qquad \mathsf{a}_{\mathcal{P}(\mathcal{M})} = 0$$

Hence, by (the proof of) Lemma B.19, there exists a compatible monotone matrix interpretation over $>^w_{\{1,2\}}$ of dimension three.

Next we show that one can get rid of the precondition $\mathcal{R} \subseteq \geqslant_{\mathcal{P}(\mathcal{M})}$ in Lemma B.19 by doubling the dimension of $\mathcal{M}$.

Lemma B.22. *Let $\mathcal{R}$ be a TRS. Then (weak) compatibility of $\mathcal{R}$ with a monotone n-dimensional matrix interpretation over $>^w_\mathcal{I}$, $\mathcal{I} \subseteq \{1,\dots,n\}$, implies (weak) compatibility with a monotone $2n$-dimensional matrix interpretation over $>^w_\mathcal{J}$, where $|\mathcal{J}| = 2 \cdot |\mathcal{I}|$, $\mathcal{J} \subseteq \{1,\dots,2n\}$.*

In order to prove Lemma B.22, we introduce a construction that combines two matrix interpretations $\mathcal{M}$ and $\mathcal{N}$ (not necessarily of the same dimension) to a matrix interpretation $\Pi(\mathcal{M},\mathcal{N})$ as follows. Assuming $\mathcal{M}$ consists of interpretation functions $f_\mathcal{M}(\vec{x}_1,\dots,\vec{x}_k) = \sum_{i=1}^k F_i\vec{x}_i + \vec{f}$ and $\mathcal{N}$ of interpretation functions $f_\mathcal{N}(\vec{x}_1,\dots,\vec{x}_k) = \sum_{i=1}^k \widetilde{F}_i\vec{x}_i + \tilde{f}$, the $\Pi(\mathcal{M},\mathcal{N})$-interpretation of each k-ary function symbol f is

$$f_{\Pi(\mathcal{M},\mathcal{N})}(\vec{x}_1,\dots,\vec{x}_k) = \sum_{i=1}^k \begin{pmatrix}F_i & 0\\ 0 & \widetilde{F}_i\end{pmatrix}\vec{x}_i + \begin{pmatrix}\vec{f}\\ \tilde{f}\end{pmatrix}$$

Lemma B.23. *Let t be an arbitrary term. Then for all variable assignments α and β,*

$$[\gamma]_{\Pi(\mathcal{M},\mathcal{N})}(t) = \begin{pmatrix} [\alpha]_{\mathcal{M}}(t) \\ [\beta]_{\mathcal{N}}(t) \end{pmatrix} \text{ for the variable assignment } \gamma(x) = \begin{pmatrix} \alpha(x) \\ \beta(x) \end{pmatrix}.$$

Proof. By induction on the structure of t. □

Proof of Lemma B.22. Assuming that $\mathcal{M}$ is an n-dimensional matrix interpretation over $>^w_{\mathcal{I}}$ (weakly) compatible with $\mathcal{R}$, we show that $\Pi(\mathcal{M},\mathcal{M})$ is (weakly) compatible as well. To this end, we let $\mathcal{J} = \mathcal{I} \cup \{\, x + n \mid x \in \mathcal{I} \,\}$ and reason as follows. By assumption, all interpretation functions of $\mathcal{M}$ are monotone with respect to $>^w_{\mathcal{I}}$, that is, for each matrix $M \in \mathcal{M}$, all column sums of $(M)_{\mathcal{I}}$ are at least one according to Lemma B.6. By construction of $\Pi(\mathcal{M},\mathcal{M})$, this implies that for each matrix $M' \in \Pi(\mathcal{M},\mathcal{M})$, all column sums of $(M')_{\mathcal{J}}$ are also at least one. Hence, all interpretation functions of $\Pi(\mathcal{M},\mathcal{M})$ are monotone with respect to $>^w_{\mathcal{J}}$. As to compatibility of $\Pi(\mathcal{M},\mathcal{M})$ with $\mathcal{R}$, for any rewrite rule $\ell \to r$, $[\gamma]_{\Pi(\mathcal{M},\mathcal{M})}(\ell) >^w_{\mathcal{J}} [\gamma]_{\Pi(\mathcal{M},\mathcal{M})}(r)$ holds for all variable assignments γ if and only if

$$\begin{pmatrix} [\alpha]_{\mathcal{M}}(\ell) \\ [\beta]_{\mathcal{M}}(\ell) \end{pmatrix} >^w_{\mathcal{J}} \begin{pmatrix} [\alpha]_{\mathcal{M}}(r) \\ [\beta]_{\mathcal{M}}(r) \end{pmatrix}$$

holds for all variable assignments α and β (cf. Lemma B.23). But this follows directly from compatibility of $\mathcal{M}$ with $\mathcal{R}$ since $[\alpha]_{\mathcal{M}}(\ell) >^w_{\mathcal{I}} [\alpha]_{\mathcal{M}}(r)$ for all assignments α. Similarly, weak compatibility of $\mathcal{M}$ with $\mathcal{R}$ implies weak compatibility of $\Pi(\mathcal{M},\mathcal{M})$ with $\mathcal{R}$. □

Summarizing the above results, every TRS that can be proved terminating (directly or incrementally) using matrix interpretations over $>^w_{\mathcal{I}}$, for some index set $\mathcal{I}$, can also be proved terminating using matrix interpretations over $>^w_{\mathcal{J}}$, for a larger index set $\mathcal{J}$, at the expense of an increased dimension.

