hybr2d Theory of Hybrid Systems Informatik 2 | RWTH AACHEN UNIVERSITY

The present work was submitted to the LuFG Theory of Hybrid Systems

BACHELOR OF SCIENCE THESIS

QUANTIFIER ELIMINATION BY CYLINDRICAL ALGEBRAIC DECOMPOSITION

Tom Neuhäuser

Examiners:

Prof. Dr. Erika Ábrahám

Priv.-Doz. Viktor Levandovskyy

Additional Advisor:

Gereon Kremer, M.Sc.

Aachen, 27.09.2018

Abstract

This thesis presents the theory needed to implement a quantifier elimination method for non-linear real arithmetic in order to extend the Satisfiability Modulo Theories Real Arithmetic Toolbox (SMT-RAT) by the capability of quantifier elimination. A quantifier elimination method for non-linear real arithmetic constructs an equivalent quantifier-free formula for a given quantified formula. The described quantifier elimination method is based on the concept of the cylindrical algebraic decomposition (CAD). A cylindrical algebraic decomposition decomposes $\mathbb{R}^n$ into sign-invariant regions with respect to a set of multivariate polynomials. The idea to use cylindrical algebraic decomposition in order to eliminate quantifiers was originally introduced by Collins. Several improvements, proposed by Hong and Brown, are used by the described quantifier elimination method in order to construct preferably simple equivalent quantifier-free formulas. Experimental results for the behavior of the implemented quantifier elimination method on a collection of exemplary quantified formulas are presented.

Acknowledgments

I would like to express my gratitude towards Erika Ábrahám for the opportunity to write my bachelor thesis on the interesting topic of quantifier elimination by cylindrical algebraic decomposition. Furthermore, I want to thank Viktor Levandovskyy for agreeing to be my second examiner. I would like to give my sincere thanks to Gereon Kremer for his continuous help while I was working on this thesis. Finally, I want to thank my family and friends for their support.

Contents

Chapter 1

Introduction

A question which already arose in school, is the question asking under which conditions a quadratic polynomial has a real root. This question can be formalized by the quantified non-linear real arithmetic formula shown below.

$$\exists x \ (a \neq 0 \ \wedge \ ax^2 + bx + c = 0)$$

An answer to this question was provided in school as well. A quadratic polynomial has a real root if, and only if, its discriminant is non-negative. This condition can be formalized by the quantifier-free non-linear real arithmetic formula given below.

$$a \neq 0 \ \wedge \ b^2 - 4ac \geq 0$$

In more general, a quantifier elimination problem for non-linear real arithmetic is a problem where a quantified non-linear real arithmetic formula is given and an equivalent quantifier-free formula needs to be found.

In 1948, Tarski introduced a quantifier elimination method for non-linear real arithmetic [17]. However, the time complexity of Tarski's method causes the impracticability of the method for all but the most trivial quantifier elimination problems. In 1973, Collins discovered a more efficient quantifier elimination method for non-linear real arithmetic based on cylindrical algebraic decomposition [7]. Collins' method has a time complexity doubly exponential in the number of variables and polynomially in the number of polynomials, the maximum degree of the polynomials, the maximum length of the coefficients of the polynomials as well as the number of atomic formulas. In 1988, Davenport and Heintz proved, that quantifier elimination for non-linear real arithmetic is doubly exponential [11]. Several refinements of Collins' method were proposed, *e.g.* by Hong [13] or Brown [2], with an aim to find simpler equivalent quantifier-free formulas.

This thesis presents the theory needed to implement a quantifier elimination method for non-linear real arithmetic in order to extend the Satisfiability Modulo Theories Real Arithmetic Toolbox (SMT-RAT) by the capability of quantifier elimination. To begin with, in chapter 2, the preliminary concepts, such as the cylindrical algebraic decomposition or the non-linear real arithmetic, are introduced. In chapter 3, a quantifier elimination method for non-linear real arithmetic is described, including a method to determine the truth values of a quantified formula, a method to assure the for quantifier elimination important property of projection-definability and a method to construct an equivalent quantifier-free formula. Finally, in chapter 4, some remarks on the implementation are made as well as experimental results are presented.

Chapter 2

Preliminaries

2.1 Non-Linear Real Arithmetic

In 1951, Tarski [17] proved, that for any quantified first-order formula in the signature $(+, \cdot, 0, 1, <)$ there exists an equivalent quantifier-free formula in the same signature. The theory of $(\mathbb{R}, +, \cdot, 0, 1, <)$ is called the theory of the real numbers or the non-linear real arithmetic (NRA). Formulas over the signature $(+, \cdot, 0, 1, <)$ with real-valued variables are called non-linear real arithmetic formulas. The terms of non-linear real arithmetic formulas, build upon $+, \cdot, 0, 1$ and variables in $\mathbb{R}$, are real-valued multivariate polynomials. If p is a multivariate polynomial in n variables $x_1, \ldots, x_n$ and i is the largest index between 1 and n such that the degree of p with respect to x_i is greater than zero, then p will be called an i-level polynomial. The atomic formulas of non-linear real arithmetic formulas are polynomial equations and inequalities. The atomic formulas will also be called constraints. If a non-linear real arithmetic formula has n free variables $x_1, \ldots, x_n$, an interpretation of the variables can be considered a vector $(a_1, \ldots, a_n)$ in $\mathbb{R}^n$. In the following, the set $\mathbb{R}^n$ will also be referred to as n-space and a vector in n-space will also be called a point.

2.2 Cylindrical Algebraic Decompositions

In 1973, Collins [7] introduced a more efficient method to eliminate quantifiers based on cylindrical algebraic decompositions. Therefore, an understanding of the term is crucial in order to examine the method. A detailed overview on cylindrical algebraic decomposition as well as the underlying theory can be found in [15]. A decomposition is a partition with the additional property that each part is a so called region. A region in $\mathbb{R}^n$ is a connected subset of $\mathbb{R}^n$.

Definition 2.2.1. *Let S be a subset of $\mathbb{R}^n$. A decomposition of S is a finite partition into regions.*

The parts of a decomposition are called cells. A decomposition can be algebraic. In order to define this property, the term *semi-algebraic set* is needed. The definition of semi-algebraic sets given below is slightly more general than needed for this work. It deals with so called real closed fields. A real closed field is a field that has the same first-order properties as the field of real numbers. For this work, the only real closed field which will be considered is the field of real numbers.

Definition 2.2.2. *Let $\mathcal{R}$ be a real closed field. A semi-algebraic set is a subset S of $\mathcal{R}^n$ defined by finitely many polynomial equations and inequalities.*

Semi-algebraic sets have an important property, they are closed under set-theoretic projection. A set-theoretic projection maps a set in n-space to a set of some lower dimension m, such that the last $n - m$ components are simply cut off. An interesting property of semi-algebraic sets is, that the set-theoretic projection of a semi-algebraic set in n-space onto m-space is a semi-algebraic set. With that being said, algebraic decompositions can be defined.

Definition 2.2.3. *A decomposition is called algebraic, if every cell is a semi-algebraic set.*

A decomposition can have the property of being cylindrical. The term *cylinder* is defined below.

Definition 2.2.4. *Let A be a region in $\mathbb{R}^n$. The set $A \times \mathbb{R}$ is called the cylinder over A.*

In order to define cylindrical decompositions, the term *stack* is needed. A stack is a decomposition of a cylinder that meets certain criteria. For the purpose of formalizing these criteria, the terms *section* and *sector* need to be defined.

Definition 2.2.5. *Let A be a region in $\mathbb{R}^n$ and let $p \in \mathbb{R}[x_1, \ldots, x_n]$. The set $\{(a, p(a)) \mid a \in A\}$ is called a p-section or just a section if the polynomial is not of interest.*

A sector is the set of elements between two sections. A formal definition is given below.

Definition 2.2.6. *Let A be a region in $\mathbb{R}^n$ and let $p_1, p_2 \in \mathbb{R}[x_1, \ldots, x_n]$. The set $\{(a, b) \mid p_1(a) < b < p_2(a),\ a \in A\}$ is called a (p_1, p_2)-sector or just a sector if the polynomials are not of interest.*

Thereby, the criteria a decomposition has to meet in order to be a stack are given below.

Definition 2.2.7. *Let A be a region in $\mathbb{R}^n$ and let $p_1, \ldots, p_m \in \mathbb{R}[x_1, \ldots, x_n]$, such that for all $a \in A$ it holds that $p_1(a) < \ldots < p_m(a)$. Furthermore, let $p_0 = -\infty$ and $p_{m+1} = +\infty$. A stack over A is a decomposition of the cylinder $A \times \mathbb{R}$ such that the cells are p_i-sections or (p_i, p_{i+1})-sectors.*

With that being said, a cylindrical decomposition can be defined.

Definition 2.2.8. *A decomposition D of $\mathbb{R}^n$ is called cylindrical if the following inductively defined property holds for D.*

$n = 1$ *D is a decomposition of $\mathbb{R}^1$ such that the cells are sections and sectors.*

$n > 1$ *There exists a decomposition D' of $\mathbb{R}^{n-1}$ such that each cell of D is an element of the stack over a cell of D'.*

A decomposition that fulfills both properties will be called a cylindrical algebraic decomposition, as seen in the definition below.

Definition 2.2.9. *A cylindrical algebraic decomposition (CAD) is a decomposition which is algebraic and cylindrical.*

Due to the inductive nature of cylindrical decompositions and the property of semi-algebraic sets being closed under set-theoretic projection, a cylindrical algebraic decomposition of n-space defines so called induced cylindrical algebraic decompositions of lower dimensions. If D' is the cylindrical algebraic decomposition of i-space induced by a cylindrical algebraic decomposition D of n-space, the cells of D' are called i-level cells of D.
For a given set of multivariate polynomials in n variables a cylindrical algebraic decomposition of n-space, such that sign-invariance is achieved, can be computed.

Definition 2.2.10. *Let A be a region in $\mathbb{R}^n$ and let $p \in \mathbb{R}[x_1, \dots, x_n]$. The region A is called sign-invariant with respect to p, if for any $a, b \in A$ the signs of p in a and b agree.*

A cylindrical algebraic decomposition defined by a set of multivariate polynomials is called sign-invariant, if each cell is sign-invariant with respect to each element of the given set of multivariate polynomials. A method to construct a sign-invariant cylindrical algebraic decomposition was originally introduced by Collins [7]. Over the years, several improvements were made, *e.g.* by McCallum [16], Hong [14] or Brown [3]. Rigorous proofs on the correctness of the method can be found in their work, which will be omitted here. A description of the method will suffice to build a solid understanding for the next chapters.