Next we elaborate on the converse of this statement. To this end, let us consider some TRS $\mathcal{R}$ and a (weakly) compatible monotone matrix interpretation $\mathcal{M}$ over $>^w_{\mathcal{I}}$ of dimension n, where $|\mathcal{I}| > 1$, consisting of interpretation functions $f_{\mathcal{M}}(\vec{x}_1, \ldots, \vec{x}_k) = \sum_{i=1}^{k} F_i \vec{x}_i + \vec{f}$ for each k-ary function symbol f in the signature. Our aim is to show that $\mathcal{R}$ is also (weakly) compatible with a monotone matrix interpretation over $>^w_{\{1\}}$ (or more generally, $>^w_{\mathcal{I}}$ for a singleton index set $\mathcal{I}$), albeit with a higher dimension.

First, we transform $\mathcal{M}$ into $\mathcal{M}' := \Psi(\mathcal{M})$, which is in turn transformed into $\mathcal{M}'' := \Phi_P(\mathcal{M}')$ for $P = I + U$, where I is the identity matrix and U is all zero except for the entries $U_{1,i+1} = 1$, for all $i \in \mathcal{I}$. As $P^{-1} = I - U$ is not non-negative, we have to ensure well-definedness of $\mathcal{M}''$, that is, make sure that all its matrices are non-negative. Now for any matrix (vector) M, PM is equal to M except for the first row, which is the sum of the rows of M with indices in $\{1\} \cup \{\, i + 1 \mid i \in \mathcal{I} \,\}$. Hence,

$$P \begin{pmatrix} f_i & 0 \\ 0 & F_i \end{pmatrix} = \begin{pmatrix} f_i & \sum_{c \in \mathcal{I}} (F_i)_{c1} & \cdots & \sum_{c \in \mathcal{I}} (F_i)_{cn} \\ 0 & (F_i)_{-1} & \cdots & (F_i)_{-n} \end{pmatrix}$$

Multiplying this matrix by P^{-1} from the right has the effect of subtracting its first column from the columns with indices in $\{ i+1 \mid i \in \mathcal{I} \}$, thus replacing $\sum_{c\in\mathcal{I}}(F_i)_{cj}$ by $\sum_{c\in\mathcal{I}}(F_i)_{cj} - f_i$ for all indices $j \in \mathcal{I}$ in the above representation. As these are the only entries that may eventually be negative, $\sum_{c\in\mathcal{I}}(F_i)_{cj} - f_i \geqslant 0$ for all $j \in \mathcal{I}$ implies well-definedness of $\mathcal{M}''$. Note, however, that if all the f_i's introduced by the transformation Ψ are one, then the latter condition is satisfied without further ado because, by assumption, all interpretation functions of $\mathcal{M}$ are monotone with respect to $>^w_{\mathcal{I}}$; hence, for all $j \in \mathcal{I}$, $\sum_{c\in\mathcal{I}}(F_i)_{cj}$ is at least one according to Lemma B.6. Moreover, note that the top-left entry of each matrix occurring in $\mathcal{M}''$ is positive since $f_i > 0$. Consequently, all interpretation functions of $\mathcal{M}''$ are monotone with respect to $>^w_{\{1\}}$.

As to (weak) compatibility of $\mathcal{M}''$ with $\mathcal{R}$, let $\ell \to r \in \mathcal{R}$ be an arbitrary rule in the variables $x_1, \dots, x_m$, let α be some variable assignment, and let $[\alpha]_{\mathcal{M}}(\ell) = L_1\alpha(x_1) + \cdots + L_m\alpha(x_m) + \vec{\ell}$ and $[\alpha]_{\mathcal{M}}(r) = R_1\alpha(x_1) + \cdots + R_m\alpha(x_m) + \vec{r}$. Likewise, let $[\beta]_{\mathcal{P}(\mathcal{M})}(\ell) = l_1\beta(x_1) + \cdots + l_m\beta(x_m)$, where $l_1, \dots, l_m \in \mathbb{N}$, and similarly for $[\beta]_{\mathcal{P}(\mathcal{M})}(r)$. By (weak) compatibility of $\mathcal{M}$, we have $\vec{\ell} >^w_{\mathcal{I}} \vec{r}$ ($\vec{\ell} \geqslant^w \vec{r}$) and $L_i \geqslant R_i$ for $i = 1, \dots, m$ (cf. Lemma B.9). Moreover, by Lemmata B.13 and B.17,

$$[\gamma]_{\mathcal{M}''}(\ell) = \sum_{i=1}^{m} P \begin{pmatrix} l_i & 0 \\ 0 & L_i \end{pmatrix} P^{-1}\gamma(x_i) + P \begin{pmatrix} 0 \\ \vec{\ell} \end{pmatrix} \quad \text{for } \gamma\colon \mathcal{V} \to \mathbb{N}^{n+1},\ x \mapsto \begin{pmatrix} \beta(x) \\ \alpha(x) \end{pmatrix}$$

Therefore, for $\triangleright \in \{>^w_{\{1\}}, \geqslant^w\}$, $[\gamma]_{\mathcal{M}''}(\ell) \triangleright [\gamma]_{\mathcal{M}''}(r)$ holds for all assignments γ if and only if

$$P \begin{pmatrix} 0 \\ \vec{\ell} \end{pmatrix} \triangleright P \begin{pmatrix} 0 \\ \vec{r} \end{pmatrix} \quad \text{and} \quad P \begin{pmatrix} l_i - r_i & 0 \\ 0 & L_i - R_i \end{pmatrix} P^{-1} \geqslant 0$$

for $i = 1, \dots, m$. The first condition follows directly from $\vec{\ell} >^w_{\mathcal{I}} \vec{r}$ ($\vec{\ell} \geqslant^w \vec{r}$) and the shape of P. Concerning the second condition, we first rewrite the corresponding matrix to

$$\begin{pmatrix} l_i - r_i & \sum_{c\in\mathcal{I}}(L_i - R_i)_{c1} & \cdots & \sum_{c\in\mathcal{I}}(L_i - R_i)_{cn} \\ 0 & (L_i - R_i)_{-1} & \cdots & (L_i - R_i)_{-n} \end{pmatrix} P^{-1}$$

Using the fact that $L_i \geqslant R_i$, the entire matrix is non-negative if and only if for all $i \in \{1, \dots, m\}$, $l_i \geqslant r_i$ and $\sum_{c\in\mathcal{I}}(L_i - R_i)_{cj} \geqslant l_i - r_i$ for all $j \in \mathcal{I}$.

Based on these observations, we establish the following lemma.

Lemma B.24. *Let $\mathcal{M}$ be a monotone n-dimensional matrix interpretation over $>^w_{\mathcal{I}}$, where $|\mathcal{I}| > 1$ for $\mathcal{I} \subseteq \{1, \dots, n\}$, and let $\mathcal{R}$ be a TRS satisfying $\mathcal{R} \subseteq \geqslant_{\mathcal{P}(\mathcal{M})}$. Moreover, assume that for each k-ary function symbol f, all column sums of each $(F_i)_{\mathcal{I}}$ are greater than or equal to f_i for all $i \in \{1, \dots, k\}$, and that for each rule $\ell \to r \in \mathcal{R}$, all column sums of each $(L_i - R_i)_{\mathcal{I}}$ are greater than or equal to $l_i - r_i$ for all $i \in \{1, \dots, m\}$. Then (weak) compatibility of $\mathcal{R}$ with $\mathcal{M}$ implies (weak) compatibility with a monotone $(n+1)$-dimensional matrix interpretation over $>^w_{\mathcal{J}}$, where $|\mathcal{J}| = 1$ for $\mathcal{J} \subseteq \{1, \dots, n+1\}$.*

Proof. By the reasoning presented above and by the fact that $\mathcal{R} \subseteq \geqslant_{\mathcal{P}(\mathcal{M})}$ holds if and only if for each rule $\ell \to r \in \mathcal{R}$, $l_i \geqslant r_i$ for all $i \in \{1, \ldots, m\}$, together with the observation that the constraint $\forall\, j \in \mathcal{I}\colon \sum_{c \in \mathcal{I}} (F_i)_{cj} - f_i \geqslant 0$ is equivalent to all column sums of $(F_i)_{\mathcal{I}}$ being greater than or equal to f_i, and likewise for the constraint $\forall\, j \in \mathcal{I}\colon \sum_{c \in \mathcal{I}} (L_i - R_i)_{cj} \geqslant l_i - r_i$. $\square$