The construction is done in two phases, projection and lifting. In the projection phase, a projection operator is used to map multivariate polynomials to multivariate polynomials of lower level by computing the discriminants, the resultants and the leading coefficients of the polynomials, such that the zero set of the resulting polynomial is the set-theoretic projection of the set of relevant points of the projected polynomials, *e.g.* the point of a vertical tangent for a single projected polynomial or the point of an intersection for two projected polynomials.

Starting from the given set of multivariate polynomials, a projection operator is applied repeatedly to the result of the previous application until the obtained set is closed under projection. The obtained set of multivariate polynomials will be called the projection factor set and its elements will be called projection factors. The projection factor set is finite and can be partitioned by the level of the projection factors. If P is the projection factor set, the set P_i will denote the set of i-level projection factors. The projection factor set needs to be defined, such that an important property is assured. Assume the projection factors of level $i+1$ and above are already computed. Then, the set of i-level projection factors, resulting from the application of a projection operator, describe the maximal regions over which the projection factors of level i and above have a constant number of real roots. The following definition formalizes this property.

Definition 2.2.11. *Let A be a region in $\mathbb{R}^{i-1}$ and let $p_1, \ldots, p_m$ be multivariate polynomials in $\mathbb{R}[x_1, \ldots, x_n]$. The p_i's are called delinable over A, if for any $a \in A$ it holds that*

- *the number of roots of $p_i(a)$ is constant,*
- *the number of different roots of $p_i(a)$ is constant,*
- *the number of common roots of $p_i(a)$ and $p_j(a)$ is constant.*

There exist several projection operators to choose from. The original projection operator was introduced by Collins [7]. Improvements to the original projection operator were made by *e.g* McCallum [16], Hong [14] or Brown [3]. In detail, the four projection operators were examined in [18]. Collins' and Hong's projection operator ensure the correctness of the constructed cylindrical algebraic decomposition, while McCallums's and Brown's projection operators might define a smaller, thus possibly incomplete, projection factor set. However, the usage of Brown's projection operator and the resulting smaller projection factor set are still desirable. The hope is, that a simpler equivalent quantifier-free formula can be constructed using a projection operator producing less projection factors. The possible incomplete projection factor set can be fixed later on, in order to ensure the construction of an actual equivalent quantifier-free formula. Thus, in this work, Brown's projection operator will be considered.

The relevant information for quantifier elimination, a cylindrical algebraic decomposition holds, is the sign of any projection factor in any cell. Thus, for a sign-invariant cylindrical algebraic decomposition, a single point per cell suffices to represent the respective cell. A point that represents a cell will be called a sample point. In the lifting phase, sample points representing a sign-invariant cylindrical algebraic decomposition are constructed successively. Beginning with sample points representing the induced sign-invariant cylindrical algebraic decomposition of 1-space, i-level sample points are extended to sample points for level $i+1$ until a representation of the sign-invariant cylindrical algebraic decomposition of n-space is obtained.

The sectors and sections defined by the 1-level projection factors decompose 1-space. The resulting cylindrical algebraic decomposition is sign-invariant since the sign of an univariate polynomial only changes at its roots. The sample points are chosen as the union of the set of roots, representing the sectors, and a set consisting of an arbitrary intermediate point for each open interval between two roots, representing the sections. Once a sample point representing an i-level cell is already computed, a set of sample points representing the stack over the considered cell can be constructed. The projection factors of level $i+1$ and above are delinable over the considered cell, since the projection operator has to assure this property. Then, the sample point representing the considered i-level cell can be substituted in the projection factors of level $i+1$ to obtain a set of univariate polynomials. The sectors and sections defined by these polynomials decompose the cylinder over the considered cell. The resulting stack is sign-invariant since the induced cylindrical algebraic decomposition of i-space was sign-invariant and because the sign of an univariate polynomial only changes at its roots. The sample points representing the stack over the considered cell are chosen similarly to the case of a decomposition of 1-space. Successively extending each sample point representing the induced cylindrical algebraic decomposition of i-space to a set of sample points of level $i+1$ as described above, results eventually in a representation of the sign-invariant cylindrical algebraic decomposition of n-space.

2.3 Minimal Hitting Sets

In 1992, Hong introduced a refinement of Collins' method which produces simpler equivalent quantifier-free formulas based on minimization [13]. In 1999, Brown took up on Hong's idea but used minimal hitting sets to simplify the constructed equivalent quantifier-free formula in multiple steps for his refinement of the quantifier elimination method [2]. The hitting set problem is one of Karp's 21 NP-complete problems.

Definition 2.3.1. *Let T be a set and $\{S_1, \ldots, S_n\}$ be a collection of subsets of T. Let $k \leq |T|$. The hitting set problem asks if there is a subset H of T, such that $|H| \leq k$ and $H \cap S_i \neq \emptyset$ for all $i = 1, \ldots, n$.*

A set H that satisfies the second property is called a hitting set. The corresponding optimization problem asks for a hitting set H such that k is minimal. A hitting set H, such that k is minimal, is called a minimal hitting set.

Chapter 3

Quantifier Elimination

3.1 Assumptions on the Quantified Formulas

This thesis is about a quantifier elimination method for non-linear real arithmetic formulas and non-linear real arithmetic formulas only. For convenience, non-linear real arithmetic formulas will also be referred to as formulas. The quantified formulas are assumed to be in prenex normal form. A quantified formula ϕ in n variables $x_1, \ldots x_n$, of which the first k variables are free and the remaining $n - k$ variables are quantified, is said to be in prenex normal form if the formula is of the form shown below.

$$Q_1 x_n \; Q_2 x_{n-1} \; \ldots \; Q_{n-k} x_{k+1} \; \phi'(x_1, \ldots, x_n)$$

Where $Q_1, \ldots Q_{n-k} \in \{\exists, \forall\}$ and where the subformula ϕ' is a quantifier-free formula in the n variables $x_1, \ldots, x_n$. In the following, the subformula ϕ' will also be referred to as the quantifier-free part. The considered quantified formulas can be assumed to be in prenex normal form without loss of generality since it can be shown, that for every first-order formula there exists an equivalent first-order formula that is in prenex normal form. As stated in section 2.1, the terms of non-linear real arithmetic formulas are multivariate polynomials and the atomic formulas are polynomial equations and inequalities, called constraints. For two multivariate polynomials p_1 and p_2 the equation $p_1 = p_2$ is equal to the normalized equation $p = 0$ with $p := p_1 - p_2$. An analogous results also holds for inequalities. The constraints in a quantified formula are assumed to only occur normalized. Consider the exemplary quantified formula given below.

$$\phi := \exists y \ (x^2 + y^2 - 1 < 0) \wedge (x + y < 0)$$

The formula has two variables x and y, whereby y is existential quantified and x is free. The quantifier-free part is given as shown below.

$$\phi' := (x^2 + y^2 - 1 < 0) \wedge (x + y < 0)$$

Both the variables x and y occur free in ϕ'. The set of constraints occurring in ϕ is given as $\{x^2 + y^2 - 1 < 0, \ x + y < 0\}$. The corresponding set of multivariate polynomials is given as $P_\phi := \{x^2 + y^2 - 1, \ x + y\}$. The set of multivariate polynomials P_ϕ defines a cylindrical algebraic decomposition of 2-space, which is shown in the figure below. See section 2.2 on cylindrical algebraic decompositions.

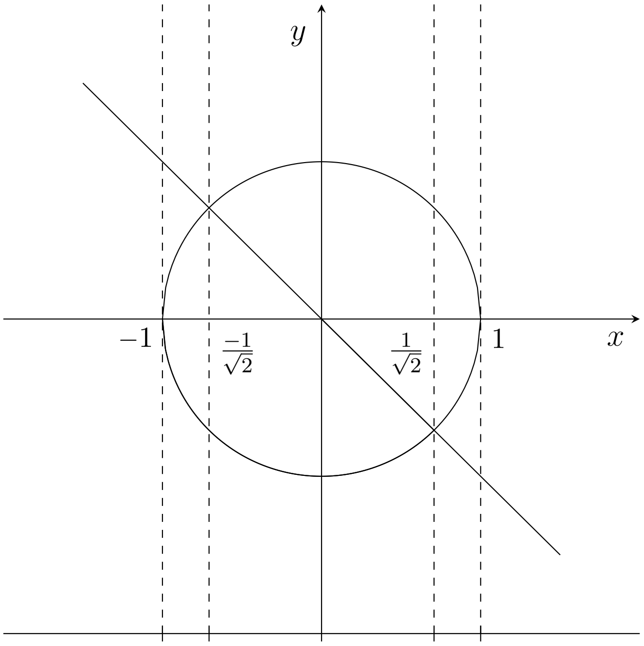

Figure 3.1: the cylindrical algebraic decomposition

The projection factor set computed by Brown's projection operator is $P = \{x^2 + y^2 - 1,\ x + y,\ x^2 - 1,\ 2x^2 - 1\}$. The cylindrical algebraic decomposition has 9 cells of level 1 and 47 cells of level 2. The 1-level cells are represented by the sample points $\{-2,\ -1,\ -{}^{7}\!/\!{}_{8},\ -{}^{1}\!/\!{}_{\sqrt{2}},\ 0,\ {}^{1}\!/\!{}_{\sqrt{2}},\ {}^{7}\!/\!{}_{8},\ 1,\ 2\}$. The enumeration of the set of sample points representing the 2-level cells is omitted.

3.2 Determine Truth Values

In order to eliminate quantifiers, a method to determine the truth values of a quantified formula ϕ depending on the interpretation of the free variables $x_1, \ldots, x_k$ needs to be refined. Originally, Collins refined such a method in [7].

As a first step, the truth values of the quantifier-free part ϕ' depending on the interpretation of all the variables $x_1, \ldots, x_n$ need to be determined. The sign-invariant cylindrical algebraic decomposition, defined by the set P_ϕ of multivariate polynomials occurring in ϕ, can be used to determine the truth values of the quantifier-free part ϕ'. The theorem shown below states, that the truth value of ϕ' is invariant in each cell of the cylindrical algebraic decomposition defined by P_ϕ. Thus, the theorem allows to reduce the problem of considering the uncountable set of all possible variable interpretations to considering a finite set of variable interpretations. The theorem was originally proved by Collins [7].