Corollary B.25. *Let $\mathcal{M}$ be a monotone n-dimensional matrix interpretation over $>^w_{\mathcal{I}}$, where $|\mathcal{I}| > 1$ for $\mathcal{I} \subseteq \{1, \ldots, n\}$, and let $\mathcal{R}$ be a non-duplicating TRS. Moreover, assume that for each $\ell \to r \in \mathcal{R}$, all column sums of each $(L_i - R_i)_{\mathcal{I}}$ are greater than or equal to $l_i - r_i$ for all $i \in \{1, \ldots, m\}$. Then (weak) compatibility of $\mathcal{R}$ with $\mathcal{M}$ implies (weak) compatibility with a monotone $(n+1)$-dimensional matrix interpretation over $>^w_{\mathcal{J}}$, where $|\mathcal{J}| = 1$ for $\mathcal{J} \subseteq \{1, \ldots, n+1\}$.*

Proof. Setting all the f_i's introduced by $\Psi(\mathcal{M})$ to one, the condition $\mathcal{R} \subseteq \geqslant_{\mathcal{P}(\mathcal{M})}$ becomes equivalent to $\mathcal{R}$ being non-duplicating according to Lemma B.18. Moreover, all column sums of each $(F_i)_{\mathcal{I}}$ are greater than or equal to $f_i = 1$ because all interpretation functions of $\mathcal{M}$ are monotone with respect to $>^w_{\mathcal{I}}$. $\square$

By restricting the class of non-duplicating TRSs further, we can get rid of the condition that all column sums of $(L_i - R_i)_{\mathcal{I}}$ are greater than or equal to $l_i - r_i$.

Corollary B.26. *Let $\mathcal{R}$ be a TRS such that for all $\ell \to r \in \mathcal{R}$, $|\ell|_x = |r|_x$ for all variables x. Then (weak) compatibility of $\mathcal{R}$ with a monotone n-dimensional matrix interpretation over $>^w_{\mathcal{I}}$, where $|\mathcal{I}| > 1$ for $\mathcal{I} \subseteq \{1, \ldots, n\}$, implies (weak) compatibility with a monotone $(n+1)$-dimensional matrix interpretation over $>^w_{\mathcal{J}}$, where $|\mathcal{J}| = 1$ for $\mathcal{J} \subseteq \{1, \ldots, n+1\}$.*

Proof. For each rewrite rule $\ell \to r \in \mathcal{R}$, $l_i - r_i = 0$ as $l_i = |\ell|_{x_i} = |r|_{x_i} = r_i$. $\square$

Together with Corollary B.20, Corollary B.26 shows that for string rewriting the various instances of matrix interpretations over $>^w_{\mathcal{I}}$ are all equivalent with respect to termination proving power if there is no restriction on the dimension of the matrices.

B.5 Matrix Interpretations and Non-weakly Decreasing Orders

In this section, we investigate the usefulness of the orders $>_\Sigma$, $>_\ell$, $>_{\mathrm{m}}$ and $>_{\mathcal{I}}$ (where $\mathcal{I}$ is a singleton set) introduced in Subsection B.2.2 for building *monotone* matrix interpretations on top of them. As these orders originated from the orders introduced in Definition B.1 by dropping the property of weak decreasingness, each of them obviously subsumes its ancestor, e.g., $>^w_\Sigma \subset >_\Sigma$, so that one is tempted to believe that these more general base orders would induce more powerful kinds of matrix interpretations. However, as already mentioned at the end of Section B.3, an inclusion like $>^w_\Sigma \subset >_\Sigma$ does not necessarily propagate to the corresponding notions of matrix interpretations because of the monotonicity requirement all interpretation functions have to satisfy. Indeed, it turns out that the monotonicity conditions with respect to $>_\Sigma$, $>_\ell$, $>_{\mathrm{m}}$ and $>_{\mathcal{I}}$ are much stronger than the ones associated with their respective weakly decreasing counterparts, ultimately resulting in weaker notions of matrix interpretations. In particular, we will see that monotone matrix interpretations over $>_{\mathcal{I}}$ and $>_{\Sigma}$ are equivalent to linear polynomial interpretations.

As already mentioned in Subsection B.2.2, all four orders are equal to $>_{\mathbb{N}}$ when the dimension n is one. Hence, matrix interpretations based on them are at least as powerful as linear polynomial interpretations. Next we show that matrix interpretations over $>_{\mathcal{I}}$ and $>_{\Sigma}$ are no more powerful than linear polynomial interpretations if monotonicity with respect to $>_{\mathcal{I}}$ and $>_{\Sigma}$ is required. Since matrix interpretations are invariant under permutations, we consider the index set $\mathcal{I} = \{1\}$ without loss of any generality.

Lemma B.27. *Let $f(\vec{x}_1, \ldots, \vec{x}_k) = \sum_{i=1}^{k} F_i \vec{x}_i + \vec{f}$ with $F_1, \ldots, F_k \in \mathbb{N}^{n \times n}$ and $\vec{f} \in \mathbb{N}^n$. Then $f(\vec{x}_1, \ldots, \vec{x}_k)$ is monotone with respect to $>_{\{1\}}$ if and only if for each F_i, $i = 1, \ldots, k$, $(F_i)_{11} \geqslant 1$ and $(F_i)_{12} = \cdots = (F_i)_{1n} = 0$.*

Proof. Let $\vec{x}_1, \ldots, \vec{x}_k$ and $\vec{y}$ be arbitrary vectors in $\mathbb{N}^n$ such that $\vec{x}_i >_{\{1\}} \vec{y}$ for some argument position $i \in \{1, \ldots, k\}$. Then $f(\ldots, \vec{x}_i, \ldots) >_{\{1\}} f(\ldots, \vec{y}, \ldots)$ holds if and only if $F_i \, \vec{x}_i >_{\{1\}} F_i \, \vec{y}$, which is equivalent to $F_i \, (\vec{x}_i - \vec{y}) >_{\{1\}} 0$. By definition of $>_{\{1\}}$, this holds if and only if the first component of $F_i \, (\vec{x}_i - \vec{y})$ is positive, that is, $\sum_{j=1}^{n} (F_i)_{1j} \, (\vec{x}_i - \vec{y})_j >_{\mathbb{N}} 0$ (for all $\vec{x}_i$ and $\vec{y} \in \mathbb{N}^n$). Obviously, the conditions $(F_i)_{11} \geqslant 1$ and $(F_i)_{12} = \cdots = (F_i)_{1n} = 0$ are sufficient for this. However, they are also necessary because if $(F_i)_{11} = 0$, then we have $\vec{x}_i >_{\{1\}} \vec{y}$ for $\vec{x}_i = (1, 0, \ldots, 0)^{\mathrm{T}}$ and $\vec{y} = 0$ but $\sum_{j=1}^{n} (F_i)_{1j} \, (\vec{x}_i - \vec{y})_j = 0 \not>_{\mathbb{N}} 0$. Similarly, if $(F_i)_{1j'} > 0$ for some $j' \in \{2, \ldots, n\}$, then let $\vec{x}_i = (1, 0, \ldots, 0)^{\mathrm{T}}$ and let $\vec{y}$ be zero everywhere except for its j'-th component, which we set to $(F_i)_{11}$. Again, we have $\vec{x}_i >_{\{1\}} \vec{y}$ but $\sum_{j=1}^{n} (F_i)_{1j} \, (\vec{x}_i - \vec{y})_j = (F_i)_{11}(1 - (F_i)_{1j'}) \not>_{\mathbb{N}} 0$. □

Intuitively, this means that the first component of a function application $f(\vec{x}_1, \ldots, \vec{x}_k)$ only depends on the respective first components of its arguments, not on the other components. Based on this observation and the fact that for comparisons with $>_{\{1\}}$ only the first components matter, we associate the following linear polynomial interpretation $\mathcal{P}$ with a given matrix interpretation $\mathcal{M}$ over $>_{\{1\}}$. For each k-ary function symbol f, if $f_{\mathcal{M}}(\vec{x}_1, \ldots, \vec{x}_k) = \sum_{i=1}^{k} F_i \vec{x}_i + \vec{f}$ is its interpretation in $\mathcal{M}$, with all matrices satisfying the conditions of Lemma B.27, then we define its $\mathcal{P}$-interpretation as $f_{\mathcal{P}}(x_1, \ldots, x_k) = \sum_{i=1}^{k} (F_i)_{11} \, x_i + (\vec{f})_1$, which is monotone with respect to $>_{\mathbb{N}}$ and $\geqslant_{\mathbb{N}}$ since $(F_i)_{11} \geqslant 1$. (So $(\mathcal{P}, >_{\mathbb{N}}, \geqslant_{\mathbb{N}})$ is an extended monotone algebra.) By construction, the $\mathcal{P}$-interpretation of an arbitrary term coincides with the first component of its $\mathcal{M}$-interpretation. The straightforward induction proof is omitted.

Lemma B.28. *Let $\mathcal{M}$ be a monotone matrix interpretation over $>_{\{1\}}$ of dimension n, $\mathcal{P}$ the associated linear polynomial interpretation as described above and t an arbitrary term. Then for any variable assignment $\alpha \colon \mathcal{V} \to \mathbb{N}^n$, $\pi_1([\alpha]_{\mathcal{M}}(t)) = [\pi_1 \circ \alpha]_{\mathcal{P}}(t)$, where π_1 projects a vector to its first component.* □

Therefore, any rewrite rule $\ell \to r$ that is (weakly) compatible with $\mathcal{M}$, is also (weakly) compatible with $\mathcal{P}$ (since $\pi_1 \circ \alpha$ covers all assignments $\mathcal{V} \to \mathbb{N}$), which shows that monotone matrix interpretations over $>_{\{1\}}$ ($>_{\mathcal{I}}$) are no more powerful than linear polynomial interpretations. The following lemma states that this is also the case for monotone matrix interpretations over $>_{\Sigma}$.