Theorem 3.2.1. *Let D be a sign-invariant cylindrical algebraic decomposition of n-space defined by the set P_ϕ of multivariate polynomials occurring in a quantified formula ϕ. Then, D is truth-invariant with respect to the quantifier-free part ϕ'.*

Proof. Let $D = (c_1, \ldots, c_m)$. Let $1 \leq l \leq m$ be arbitrary and choose $a := (a_1, \ldots, a_n) \in c_l$. Let $p \in P_\phi$ be arbitrary and let $\mathcal{C}_p$ be the corresponding constraint occurring in ϕ'. Assume $\mathcal{C}_p$ is *true* in a. Let $b := (b_1, \ldots, b_n) \in c_l$ be arbitrary. Since D is sign-invariant, p has the same sign in b as in a. Thus, $\mathcal{C}_p$ is *true* in b. Analogous, one can see that $\mathcal{C}_p$ would be *false* in

b, if assumed $\mathcal{C}_p$ is *false* in a. Since b was arbitrary, c_l is truth-invariant with respect to $\mathcal{C}_p$. Since p was arbitrary, c_l is truth-invariant with respect to any constraint occurring in ϕ. Since ϕ' is a Boolean combination of the constraints, c_l is truth-invariant with respect to ϕ'. Since $1 \leq l \leq m$ was arbitrary, D is truth-invariant with respect to ϕ'. □

In consequence of theorem 3.2.1 it suffices to determine the truth value of ϕ' in a sample point representing the cell. The following definition formalizes the idea of assigning truth values to cells.

Definition 3.2.2. *Let D be a truth-invariant cylindrical algebraic decomposition with respect to a formula ψ. A mapping*

$$\nu : D \mapsto \{\textit{true}, \textit{false}\}$$

will be called an evaluation. Let $c \in D$ be arbitrary and $a \in c$ a sample point. An evaluation ν is called the evaluation of ψ if, and only if, $\nu(c) = \psi(a)$. An evaluation of ψ will be denoted as ν_ψ.

So far, an evaluation $\nu_{\phi'}$ of the quantifier-free part ϕ' can be defined. Because of theorem 3.2.1, the mapping is well-defined. However, in a quantified formula, the last $n-k$ of the n variables $x_1, \ldots, x_n$ are quantified. Thus, the truth values of the quantified formula only depends on the interpretation of the free variables $x_1, \ldots, x_k$. In the following, the truth values of the quantifier-free part ϕ', depending on the interpretation of all the variables $x_1, \ldots, x_n$, will be used to determine the truth values of the quantified formula ϕ. In order to illustrate the idea how the truth values of ϕ can be determined using the truth values of ϕ', *i.e.* the evaluation of k-level cells can be defined using the evaluation of n-level cells, consider the exemplary quantified formula introduced in section 3.1.

$$\phi := \exists y \ (x^2 + y^2 - 1 < 0) \wedge (x + y < 0)$$

The cylindrical algebraic decomposition defined by P_ϕ is shown in the figure below. The 2-level cells in which ϕ' evaluates to true are hatched in red.

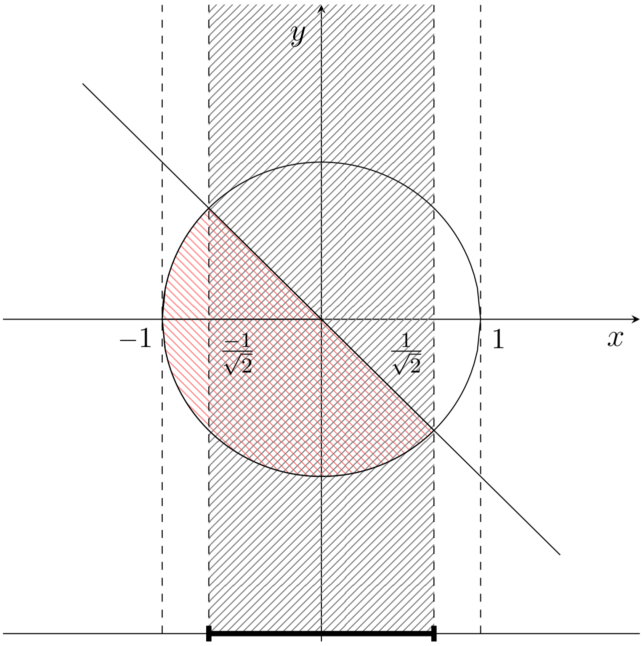

Figure 3.2: the cylindrical algebraic decomposition

Let c denote the 1-level cell $(-{}^1\!/\!\sqrt{2},\ {}^1\!/\!\sqrt{2})$. The truth value of c can be determined considering all 2-level cells in the stack over c. The stack over c is hatched in black. Since the variable y is existential quantified, the truth value of the quantified formula ϕ in c is *true*. That is because there exists a 2-level cell in the stack over c in which the quantifier-free part ϕ' is *true*.

The idea used in the example can be generalized. Beginning with the evaluation $\nu_{\phi'}$ of the quantifier-free part ϕ', the evaluation of the cells of level $i+1$ will be used to define the evaluation on the induced cylindrical algebraic decomposition of i-space, until level k is reached and the evaluation ν_ϕ of the quantified formula ϕ is obtained. Let $k \leq i < n$. In order to obtain the evaluation ν_ϕ as described, a method to define the evaluation on the induced cylindrical algebraic decomposition of i-space, provided the evaluation of the cells of level $i+1$, needs to be refined. In more general, let D' be a cylindrical algebraic decomposition of $(i+1)$-space and let D be the cylindrical algebraic decomposition of i-space induced by D'. The theorem shown below states, that the induced cylindrical algebraic decomposition inherits the property of truth-invariance.

Theorem 3.2.3. *Let $\psi(x_1, \ldots, x_i) := Qx_{i+1}\ \psi'(x_1, \ldots, x_i, x_{i+1})$ be a quantified formula. Let D' be a truth-invariant cylindrical algebraic decomposition of $(i+1)$-space with respect to ψ'. Let D be the cylindrical algebraic decomposition of i-space induced by D'. Then, D is truth-invariant regarding ψ.*

Proof. The theorem will be proved for $Q = \exists$. The prove for $Q = \forall$ is similar and can be found in [7]. Let $D' = (c_{1,1}, \ldots c_{1,m_1}, \ldots, c_{m,1}, \ldots, c_{m,m_m})$ and $D = (c_1, \ldots, c_m)$, such that $(c_{l,1}, \ldots, c_{l,m_l})$ is the stack over c_l. Let $1 \leq l \leq m$ be arbitrary and let $(a_1, \ldots, a_i) \in c_l$. Assume $\psi(a_1, \ldots, a_i)$ is *false*. Let $(b_1, \ldots, b_i) \in c_l$ and $b_{i+1} \in \mathbb{R}$ be arbitrary. Then, for some $1 \leq j \leq m_l$, $(b_1, \ldots, b_{i+1}) \in c_{l,j}$. Since $\psi(a_1, \ldots, a_i)$ is *false*, $\psi(a_1, \ldots, a_i, a_{i+1})$ is *false* for all $a_{i+1} \in \mathbb{R}$. That is because $Q = \exists$. Choose $a_{i+1} \in \mathbb{R}$ such that $(a_1, \ldots, a_{i+1}) \in c_{l,j}$. Since D is truth-invariant regarding ψ, $\psi(b_1, \ldots, b_{i+1})$ is *false*. Since b_{i+1} was arbitrary, $\psi(b_1, \ldots, b_i)$ is *false*. Since $(b_1, \ldots, b_i)$ was arbitrary in c_l, c_l is truth-invariant regarding ψ. Since $1 \leq l \leq m$ was arbitrary, D' is truth-invariant regarding ψ. □

In consequence of theorem 3.2.3, the mapping, defined by a method that defines the evaluation on the induced cylindrical algebraic decomposition of i-space using the evaluation of the cells of level $i+1$, is well defined. It remains to refine such a method. The idea is, that the truth value of a cell $c \in D$ can be determined considering each cell $c' \in D'$ in the stack over c. In the example, the variable y was existential quantified. Thus, the considered 1-level cell c was assigned to be *true*, since there existed a *true* cell in the stack over c. Analogous, if y would have been universal quantified, it would have been required for all cells in the stack over c to be *true* in order to assign c to be *true* as well. The theorem shown below formalizes this idea.

Theorem 3.2.4. *Let $\psi(x_1, \ldots, x_i) := Qx_{i+1}\ \psi'(x_1, \ldots, x_i, x_{i+1})$ be a quantified formula. Let D' be a truth-invariant cylindrical algebraic decomposition of $(i+1)$-space regarding ψ' and let D be the truth-invariant cylindrical algebraic decomposition of i-space induced by D'. Denote D and D' as in the proof of Theorem 3.2.3. Furthermore, let $\nu_{\psi'} : D' \mapsto \{true, false\}$*

be the evaluation of ψ'*. Then* $\nu_\psi : D \mapsto \{$*true*, *false*$\}$ *can be defined as* $\nu_\psi(c_l) = \vee_{j=1}^{m_l} \nu_{\psi'}(c_{l,j})$*, if* $Q = \exists$ *and as* $\nu_\psi(c_l) = \wedge_{j=1}^{m_l} \nu_{\psi'}(c_{l,j})$*, if* $Q = \forall$*.*

Proof. The theorem will be proved for $Q = \exists$. The prove for $Q = \forall$ is similar and can be found in [7]. Let $1 \leq l \leq m$. Choose $(a_1, \ldots, a_i) \in c_l$. First, assume $\psi(a_1, \ldots, a_i)$ is *false*. Then, $\psi'(a_1, \ldots, a_i, a_{i+1})$ is *false* for all $a_{i+1} \in \mathbb{R}$. That is because $Q = \exists$. Thus, for all $1 \leq j \leq m_l$ it is, that $\nu_{\psi'}(c_{l,j}) =$ *false*, since $\nu_{\psi'}$ is the evaluation of ψ'. Therefore, $\vee_{j=1}^{m_l} \nu_{\psi'}(c_{l,j})$ is *false* and thus $\nu_\psi(c_l) =$ *false*. Now, assume $\psi(a_1, \ldots, a_i)$ is *true*. Then, since $Q = \exists$, there exists an $a_{i+1} \in \mathbb{R}$ such that $\psi'(a_1, \ldots, a_i, a_{i+1})$ is *true*. Thus, there is $1 \leq j \leq m_l$ such that $(a_1, \ldots, a_i, a_{i+1}) \in c_{l,j}$ and $\nu_{\psi'}(c_{l,j}) =$ *true*. That is because $\nu_{\psi'}$ is the evaluation of ψ'. Therefore, $\vee_{j=1}^{m_l} \nu_{\psi'}(c_{l,j})$ is *true* and thus $\nu_\psi(c_l) =$ *true*. □

In consequence of theorem 3.2.4, a method to define an evaluation on the induced cylindrical algebraic decomposition of i-space, provided the evaluation of the cells of level $i+1$, is refined. In order to assign a truth value to a i-level cell c of D, simply consider the conjunction or disjunction of the truth values assigned to all $(i+1)$-level cells c' of D' in the stack over c depending whether the $(n-i)$'th variable is universal quantified or existential quantified. The algorithm given below determines the truth value of a quantified formula ϕ depending on the interpretation of the free variables $x_1, \ldots, x_k$ by successively applying the method described above, beginning with the evaluation of the quantifier-free part ϕ' on the cylindrical algebraic decomposition of n-space, defined by the set of multivariate polynomials occurring in ϕ and until an evaluation of ϕ on the induced cylindrical algebraic decomposition of k-space, is obtained.

Algorithm 3.2.1 Determining Truth Values

Input: A quantified formula ϕ

Output: The evaluation ν_ϕ of ϕ

```
$P_\phi \leftarrow$ set of multivariate polynomials occurring in $\phi$
$D \leftarrow$ sign-invariant cylindrical algebraic decomposition defined by $P_\phi$

for $c \in D$ do
    for $p \in P_\phi$ do
        evaluate the sign of $p$ in $c$
    end for
    $\nu_n(c) \leftarrow$ evaluate the truth of $\phi'$ in $c$ using the signs of the $p$'s
end for

for $i = n$ down to $k$ do
    $D_i \leftarrow$ induced cylindrical algebraic decomposition of $i$-space
    for $c \in D_i$ do
        if $Q_i = \exists$ then
            $\nu_i(c) \leftarrow$ false
            for $c'$ in the stack over $c$ do
                if $\nu_{i+1}(c') =$ true then
                    $\nu_i(c) \leftarrow$ true
                end if
            end for
        else
            $\nu_i(c) \leftarrow$ true
            for $c'$ in the stack over $c$ do
                if $\nu_{i+1}(c') =$ false then
                    $\nu_i(c) \leftarrow$ false
                end if
            end for
        end if
    end for
end for

return $\nu_\phi \leftarrow \nu_k$
```

3.3 Simplifications

In the following, a method to simplify the cylindrical algebraic decompositions of k-space and below, induced by a cylindrical algebraic decomposition, defined by the set of multivariate polynomials occurring in a quantified formula, is refined. The hope is, that a simpler equivalent quantifier-free formula can be constructed from a simpler cylindrical algebraic decomposition. Originally, Brown refined such a method [2].

A cylindrical algebraic decomposition D' is called simpler than a cylindrical algebraic decomposition D, if D' results from a merge of cells in D, *i.e.* each cell of D' is a union of cells of D. The removal of a section results in the merge of the considered section and its two adjacent sectors. A section is removed by the removal of the projection factors defining the considered section, *i.e.* the projection factors being zero in the considered section. Therefore, a cylindrical algebraic decomposition is simplified by the removal of projection factors.

A result from section 3.2 was, that the cylindrical algebraic decomposition of k-space, induced by a cylindrical algebraic decomposition, defined by the set of multivariate polynomials occurring in a quantified formula ϕ, is truth-invariant with respect to ϕ. A simpler cylindrical algebraic decomposition still needs to assure the truth-invariance with respect to ϕ. The definition shown below introduces a type of cell which must not be removed in order to assure truth-invariance.

Definition 3.3.1. *A k-level section is called a truth-boundary cell, if the truth values of the considered k-level section and its two adjacent k-level sectors do not agree.*

Since a section is removed by the removal of the projection factors defining the considered section, each k-level truth-boundary cell defines a set of k-level projection factors of which at least one must be kept in order to keep the considered truth-boundary cell. Therefore, a minimal hitting set of the collection of these sets defines a set of k-level projection factors which, if kept, assure truth-invariance. Any other k-level projection factor can be safely removed in order to simplify the cylindrical algebraic decomposition without consequently violating truth-invariance.

The induced cylindrical algebraic decompositions of $(k-1)$-space or below may be simplified as well, hoping for the possibility to construct an even simpler equivalent quantifier-free formula. For level $k-1$ or below, besides assuring truth-invariance, it must be assured that the projection factor set will still be closed under projection after the simplifications are done. Additionally, truth-boundary cells of level $k-1$ or below can not be as easily defined as for level k. That is because only k-level cells have truth values assigned to them. In order to still provide a definition of truth-boundary cells of level $k-1$ or below, a k-level cell c is said to be above a cell c' of level $k-1$ or below, if the set-theoretic projection of c onto the respective level is c'. Furthermore, the k-level cells above a section of level $k-1$ or below and its two adjacent sectors of the respective level are said to correspond, if the projection factors of higher level are delinable over the union of the considered section and its two adjacent sectors.

Definition 3.3.2. *A section of level $k-1$ or below is called a truth-boundary cell, if there are corresponding k-level cells above the considered section and its two adjacent sectors, such that their truth values do not agree.*

The same as for level k, each truth-boundary cell of level $k-1$ or below defines a set of projection factors of the respective level of which at least one must be kept in order to keep the considered truth-boundary cell. A minimal hitting set of the collection of these sets defines a set of projection factors of the respective level which, if kept, assure truth-invariance. In addition, each projection factor of the respective level, being the result of the application of the projection operator on projection factors of higher level, need to be kept in order to assure, that the projection factor set will still be closed under projection. The algorithm shown below simplifies a cylindrical algebraic decomposition, defined by the set of multivariate polynomials occurring in a quantified formula, beginning with level k, by successively simplifying the induced cylindrical algebraic decompositions, until level 1 is reached.

Algorithm 3.3.1 Simplify Cylindrical Algebraic Decompositions

Input: A cylindrical algebraic decomposition D

Output: A simpler cylindrical algebraic decomposition D'

```
$D' \leftarrow D$
for $i = k$ down to $1$ do
    $C \leftarrow$ the set of $i$-level truth-boundary cells in $D'$
    $S_1, S_2 \leftarrow \emptyset$
    for $c$ in $C$ do
        $S_1 \leftarrow S_1 \cup \{p \in P_i \mid p$ is zero in $c\}$
    end for
    $H \leftarrow$ a minimal hitting set for $S_1$
    if $i \neq k$ then
        $S_2 \leftarrow \{p \in P_i \mid p$ is in the closure under projection of $P'_{i+1} \cup \ldots \cup P'_k\}$
    end if
    $P'_i \leftarrow H \cup S_2$
    $P' \leftarrow P_1 \cup \ldots \cup P_{i-1} \cup P'_i \cup \ldots \cup P'_k$
    $D' \leftarrow$ the cylindrical algebraic decomposition defined by $P'$
end for

return $D'$
```

The obtained decomposition is still cylindrical as well as algebraic, thus is a cylindrical algebraic decomposition. That is because it was assured that the set of multivariate polynomials resulting from the removal of projection factors is still closed under projection. The obtained cylindrical algebraic decomposition is still truth-invariant, since no truth-boundary cells were removed. The obtained truth-invariant cylindrical algebraic decomposition is simpler or unaltered, since projection factors were possibly removed, resulting in the possible merge of a section defined by the removed projection factor and its two adjacent sectors.

Consider the slightly modified exemplary quantified formula introduced in section 3.1 given below.

$$\phi := \exists y \; (x^2 + y^2 - 1 < 0) \wedge (x + y < 0) \wedge (y - x - 1 < 0)$$

The cylindrical algebraic decomposition defined by the multivariate polynomials occurring in ϕ is shown below. Adding the multivariate polynomial $y - x - 1$ results in two additional 1-level projection factors $2x + 1$ and $x^2 + x$. The additional 1-level projection factors define additional 1-level sections and sectors. The 1-level cell $(-1/\sqrt{2}, 1/\sqrt{2})$ in the cylindrical algebraic decomposition, defined by the multivariate polynomials occurring in the exemplary quantified formula introduced in section 3.1, is the union of the 1-level cells $(-1/\sqrt{2}, -1/2)$, $\{-1/2\}$, $(-1/2, 0)$, $\{0\}$ and $(0, 1/\sqrt{2})$ in the cylindrical algebraic decomposition shown below.

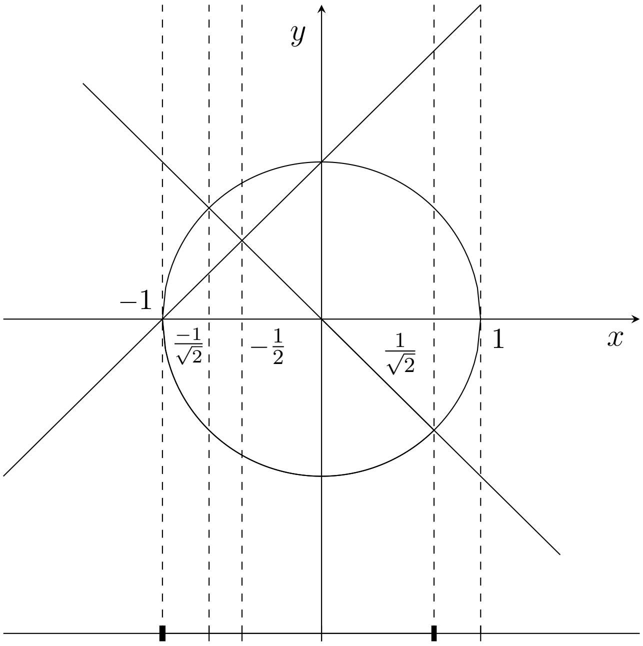

Figure 3.3: the cylindrical algebraic decomposition

The 1-level truth-boundary cells $\{-1\}$ and $\{1/\sqrt{2}\}$ are marked. The sets of 1-level projection factors defining the sections $\{-1\}$ and $\{1/\sqrt{2}\}$ are $\{x^2 - 1,\ x^2 + x\}$ respectively $\{2x^2 - 1\}$. A hitting set is $\{x^2 - 1,\ 2x^2 - 1\}$. As a result, the two additional 1-level projection factors $2x + 1$ and $x^2 + x$ can be safely removed without consequently violating truth-invariance.

3.4 Signatures

Section 3.2 refined a method to determine the truth values of a quantified formula ϕ depending on the interpretation of the free variables $x_1, \ldots, x_k$. This information is necessary in order to eliminate quantifiers. The present section introduces a crucial term in order to talk about a method determining whether this information already suffices to eliminate quantifiers, which is to be refined in the next section, or if additional work needs to be done.