Lemma B.29. *Let $f(\vec{x}_1, \dots, \vec{x}_k) = \sum_{i=1}^{k} F_i \vec{x}_i + \vec{f}$ with $F_1, \dots, F_k \in \mathbb{N}^{n \times n}$ and $\vec{f} \in \mathbb{N}^n$. Then $f(\vec{x}_1, \dots, \vec{x}_k)$ is monotone with respect to $>_{\Sigma}$ if and only if for each F_i, $i = 1, \dots, k$, all column sums are equal and at least one.*

Proof. According to Lemma B.3, we have $>^w_{\Sigma} = >^w_{\mathcal{I}}$ for $\mathcal{I} = \{1, \dots, n\}$. Due to $>^w_{\Sigma} \subseteq >_{\Sigma}$ and the monotonicity of f with respect to $\geqslant^w$, monotonicity of f with respect to $>_{\Sigma}$ implies monotonicity with respect to $>^w_{\Sigma}$. Hence, the condition that all column sums are at least one is a necessary condition according to Lemma B.6. Next we show that equality of all column sums is necessary, too. Assume to the contrary that some matrix F_i, $i \in \{1, \dots, k\}$, has two columns with a differing column sum. Without loss of generality, let j and j' be the corresponding column indices such that the sum s_j of $(F_i)_{-j}$ is greater than the sum $s_{j'}$ of $(F_i)_{-j'}$. Then let $\vec{x}_i$ be zero everywhere except for its j'-th component, which we set to s_j, and let $\vec{y}$ be zero everywhere except for its j-th component, which we set to $s_{j'}$. As a consequence, we have $\vec{x}_i >_{\Sigma} \vec{y}$ but $f(\dots, 0, \vec{x}_i, 0, \dots) \not>_{\Sigma} f(\dots, 0, \vec{y}, 0, \dots)$.

Sufficiency of the conditions follows by the following chain of reasoning. Assuming that the column sums of each matrix are equal and at least one, $\vec{x}_i >_{\Sigma} \vec{y}$ implies $F_i\, \vec{x}_i >_{\Sigma} F_i\, \vec{y}$ because for any vector $\vec{v} = (v_1, \dots, v_n)^{\mathrm{T}} \in \mathbb{N}^n$,

$$\sum_{j=1}^{n} (F_i\, \vec{v})_j = v_1 \sum_{j=1}^{n} (F_i)_{j1} + \dots + v_n \sum_{j=1}^{n} (F_i)_{jn} = \left(\sum_{j=1}^{n} v_j \right) \cdot \left(\sum_{j=1}^{n} (F_i)_{j1} \right)$$

Finally, $f(\dots, \vec{x}_i, \dots) >_{\Sigma} f(\dots, \vec{y}, \dots)$ follows from $F_i\, \vec{x}_i >_{\Sigma} F_i\, \vec{y}$ by adding equal vectors to both sides of the inequality. □

In analogy to the treatment of matrix interpretations over $>_{\{1\}}$, given an n-dimensional matrix interpretation $\mathcal{M}$ over $>_{\Sigma}$, we again associate a linear polynomial interpretation $\mathcal{P}$ with $\mathcal{M}$ as follows. For each k-ary function symbol f, if $f_{\mathcal{M}}(\vec{x}_1, \dots, \vec{x}_k) = \sum_{i=1}^{k} F_i \vec{x}_i + \vec{f}$ is its interpretation in $\mathcal{M}$, with all matrices satisfying the conditions of Lemma B.29, then its $\mathcal{P}$-interpretation is defined as $f_{\mathcal{P}}(x_1, \dots, x_k) = \sum_{i=1}^{k} F_i^{\Sigma} x_i + \sum_{j=1}^{n} (\vec{f})_j$, where F_i^{Σ} denotes the column sum of F_i, which is equal for all columns of F_i and at least one. (So $(\mathcal{P}, >_{\mathbb{N}}, \geqslant_{\mathbb{N}})$ is an extended monotone algebra.) By construction, the $\mathcal{P}$-interpretation of an arbitrary term coincides with the sum of the components of its $\mathcal{M}$-interpretation. (Again, the straightforward induction proof is omitted.)

Lemma B.30. *Let $\mathcal{M}$ be a monotone matrix interpretation over $>_{\Sigma}$ of dimension n, $\mathcal{P}$ the associated linear polynomial interpretation as described above, and t an arbitrary term. Then for any variable assignment $\alpha \colon \mathcal{V} \to \mathbb{N}^n$, $\sum_{j=1}^{n} ([\alpha]_{\mathcal{M}}(t))_j = [\alpha']_{\mathcal{P}}(t)$, where $\alpha'(x) = \sum_{j=1}^{n} (\alpha(x))_j$ for all $x \in \mathcal{V}$.* □

As a consequence, if a rewrite rule $\ell \to r$ is (weakly) compatible with $\mathcal{M}$, then it is also (weakly) compatible with $\mathcal{P}$ (since α' covers all assignments $\mathcal{V} \to \mathbb{N}$), which shows that monotone matrix interpretations over $>_\Sigma$ are no more powerful than linear polynomial interpretations.