The projection factor set P for a set of multivariate polynomials P_ϕ occurring in a quantified formula ϕ is finite and can be partitioned by the level of the projection factors. Therefore, the projection factor set P can be written as $P = P_1 \cup \cdots \cup P_n$, where the set P_i denotes the set of i-level projection factors. Since P is finite, the set P_i is finite and can be written as $P_i = \{p_{i,1}, \ldots, p_{i,m_i}\}$. Accordingly, the set of projection factors of level at most k can be written as shown below.

$$P_1 \cup \cdots \cup P_k = \{p_{1,1}, \ldots, p_{1,m_1}, \ldots, p_{k,1}, \ldots, p_{k,m_k}\}$$

Given a point in k-space, a projection factor p of level at most k is either negative, positive or zero in that point, depending whether the result is negative, positive or zero when substituting the point for the variables of the projection factor. Since sign-invariant cylindrical algebraic decompositions are considered, a pair of a k-level cell c and a projection factor p of level at most k can be assigned a sign. If the context clearly states which k-level cell c is considered, *sgn*(p) denotes the sign p has in c.

Definition 3.4.1. *Let c be a k-level cell. The tuple of signs sgn(p), any projection factor p of level at most k has in c, is called the signature of the cell c.*

$$(\textit{sgn}(p_{1,1}), \ldots, \textit{sgn}(p_{1,m_1}), \ldots, \textit{sgn}(p_{k,1}), \ldots, \textit{sgn}(p_{k,m_k}))$$

For the example introduced in section 3.1 the set of projection factors of level at most 1 is $\{x^2 - 1,\ 2x^2 - 1\}$. The signature of the exemplary 1-level cell $(-1, -{}^1\!/\!\sqrt{2})$ is $(-,+)$. That is because $-$ and $+$ are the signs of the values obtained by substituting the sample point $-{}^7\!/\!{}_8$ for x. The calculation is shown below.

$$\begin{aligned}(-,+) &= \left(\mathit{sgn}\left(-\frac{15}{64}\right),\ \mathit{sgn}\left(\frac{17}{32}\right)\right)\\ &= \left(\mathit{sgn}\left(\left(-\frac{7}{8}\right)^2-1\right),\ \mathit{sgn}\left(2\cdot\left(-\frac{7}{8}\right)^2-1\right)\right)\\ &= \left(\mathit{sgn}\left(x^2-1\right)\Big|_{-7/8},\ \mathit{sgn}\left(2x^2-1\right)\Big|_{-7/8}\right)\end{aligned}$$

Likewise, the signatures for each 1-level cell can be computed. The signatures of the 1-level cells are shown in the table given below.

1-level cell	sample point	signature
$(-\infty,-1)$	-2	$(+,+)$
$\{-1\}$	-1	$(0,+)$
$(-1, -{}^1\!/\!\sqrt{2})$	$-{}^7\!/\!{}_8$	$(-,+)$
$\{-{}^1\!/\!\sqrt{2}\}$	$-{}^1\!/\!\sqrt{2}$	$(-,0)$
$(-{}^1\!/\!\sqrt{2}, {}^1\!/\!\sqrt{2})$	0	$(-,-)$
$\{{}^1\!/\!\sqrt{2}\}$	${}^1\!/\!\sqrt{2}$	$(-,0)$
$({}^1\!/\!\sqrt{2}, 1)$	${}^7\!/\!{}_8$	$(-,+)$
$\{1\}$	1	$(0,+)$
$(1,\infty)$	2	$(+,+)$

As one can easily see, the signatures of k-level cells are in general not unique. For example, consider the two distinct 1-level cells $(-1, -{}^1\!/\!\sqrt{2})$ and $({}^1\!/\!\sqrt{2}, 1)$ from the table shown above, both have the signature $(-,+)$.

3.5 Projection-Definability Test

The quantifier elimination method by cylindrical algebraic decomposition introduced by Collins [7] guaranteed to define an equivalent quantifier-free formula solely by using the projection factors and the information provided by the truth values of the considered quantified formula depending on the interpretation of the free variables. This is due to the so called augmented-projection used by Collin's method. In this work, Brown's projection operator [3] is considered. The usage of Brown's projection operator results in a smaller projection factor set. The hope is, that a simpler equivalent quantifier-free formula can be constructed using a projection operator producing less projection factors. However, a smaller projection factor set may lack some projection factors needed to construct an equivalent quantifier-free formula solely from the projection factors. Below, a property a cylindrical algebraic decomposition can have, formalizing the idea described above, is defined.

Definition 3.5.1. *Let ϕ be a quantified formula. The cylindrical algebraic decomposition, defined by the set P_ϕ of the multivariate polynomials occurring in ϕ, is called projection-definable if an equivalent quantifier-free formula for ϕ can be constructed solely from the projection factors.*

In order to understand projection-definability better, the definition is rephrased in the theorem shown below. In section 3.4 it was already stated that, in general, the signatures of k-level cells are not unique. Furthermore, k-level cells which agree in their signature do, in general, not even need to agree in their truth value. For example, consider the two distinct 1-level cells $(-1, -{}^1/\sqrt{2})$ and $({}^1/\sqrt{2}, 1)$ in the cylindrical algebraic decomposition introduced with the exemplary quantified formula ϕ from section 3.1. Both cells have the signature $(-, +)$, as seen in section 3.4. Yet it was seen in section 3.2, that ϕ is *true* in $(-1, -{}^1/\sqrt{2})$ but *false* in $({}^1/\sqrt{2}, 1)$. The existence of two distinct k-level cells agreeing in their signature but disagreeing in their truth value cause a cylindrical algebraic decomposition to be projection-undefinable, as the following theorem, originally proved by Brown [2], shows.

Theorem 3.5.2. *Let ϕ be a quantified formula. The cylindrical algebraic decomposition D, defined by the set P_ϕ of the multivariate polynomials occurring in ϕ, is projection-definable if there are no k-level cells c and c' agreeing in their signature but disagreeing in their truth values.*

Proof. Assume D is projection-definable. By the definition of projection-definability, there exists an equivalent quantifier-free formula ψ constructed solely from the projection factors. Assume, for the sake of contradiction, there exist two distinct k-level cells c and c' agreeing in their signature but disagreeing in their truth values. In particular, the signs of all projection factors of level at most k are the same in c and c'. Thus, the truth values of any formula constructed solely from the projection factors of level at most k are the same in c and c'. Without loss of generality, assume ϕ is *true* in c and *false* in c'. Since the truth values of ψ are the same in c and c', ψ is either *false* in c or *true* in c, a contradiction.

Now, assume there are no k-level cells c and c' agreeing in their signature but disagreeing in their truth values. Below, a quantifier-free formula ψ, which is *true* if ϕ is *true* for any interpretation of the free variables $x_1, \ldots, x_k$, will be constructed. Let c be a k-level cell and let p be a projection factor of level at most k. An atomic formula $\psi_{c,p}$ describing which sign p has in c can be defined as below.

$$\psi_{c,p} := \begin{cases} p < 0 & \text{if sgn}(p) \text{ is negative in } c \\ p = 0 & \text{if sgn}(p) \text{ is zero in } c \\ p > 0 & \text{if sgn}(p) \text{ is positive in } c \end{cases}$$

A formula ψ_c describing the signature of c can be defined using the atomic formulas $\psi_{c,p}$ describing which signs the projection factors of level at most k have in c.

$$\psi_c := \bigwedge_{p \in P_1 \cup \cdots \cup P_k} \psi_{c,p}$$

Clearly, ψ_c is *true* in c. Therefore, the formula ψ, defined as the disjunction of the formulas ψ_c for all k-level cells c in which ϕ is *true*, is a quantifier-free formula which is *true* if ϕ is *true*. It remains to be shown, that ψ is *true* only if ϕ is *true*. Since there is no k-level cell c in which ϕ is *true* and which agrees in its signature with any k-level cell c' in which ϕ is *false*, there is no disjunct ψ_c in ψ such that ψ_c is satisfied by any point in any cell c' in which ϕ is *false*. Thus, ψ is *true* if, and only if, ϕ is *true*. Therefore, ψ is an equivalent quantifier-free formula. □

As a consequence of theorem 3.5.2, a cylindrical algebraic decomposition, defined by the set P_ϕ of the multivariate polynomials occurring in a quantified formula ϕ, can be tested for projection-definability.

Algorithm 3.5.1 Projection Definability Test

Input: A cylindrical algebraic decomposition D

Output: An answer whether D is projection definable or not

```
$L \leftarrow$ list of all $k$-level cells, sorted by signature
for $i = 1$ up to $|L|$ do
    if signatures for $L[i]$, $L[i+1]$ agree then
        if truth values for $L[i]$, $L[i+1]$ disagree then
            return false
        end if
    end if
end for

return true
```

As it was already seen above, the cylindrical algebraic decomposition introduced with the exemplary quantified formula ϕ from section 3.1 is projection-undefinable. That is because the distinct 1-level cells $(-1, -1/\sqrt{2})$ and $(1/\sqrt{2}, 1)$ agree in their signatures but disagree in their truth values. In fact, there are even more 1-level cells causing the cylindrical algebraic decomposition to be projection-undefinable.

3.6 Assure Projection-Definablility

If the method refined in section 3.5 determines a cylindrical algebraic decomposition, defined by a set of multivariate polynomials occurring in a quantified formula ϕ, to be projection-undefinable, the cylindrical algebraic decomposition needs to be modified such that projection-definability can be assured after the modifications are done. Originally, Brown refined such a method [2]. In section 3.2 it was shown, that a cylindrical algebraic decomposition, defined by the set of multivariate polynomials occurring in a quantified formula ϕ, is truth-invariant with respect to ϕ. Thus, the truth values of cells can not be altered in order to achieve projection-definability. Instead, the approach of adding new polynomials in order to distinguish the problem-causing cells by the signs of the newly added polynomials will be taken, such that cells former causing the cylindrical algebraic decomposition to be projection-undefinable will not agree in their signature anymore.

The generalized Thom's Lemma [10] gives a hint on which polynomials to add. It introduces so called stratisfying families of multivariate polynomials. A stratisfying family is a family $(p_{i,j})$ of multivariate polynomials, for $i = 1, \ldots, m$ and $j = 1, \ldots, l_i$, such that the family is closed under projection and for any fixed i the subfamily $p_{i,j}$ for $j = 1, \ldots, l_i$ is closed under partial derivation with respect to x_i. Consider the semi-algebraic set c for a family$(\prec_{i,j})$ of relation symbols in $\{=, <, >\}$ given below.

$$c = \bigcap_{i=1}^{m} \bigcap_{j=1}^{l_i} \{x \in \mathbb{R}^n \mid \mathsf{sgn}(p_{i,j}) \prec_{i,j} 0\}$$

The generalized Thom's Lemma states, that c is either empty or connected. Furthermore, the decomposition given by the semi-algebraic sets for the families of relation symbols resulting in a non-empty set yield a cylindrical algebraic decomposition of n-space. As a consequence, the cells of the resulting cylindrical algebraic decomposition are distinguishable by the signs of the multivariate polynomials $p_{i,j}$ in the stratisfying family. That is because the semi-algebraic sets above are defined as connected sets fulfilling conjunctions of sign conditions. Thus, there can not exist cells which agree in their signatures but disagree in their truth values, since there are no distinct cells agreeing in their signature. With Thom's Lemma in mind, a simple algorithm to make a cylindrical algebraic decomposition projection-definable can be given. Simply construct the closure under derivation and projection of the set of multivariate polynomials occurring in a quantified formula ϕ. The resulting cylindrical algebraic decomposition defined by the obtained set of multivariate polynomials will be projection-definable.

In section 3.5 it was already shown, that the cylindrical algebraic decomposition introduced with the exemplary quantified formula ϕ from section 3.1 is projection-undefinable. Adding the polynomial x, the normalized first derivative of the projection factor $2x^2 - 1$, respectively $x^2 - 1$, already suffices to distinguish all the cells causing the cylindrical algebraic decomposition to be projection-undefinable by the sign of the additional polynomial. The table given below shows, that after adding the normalized first derivative of the projection factor $2x^2-1$, respectively x^2-1, there are no cells agreeing in their signature but disagreeing in their truth values anymore.

1-level cell	sample point	signature	Truth Value
$(-\infty,-1)$	-2	$(+,+,-)$	*false*
$\{-1\}$	-1	$(+,0,-)$	*false*
$(-1,-{}^1/\sqrt{2})$	$-{}^7/_8$	$(+,-,-)$	*true*
$\{-{}^1/\sqrt{2}\}$	$-{}^1/\sqrt{2}$	$(0,-,-)$	*true*
$(-{}^1/\sqrt{2},0)$	$-{}^1/_4$	$(-,-,-)$	*true*
$\{0\}$	0	$(-,-,0)$	*true*
$(0,{}^1/\sqrt{2})$	${}^1/_4$	$(-,-,+)$	*true*
$\{{}^1/\sqrt{2}\}$	${}^1/\sqrt{2}$	$(0,-,+)$	*false*
$({}^1/\sqrt{2},1)$	${}^7/_8$	$(+,-,+)$	*false*
$\{1\}$	1	$(+,0,+)$	*false*
$(1,\infty)$	2	$(+,+,+)$	*false*

In practice, the closure under derivation and projection can be rather large, resulting in the practical impossibility of constructing the cylindrical algebraic decomposition. Below, a method to determine a preferably small set of polynomials to add in order to achieve projection-definability is refined as well as the underlying theory is introduced. The definition given below introduces a term that will help to do so.

Definition 3.6.1. *Let D be the cylindrical algebraic decomposition defined by the set P_ϕ of the multivariate polynomials occurring in a quantified formula ϕ. Let there be two k-level cells c and d agreeing in their signatures but disagreeing in their truth values. Let $1 \leq i \leq k$. Two distinct i-level cells c' and d' are called a conflicting pair for c and d if they are the result of the projection of c and d onto i-space and if they are in the same stack.*

The theorem given below shows, that there is always a unique conflicting pair for two distinct k-level cells agreeing in their signatures but disagreeing in their truth values. The theorem was originally proved by Brown [2].

Theorem 3.6.2. *Let D be the cylindrical algebraic decomposition defined by the set P_ϕ of the multivariate polynomials occurring in a quantified formula ϕ. If there are two k-level cells c and d agreeing in their signature but disagreeing in their truth values, then there is a unique conflicting pair for c and d.*

Proof. All cells in 1-space are in the same stack. Therefore, the projection of c and d onto 1-space must be in the same stack. Let $1 \leq i \leq k$ be the greatest level such that the projection of c and d onto i-space are in the same stack. Let c' and d' be the result of the projection of c and d onto i-space, then c' and d' are distinct. For the sake of contradiction, assume they are not distinct. Since c' and d' are in the same stack but not distinct, they are equal. Thus, the resulting cells of the projection of c and d onto $(i+1)$-space are in the same stack, a contradiction. □

The next theorem, originally proved by Brown [2], assures that no new conflicting pairs are created by adding new polynomials.

Theorem 3.6.3. *Let D_1 be the cylindrical algebraic decomposition defined by the set $P_{1,\phi}$ of all multivariate polynomials occurring in a quantified formula ϕ. Let $P_{2,\phi}$ be a superset of $P_{1,\phi}$ and D_2 the cylindrical algebraic decomposition defined by $P_{2,\phi}$. If c'_2 and d'_2 are an i-level conflicting pair in D_2, then there are c'_1 and d'_1, a j-level conflicting pair in D_1 for a $i \leq j \leq k$. Furthermore, c'_2 and d'_2 are subsets of the projections of c'_1 and d'_1 onto i-space.*

Proof. Since c'_2 and d'_2 are a conflicting pair, there are k-level cells c_2 and d_2 agreeing in their signature but disagreeing in their truth values. Recall that P_2 is the closure under the projection operator of $P_{2,\phi}$. With that said, it is that each element in P_2 of level at most k has the same sign in c_2 as in d_2. Without loss of generality, let ϕ be *true* in c_2 and *false* in d_2. Since $P_{1,\phi} \subseteq P_{2,\phi}$, it is that $P_1 \subseteq P_2$ as well. Therefore, each cell of D_1 is a union of cells from D_2. In particular, it is that each k-level cell of D_1 is the union of k-level cells from D_2. Thus, there is a k-level cell c_1 that is a superset of c_2 and a k-level cell d_1 that is a superset of d_2. Since sign-invariant cylindrical algebraic decomposition are considered and P_1 is a subset of P_2, it is that each element in P_1 of level at most k has the same sign in c_1 as in d_1. Since truth-invariant cylindrical algebraic decompositions are considered, ϕ is *true* in c_1 and *false* in d_1. Thus, c_1 and d_1 are k-level cells agreeing in their signature but disagreeing in their truth values. Therefore, there is a unique j-level conflicting pair c_1' and d_1' for c_1 and d_1, by theorem 3.6.2. Since c_1 is a superset of c_2, it is that for any $1 \leq l \leq j$ the projection onto l-space of c_1 is a superset of the projection onto l-space of c_2. Analogous, the same result applies for d_1 and d_2. In particular, this statement holds for level i. Therefore, since c_2 and d_2 are in the same stack, the projection of c_1 and d_1 onto i-space are in the same stack. The projections of c_1 and d_1 onto i-space might not be distinct, but there is a $i \leq j \leq k$ such that the projection of c_1 and d_1 onto j-space is distinct, as seen in theorem 3.6.2. □

It remains to find an answer to the question which polynomials to add in order to remove a conflicting pair. The definition given below will help to answer this question.

Definition 3.6.4. *Let P be a set of multivariate polynomials of level at most i. The set P^* denotes the closure under derivation with respect to x_i. For an i-level polynomial p the set $\{p\}^*$ is denoted as p^*.*

The application of the lemma shown below will be useful to prove the following theorem.

Lemma 3.6.5. *Let A be a region in $\mathbb{R}^{i-1}$ and let P be a set of multivariate polynomials of level at most i. If P is delinable over A, the cells in the decomposition of the cylinder $A \times \mathbb{R}$ defined by P are distinguishable by the signs of the elements of P.*

Proof. The lemma will be proved by induction on d, the maximal degree with respect to x_i. The case $d = 1$ is trivial, thus consider $d > 1$. Assume, for the sake of contradiction, the lemma does not hold. Then, there exist two distinct cells c and d in which the signs of all elements of P agree. Let $P' = \{p \in P \mid$ the degree of p with respect to x_i is smaller than $d\}$.

Furthermore, let D' be the decomposition of the cylinder $A \times \mathbb{R}$ defined by P'. Note that $P' \subseteq P$. Thus, there exist two cells c' and d' in D', such that c is a subset of c' and d is a subset of d'. By induction, there is a $p' \in P'$ such that the sign of p' disagrees in c' and d', if c' and d' are distinct. Since $P' \subseteq P$ and c is a subset of c' as well as d is a subset of d', a contradiction. It remains to be shown that c' and d' are distinct. It suffices to show that there is an element in P' which is zero in some cell between c' and d'. Since c and d are distinct cells, there is a $p \in P$ such that p is either zero in both c and d or zero in at least one cell between c and d. In the following, it will be shown that the first derivative of p with respect to x_i is an element in P' which is zero in some cell between c' and d'. If p is zero in exactly one cell between c and d and not zero in neither c nor d, the multiplicity of the respective root is at least 2. Thus, the first derivative of p with respect to x_i is zero in at least one cell between c and d. If p is zero in at least two cells between c and d or zero in both c and d, Rolle's theorem implies that the first derivative of p with respect to x_i is zero in at least one cell between c and d. In either case, there is an element in P' which is zero in some cell between c' and d'. That is because the first derivative of p with respect to x_i is an element of P' and c is a subset of c' as well as d is a subset of d'. □

Finally, the following theorem shows which polynomials to add in order to remove a conflicting pair. The theorem was originally proved by Brown [2].