Finally, concerning monotone matrix interpretations over $>_{\mathrm{m}}$ and $>_\ell$, the situation is similar as for $>_\Sigma$ and $>_{\mathcal{I}}$. That is to say that the respective monotonicity conditions are too strong, thus reducing the set of potential interpretation functions down to a size that renders matrix interpretations over $>_{\mathrm{m}}$ and $>_\ell$ useless. For example, one can show that for monotonicity of a function $A\vec{x} + \vec{b}$ with respect to $>_{\mathrm{m}}$, it is necessary that the matrix A satisfies the conditions of Lemma B.29, that is, all column sums of A must be equal and at least one; e.g. by considering vectors $\vec{x}$ and $\vec{y}$ such that all components of $\vec{y}$ are equal to some $y \in \mathbb{N}$ and $\vec{x}$ is zero everywhere except for its j-th component, $j \in \{1, \ldots, n\}$, which contains the value $y + 1$. Similarly, one can show that for monotonicity of $A\vec{x} + \vec{b}$ with respect to $>_\ell$, it is necessary that all column vectors of A are non-zero and have the same (Euclidean) length; e.g., for dimension two and higher, by considering vectors $\vec{x} = (y \mp 1, y \pm 1, 0, \ldots, 0)^{\mathrm{T}}$ and $\vec{y} = (y, y, 0, \ldots, 0)^{\mathrm{T}}$, where $y \in \mathbb{N} \setminus \{0\}$. However, these conditions are not sufficient. Even if A is the identity matrix, $A\vec{x} + \vec{b}$ is not necessarily monotone with respect to $>_\ell$.

B.6 Improved Matrix Interpretations

According to the results presented in Section B.4, when considering monotone matrix interpretations over $>_{\mathcal{I}}^w$, then for any index set $\mathcal{I}$ there exists a larger index set $\mathcal{J}$ such that matrix interpretations over $>_{\mathcal{J}}^w$ subsume matrix interpretations over $>_{\mathcal{I}}^w$ if one does not impose a restriction on the dimension of the matrices. In practice, however, due to computational restrictions the dimension is limited. But then the various instances of matrix interpretations over $>_{\mathcal{I}}^w$ are incomparable as witnessed by Example B.11 and by the experiments we performed. Therefore, an implementation should try all instances (cf. also [17]). Apart from parallelization, one could try to combine the constraints associated with each instance into a single disjunctive constraint and let the constraint solver figure out which instance to pursue. This approach was chosen in [17]. However, according to our experiments, it does not yield an efficient implementation (cf. the experimental results below). Therefore, we propose a different approach, which generalizes traditional matrix interpretations.

Given some signature $\mathcal{F}$, we define an $\mathcal{F}$-algebra $\mathcal{M}$ with carrier $\mathbb{N}^n$, where each k-ary function symbol $f \in \mathcal{F}$ is interpreted by a linear function as in Definition B.7. Concerning monotonicity, we demand that

- all functions are monotone with respect to $>_{\mathcal{I}_1}^w$, or
- all functions are monotone with respect to $>_{\mathcal{I}_2}^w$, or

  $\vdots$
- all functions are monotone with respect to $>_{\mathcal{I}_n}^w$.

(This is meant to be a single disjunctive constraint involving the conditions given in Lemma B.6.) Compatibility with a given TRS $\mathcal{R}$ is established by

Table B.1: Experimental results for various matrix interpretations.

dimension 2	# bits	SCORE
$>^w_{\mathcal{I}_1}$	3	242
$>^w_{\mathcal{I}_2}$	3	247
$>^{\text{ext}}_{\{1\}}$	3	254
[17]	3	250

dimension 3	# bits	SCORE
$>^w_{\mathcal{I}_1}$	2|3	266|285
$>^w_{\mathcal{I}_2}$	2|3	252|264
$>^w_{\mathcal{I}_3}$	2|3	249|269
$>^{\text{ext}}_{\{1\}}$	2|3	276|287
[17]	2|3	267|270

demanding that for each rewrite rule $\ell \to r \in \mathcal{R}$, $[\alpha]_{\mathcal{M}}(\ell) >^w_{\mathcal{I}_1} [\alpha]_{\mathcal{M}}(r)$ for all variable assignments α, i.e., every rewrite rule gives rise to a strict decrease in the first components of the vectors associated with it. Weak compatibility is established in the usual way (using the order $\geqslant^w$).