Theorem 3.6.6. *Let D_1 be the cylindrical algebraic decomposition defined by the set $P_{1,\phi}$ of the multivariate polynomials occurring in a quantified formula ϕ. Let $1 \leq i \leq k$. Furthermore, let c'_1 and d'_1 be an i-level conflicting pair in D_1 and let p be an i-level projection factor, such that p is zero in c'_1 and d'_1 or in some cell between c'_1 and d'_1, but not identically zero in the stack between c'_1 and d'_1. Let $P_{2,\phi} = P_{1,\phi} \cup p^*$ and let D_2 be the cylindrical algebraic decomposition defined by $P_{2,\phi}$. Then, there is no conflicting pair c'_2 and d'_2 in D_2 such that c'_2 is a subset of c'_1 and d'_2 is a subset of d'_1.*

Proof. Assume there is a conflicting pair c'_2 and d'_2 in D_2 such that c'_2 is a subset of c'_1 and d'_2 is a subset of d'_1. Note that c'_2 and d'_2 are an i-level conflicting pair, since c'_1 and d'_1 are. Let A be the union of the set-theoretic projections of c'_2 and d'_2 onto $(i-1)$-space. Since, by the definition of a conflicting pair, c'_2 and d'_2 are i-level cells and in the same stack, the set-theoretic projection of c'_2 and the set-theoretic projection of d'_2 onto $(i-1)$-space are equal. Thus, A is an $(i-1)$-level cell of D_2. Note that A is a region of $\mathbb{R}^{i-1}$. Let D be the decomposition of the cylinder $A \times \mathbb{R}$ defined by p^*. Note that p^* is delinable over A, since A is an $(i-1)$-level cell of D_2 and p^* is a subset of $P_{2,\phi}$. By lemma 3.6.5, the cells of D are distinguishable by the signs of the elements of p^*. Since p^* is a subset of $P_{2,\phi}$ there are cells c' and d' in D such that c'_2 is a subset of c' and d'_2 is a subset of d'. Note that by the assumptions on p, p is zero in c'_1 and d'_1 or in some cell between c'_1 and d'_1, but not identically zero in the stack between c'_1 and d'_1. Thus, p is zero in c' and d' or in some cell between c' and d', but not identically zero in the stack between c' and d'. Since p is in p^*, c' and d' are distinct. Therefore, there is an element in p^* having different signs in c' and d' and thus different signs in c'_2 and d'_2. Since all elements of p^* are of level at most k by the assumptions on p, c'_2 and d'_2 disagree in their signature, a contradiction. $\square$

The theory introduced above can be used to refine a method that is capable of modifying a given cylindrical algebraic decomposition such that projection-definablility can be assured after the modifications are done. The method successively eliminates i-level conflicting pairs, beginning with level k down to level 1. Such that, after the i'th iteration, there are no conflicting pairs of level i or higher left. The absence of conflicting pairs is equal to projection-definability, as theorem 3.5.2 stated. In the i'th iteration, polynomials are added in order to remove the i-level conflicting pairs. Theorem 3.6.3 assured that no new conflicting pairs are created by adding new polynomials. For an i-level conflicting pair, theorem 3.6.6 gave an answer to the question which polynomials to add, in order to remove the considered conflicting pair. The algorithm shown below realises the method described above in order to modify a cylindrical algebraic decomposition, such that projection-definability can be assured after the modifications are done.

Algorithm 3.6.1 Assure Projection-Definability

```
Input: A cylindrical algebraic decomposition $D$
Output: A projection-definable cylindrical algebraic decomposition $D'$
$D' \leftarrow D$
for $l = k$ down to $1$ do
    $C \leftarrow$ set of all $i$-level conflicting pairs $c$ and $d$ in $D'$
    $S \leftarrow \emptyset$
    for each conflicting pair $c$ and $d$ in $C$ do
        $P_{c,d} \leftarrow$ set of $i$-level projection factors to add for $c$ and $d$
        $S \leftarrow S \cup \{P_{c,d}\}$
    end for
    $H \leftarrow$ a hitting set for $S$
    $D' \leftarrow$ the cylindrical algebraic decomposition defined by $P \cup H^*$
end for

return $D'$
```

3.7 Equivalent Quantifier-Free Formulas

In the following, a method to construct an equivalent quantifier-free formula is refined. Originally, Collins introduced such a method [7]. Hong described a refinement of Collins' method which produces simpler equivalent quantifier-free formulas based on minimization [13]. Brown took up on Hong's idea but used minimal hitting sets to simplify the constructed equivalent quantifier-free formulas [2].

Let ϕ be a quantified formula in n variables $x_1, \ldots, x_n$, of which the first k variables are free and the remaining $n-k$ variables are quantified. Furthermore, let D be the cylindrical algebraic decomposition defined by P_ϕ, the set of the multivariate polynomials occurring in ϕ. Let a be a point in k-space, *i.e.* an interpretation of the free variables $x_1, \ldots, x_k$. A formula ψ is equivalent to ϕ, if a satisfies ψ if, and only if, a satisfies ϕ. The point a satisfies ϕ if, and only if, there is a k-level cell c, which is determined to be *true* by the method refined in section 3.2, such that a is in c. As a consequence, a formula ψ is equivalent to ϕ, if a satisfies ψ if, and only if, there is a *true* k-level cell c, such that a is in c. Thus, a formula ψ describing exactly the *true* k-level cells is equivalent to ϕ. If D is determined to be projection-definable by the method refined in section 3.5 or, if not, is modified to be projection-definable by the method refined in section 3.6, the *true* and *false* k-level cells can be distinguished by the signs of the projection factors. Thus, a formula ψ describing exactly the *true* k-level cells can be constructed solely from the atomic formulas in the set $\mathcal{A}$, defined as below. The atomic formulas in the set $\mathcal{A}$ are sign conditions.

$$\mathcal{A} := \left\{ \begin{cases} p = 0 \\ p < 0 \\ p > 0 \end{cases} \middle| \; p \in P_1 \cup \cdots \cup P_k \right\}$$

For the projection factor set introduced with the exemplary quantified formula ϕ from section 3.1, the set of atomic formulas $\mathcal{A}$ is shown below.

$$\mathcal{A} = \{2x^2 - 1 = 0,\ 2x^2 - 1 < 0,\ 2x^2 - 1 > 0, x^2 - 1 = 0,$$
$$x^2 - 1 < 0,\ x^2 - 1 > 0, x = 0,\ x < 0,\ x > 0\}$$

The definition given below introduces a type of formula, which are constructed from the atomic formulas in $\mathcal{A}$ and which describe *true* k-level cells.

Definition 3.7.1. *Let $\mathcal{A}$ be the set of sign conditions on the projection factors of level at most k. An implicant is a conjunction of atomic formulas in $\mathcal{A}$, such that at least one true k-level cell satisfies the conjunction but no false k-level cell does. A prime implicant is an implicant such that removing any conjunct, the resulting formula would not be an implicant anymore.*

The definition shown below introduces a term describing an implicant constructed such that a given k-level cell satisfies it.

Definition 3.7.2. *Let c be a true k-level cell and let $\mathcal{I}$ be an implicant. The implicant $\mathcal{I}$ is said to capture c, if c satisfies $\mathcal{I}$.*

Since any cylindrical algebraic decompositions, defined by the set of multivariate polynomials occurring in a quantified formula, can be modified to be projection-definable, there exists an implicant for all *true* k-level cells capturing it. That is because there are projection factors of level at most k distinguishing the considered *true* k-level cell from all *false* k-level cells. Thus, a method that constructs an implicant capturing a given *true* k-level cell c can be refined. At first, the set of atomic formulas in $\mathcal{A}$ which hold in c is build. Then, for each *false* k-cell c', the subset of the atomic formulas in $\mathcal{A}$ which hold in c but do not hold in c' is determined. Finally, constructing a minimal hitting set for the obtained collection of subsets results in a set of atomic formulas which conjunction is a prime implicant capturing c. The algorithm shown below, originally introduced by Brown [2], realises the method described above in order to construct a prime implicant capturing a given k-level cell.

Algorithm 3.7.1 Construct an implicant capturing a cell

```
Input: A cylindrical algebraic decomposition $D$, a $k$-level cell $c$
Output: An Implicant $\mathcal{I}$ capturing $c$
$\mathcal{A}_c \leftarrow$ all atomic formulas in $\mathcal{A}$ that evaluate to true in $c$

$S \leftarrow \emptyset$
for each false cell $c'$ do
    $S_{c'} \leftarrow$ all atomic formulas in $\mathcal{A}_c$ that evaluate to false in $c'$
    $S \leftarrow S \cup \{S_{c'}\}$
end for

$H \leftarrow$ a minimal hitting set for $S$
$\mathcal{I} \leftarrow$ the conjunction of the elements of $H$
return $\mathcal{I}$
```

The obtained conjunction of atomic formulas in $\mathcal{A}$ is an implicant because c satisfies all conjuncts, but for any *false* k-level cell there is at least one conjunct that is not satisfied by c. The implicant is prime since a minimal hitting set was constructed.

Consider the *true* 1-level cell $(-1, -1/\sqrt{2})$ in the cylindrical algebraic decomposition introduced with the exemplary formula ϕ from section 3.1. The set of atomic formulas in $\mathcal{A}$, that evaluate to *true* in the 1-level cell $(-1, -1/\sqrt{2})$, can be determined by substituting the sample point for x. The set described above is $\{2x^2 - 1 > 0,\ x^2 - 1 < 0,\ x < 0\}$. The collection of subsets of atomic formulas in the set above, which do not hold in the *false* k-level cells $(-\infty, 1)$, $\{-1\}$, $\{1/\sqrt{2}\}$, $(1/\sqrt{2}, 1)$, $\{1\}$ respectively $(1, \infty)$, is shown below.

$$\{\{x^2 - 1 < 0\}, \{x^2 - 1 < 0\}, \{2x^2 - 1 > 0, x < 0\},$$
$$\{x < 0\}, \{x^2 - 1 < 0, x < 0\}, \{x^2 - 1 < 0, x < 0\}\}$$

The set $\{x^2 - 1 < 0,\ x < 0\}$ is a hitting set for the collection of subsets given above. Therefore, a prime implicant capturing the 1-level cell $(-1, -1/\sqrt{2})$ is $x^2 - 1 < 0 \wedge x < 0$. The implicant $x^2 - 1 < 0 \wedge x < 0$ also captures the 1-level cells $\{-1/\sqrt{2}\}$ and $(-1/\sqrt{2}, 0)$.

The algorithm above can be used to refine a method that constructs an equivalent quantifier-free formula for a given quantified formula. If a *true* k-level cell is not captured yet, a prime implicant capturing the considered cell is constructed. Since an implicant could possibly capture multiple *true* k-level cells at once, a minimal subset of the set of constructed implicants, such that all *true* cells are captured, is determined. The disjunction of the elements of such a subset is an equivalent quantifier-free formula. The construction of such a minimal subset can be implemented as a minimal hitting set problem. The algorithm shown below, originally introduced by Brown [2], realises the method described above.