Clearly, if all interpretation functions of $\mathcal{M}$ are monotone with respect to $>^w_{\mathcal{I}_1}$, then $\mathcal{M}$ corresponds to a traditional matrix interpretation [20]. More generally, $\mathcal{M}$ always is a monotone matrix interpretation over $>^w_{\mathcal{I}_d}$, $d \in \{1, \ldots, n\}$, in the sense of Definition B.7 because of the inclusions $>^w_{\mathcal{I}_1} \subset >^w_{\mathcal{I}_2} \subset \cdots \subset >^w_{\mathcal{I}_n}$.

Next we provide some experimental data. We implemented the variants of matrix interpretations considered in this appendix in the termination prover $\mathsf{T_TT_2}$ [41] and analyzed their performance on TPDB [72] version 7.0.2. All tests have been performed on a laptop equipped with 2 GB of main memory and one dual-core INTEL® Core 2 Duo T7500 processor running at a clock rate of 2.2 GHz with a time limit of 60 seconds per system.[1]

Table B.1 summarizes our results for establishing direct termination (using monotone matrix interpretations that are (strictly) compatible with all rules of a given TRS). We searched for matrix interpretations of dimensions two and three by encoding the constraints as an SMT problem (quantifier-free non-linear arithmetic), which is solved by bit-blasting. The table lists the number of bits used to represent matrix/vector coefficients, the number of bits for intermediate results is one higher than that. The entry $>^{\text{ext}}_{\{1\}}$ in the first column refers to the notion of matrix interpretations presented above, whereas the entry [17] refers to the approach proposed in [17]. For the experiments presented in the table the time limit was hardly ever consumed. Typically, a termination proof is obtained in about 2 (5) seconds for dimension 2 (3). For dimensions 4 and higher, however, there are many more timeouts, resulting in inferior performance scores; e.g., for matrix interpretations over $>^w_{\mathcal{I}_1}$ of dimension 4 (with 3 bits) one loses more than 40 of the 285 systems for dimension 3.

B.7 Conclusion and Future Work

In this appendix, we studied various alternative well-founded orders on vectors of natural numbers based on vector norms. Most of them turned out to be equivalent

[1] For full details see `http://colo6-c703.uibk.ac.at/ttt2/fn/matrix/`.

to or subsumed by an instance of $>^w_{\mathcal{I}}$, an order which already appeared in [17]. In this respect, our main contribution are the theoretical comparisons presented in Section B.4, as well as the variant of matrix interpretations introduced in Section B.6. We do note, however, that the situation is quite different when switching from the natural numbers to the rationals and reals. Then it is not the case anymore that almost all of the orders of Section B.2 (suitably adapted) are equivalent. In particular, one could imagine interpretation functions, all of whose matrices have entries less than one, but which are still monotone. We leave this issue for the near future.

We also plan to investigate the ramifications of the kinds of matrix interpretations proposed in this appendix with respect to recent results on the derivational complexity of TRSs [59] (cf. also Chapter 7). For example, if a matrix has a diagonal of all zeros, then its trace, the sum of the diagonal entries, is also zero. As the trace of a matrix is the sum of its eigenvalues, which have been shown to be the determining factor for the derivational complexity of TRSs, a lower trace might be beneficial in this context.

In the near future work we will address alternative matrix interpretations in the context of the DP framework, where it suffices to consider weakly monotone algebras (cf. Definition 1.4). Based on the results of the previous sections, the following instances of a weakly monotone algebra $(\mathbb{N}^n, >, \gtrsim)$, where $\gtrsim \; = \; \geqslant^w$ and (1) $> \; = \; >^w_{\mathcal{I}}$ for $\mathcal{I} = \{1, \ldots, n\}$, (2) $> \; = \; >_\Sigma$, (3) $> \; = \; >_\mathrm{m}$, or (4) $> \; = \; >_\ell$ need to be considered. As to the first instance, we note that $>$ is the strict part of $\gtrsim$ according to Lemma B.3. Yet this case was already considered in [20], apart from a refinement that reduces the search space in an implementation. Moreover, by Corollary B.5, no other weakly decreasing orders need to be considered for $>$. However, observing that weak decreasingness is not really needed to obtain a weakly monotone algebra, one might as well drop it, thus obtaining a weakly monotone algebra, where $> \; = \; >_\Sigma$ (instance (2) above), which is a proper generalization of the first one since $>^w_{\mathcal{I}} \; = \; >^w_\Sigma \; \subset \; >_\Sigma$ (cf. Lemma B.3). Similarly, one can use the non-weakly decreasing orders $>_\mathrm{m}$ and $>_\ell$ to obtain other instances of weakly monotone algebras. They are all incomparable since the orders $>_\Sigma$, $>_\mathrm{m}$ and $>_\ell$ are incomparable.