Algorithm 3.7.2 Construct an equivalent quantifier-free formula

Input: A cylindrical algebraic decomposition D, a quantified formula ϕ
Output: An equivalent quantifier-free formula ψ

```
$\mathcal{S} \leftarrow \emptyset$
for each true cell $c$ do
    if $c$ is not captured by any implicant in $S$ then
        $\mathcal{I} \leftarrow$ an implicant capturing the cell $c$
        $S \leftarrow S \cup \{\mathcal{I}\}$
    end if
end for

$H \leftarrow$ a minimal subset of $S$ (in terms of capturing true cells)
$\psi \leftarrow$ the disjunction of the elements of $H$
return $\psi$
```

The obtained formula is quantifier-free, since it is in disjunctive normal form. Let a be a point in k-space, *i.e.* an interpretation of the free variables $x_1, \ldots, x_k$. If a satisfies the quantified formula, there is at least one implicant capturing the *true* k-level cell c, in which a is. Thus, a satisfies the obtained formula. If a does not satisfy the quantified formula, a does not satisfy at least one atomic formula in $\mathcal{A}$ for each implicant. Thus, a does not satisfy the obtained formula. Therefore, the obtained formula is equivalent to the given quantified formula.

Above it was described how a prime implicant capturing the 1-level cell $(-1, -{}^1/\sqrt{2})$ in the cylindrical algebraic decomposition introduced with the exemplary formula ϕ from section 3.1 is constructed. The formula $x^2 - 1 < 0 \wedge x < 0$ is a prime implicant capturing the 1-level cell $(-1, -{}^1/\sqrt{2})$. The implicant also captures the *true* 1-level cells $\{-{}^1/\sqrt{2}\}$ and $(-{}^1/\sqrt{2}, 0)$. Analogous, the prime implicant $2x^2 - 1 < 0$ capturing the *true* 1-level cell $\{0\}$ can be constructed. The implicant $2x^2 - 1 < 0$ also captures the *true* 1-level cell $(0, {}^1/\sqrt{2})$. Thus, all *true* 1-level cells are captured by these two implicants. Therefore, the formula shown below is an equivalent quantifier-free formula for the exemplary formula ϕ introduced in section 3.1.

$$\psi := (x^2 - 1 < 0 \wedge x < 0) \vee (2x^2 - 1 < 0)$$

Chapter 4

Implementation & Experimental Results

4.1 Implementation

The purpose of this thesis was to extend the Satisfiability Modulo Theories Real Arithmetic Toolbox (SMT-RAT) by the capability to eliminate quantifiers for NRA-formulas. The toolbox is an open source C++ project maintained by the Theory of Hybrid Systems research group at RWTH Aachen University [9]. It consists of implementations of methods for solving quantifier-free (non-)linear real and integer arithmetic formulas, called modules. In particular, SMT-RAT already provides a CAD module.

However, for this thesis, a slightly modified version of the existing CAD module was implemented providing two methods `void project()` and `void lift()` implementing the projection and lifting phase for the construction of a cylindrical algebraic decomposition. In addition, the modified implementation provides a method to remove a single projection factor, a feature not supported by the existing implementation but needed for the simplification of cylindrical algebraic decompositions. In order to extend SMT-RAT by the capability to eliminate quantifiers for non-linear real arithmetic formulas, a `QE` class was implemented providing a collection of methods implementing the algorithms described in chapter 3 as well as several datastructures storing important information. An example is a `std::map` implementing the assignment of truth values to k-level cells. The class also provides a method implementing a greedy algorithm capable of computing a hitting set for a given collection of sets. In order to speed up the computation, the greedy approach to approximate a minimal hitting set is chosen instead of the computation of an actual minimal hitting set.

In order to describe an input formula, SMT-RAT accepts `.smt2`-files, as specified as in the SMT-LIB [1]. For this thesis, SMT-RAT was extended by a new SMT-LIB command to describe a list of the to be eliminated quantifiers along with the variables quantified by them. Furthermore, the command invokes the execution of the implemented quantifier elimination method.

4.2 Experimental Results

The implementation of the quantifier elimination method described in chapter 3 was tested on a collection of exemplary quantified formulas, carried together by John Wilson [19]. Overall, the implementation was tested on 30 exemplary quantified formulas. The results for 10 of them are presented below. Out of 10, 2 of the quantified formulas are sentences. Thus, for these 2, the result simply is whether the considered sentence is either *true* or *false*, *i.e.* equivalent to either the constant *true* formula $\top$ or the constant *false* formula $\bot$. For the other 8, equivalent quantifier-free formulas are constructed. In addition, 2 out of these 8 formulas were determined to be equivalent to either $\top$ or $\bot$ as well. The formulas constructed by the implemented quantifier elimination method are compared to those constructed by QEPCAD [4]. QEPCAD is an implementation of a quantifier elimination method by partial cylindrical algebraic decomposition. It is originally due to Hong, but extended by many others *e.g.* Brown. An examination of the constructed equivalent quantifier-free formulas shows, that the results produced by QEPCAD are slightly better for some of the examples. Additionally, the formulas constructed by the implemented quantifier elimination method are compared to those constructed by the original quantifier elimination method by cylindrical algebraic decomposition, introduced by Collins [7].

Real Implicitization

The quantified formula considered by the real implicitization problem [12] is shown below.

$$\exists u\ \exists v\ (-x + uv = 0\ \wedge\ -y + uv^2 = 0\ \wedge\ -z + u^2 = 0)$$

The equivalent quantifier-free formula constructed by the implemented quantifier elimination method is shown below. The formula was constructed in 1 second.

$$(z = 0\ \wedge\ y = 0\ \wedge\ -y^2z + x^4 = 0) \vee (-z < 0\ \wedge\ -y^2z + x^4 = 0))$$

The atomic formula $-y^2z + x^4 = 0$ occurs in both disjuncts, it could be factored out in order to obtain a formula similar to the equivalent quantifier-free formula constructed by QEPCAD shown below.

$$z \geq 0\ \wedge\ y^2z - x^4 = 0\ \wedge\ (y = 0\ \vee\ z > 0)$$

The equivalent quantifier-free formula constructed by the original quantifier elimination method is shown below.

$$\begin{gathered}(z = 0\ \wedge\ y = 0\ \wedge\ x = 0\ \wedge\ -y^2z + x^4 = 0) \vee (-y^2z + x^4 = 0\ \wedge\ -z < 0\ \wedge\ y < 0\ \wedge\ x < 0) \vee \\ (-y^2z + x^4 = 0\ \wedge\ -z < 0\ \wedge\ y < 0\ \wedge\ -x < 0) \vee (y = 0\ \wedge\ x = 0\ \wedge\ -y^2z + x^4 = 0\ \wedge\ -z < 0) \vee \\ (-y^2z + x^4 = 0\ \wedge\ -z < 0\ \wedge\ x < 0\ \wedge\ -y < 0) \vee (-y^2z + x^4 = 0\ \wedge\ -z < 0\ \wedge\ -x < 0\ \wedge\ -y < 0)\end{gathered}$$

Termination of Term Rewrite Systems

The quantified formula considered by the termination of term rewrite systems problem [8] is given below. Note that the formula shown below is a sentence.

$$\exists r\ \forall x\ \forall y\ ((r - x < 0)\ \wedge\ ((r - y < 0) \Rightarrow (y^2 - x^2 - 4x^2y - 2x^2y^2 < 0))$$

The sentence was determined to be *false*. The sentence was determined to be *false* in 5 seconds. QEPCAD determined the sentence to be *false* as well.

Parametric Parabola

The quantified formula considered by the parametric parabola problem [6] is shown below.

$$\exists x\ (c + bx + ax^2 = 0)$$

The equivalent quantifier-free formula constructed by the implemented quantifier elimination method is shown below. The formula was constructed in less than a second.

$$\begin{array}{l}
(c = 0) \lor (c < 0\ \land\ b < 0\ \land\ b^2 - 4ac = 0) \lor (c < 0\ \land\ 4ac - b^2 < 0)\ \lor \\
(c < 0\ \land\ -b < 0\ \land\ b^2 - 4ac = 0) \lor (-c \leq 0\ \land\ b < 0\ \land\ b^2 - 4ac = 0) \\
(-c \leq 0\ \land\ 4ac - b^2 < 0) \lor (-c \leq 0\ \land\ -b < 0\ \land\ b^2 - 4ac = 0)
\end{array}$$

The equivalent quantifier-free formula constructed by QEPCAD is shown below.

$$ac - b^2 \leq 0\ \land\ (c = 0\ \lor\ a \neq 0\ \lor\ 4ac - b^2 < 0)$$

The equivalent quantifier-free formula constructed by the original quantifier elimination method is shown below.

$$\begin{array}{l}
(c < 0\ \land\ b < 0\ \land\ a < 0\ \land\ -4ac + b^2 = 0) \lor (c < 0\ \land\ b < 0\ \land\ a < 0\ \land\ 4ac - b^2 < 0)\ \lor \\
(c < 0\ \land\ b < 0\ \land\ 4ac - b^2 < 0\ \land\ a = 0) \lor (c < 0\ \land\ b < 0\ \land\ 4ac - b^2 < 0\ \land\ -a < 0)\ \lor \\
(c < 0\ \land\ 4ac - b^2 < 0\ \land\ -a < 0\ \land\ b = 0) \lor (c < 0\ \land\ a < 0\ \land\ -4ac + b^2 = 0\ \land\ -b < 0)\ \lor \\
(c < 0\ \land\ a < 0\ \land\ 4ac - b^2 < 0\ \land\ -b < 0) \lor (c < 0\ \land\ 4ac - b^2 < 0\ \land\ a = 0\ \land\ -b < 0)\ \lor \\
(c < 0\ \land\ 4ac - b^2 < 0\ \land\ -a < 0\ \land\ -b < 0) \lor (b < 0\ \land\ a < 0\ \land\ 4ac - b^2 < 0\ \land\ c = 0)\ \lor \\
(b < 0 \land 4ac - b^2 < 0\ \land\ a = 0\ \land\ c = 0) \lor (b < 0\ \land\ 4ac - b^2 < 0\ \land\ -a < 0\ \land\ c = 0)\ \lor \\
(a < 0\ \land\ -4ac + b^2 = 0\ \land\ b = 0\ \land\ c = 0) \lor (-4ac + b^2 = 0\ \land\ a = 0\ \land\ b = 0\ \land\ c = 0)\ \lor \\
(-4ac + b^2 = 0\ \land\ -a < 0\ \land\ b = 0\ \land\ c = 0) \lor (a < 0\ \land\ 4ac - b^2 < 0\ \land\ -b < 0\ \land\ c = 0)\ \lor \\
(4ac - b^2 < 0\ \land\ a = 0\ \land\ -b < 0\ \land\ c = 0) \lor (4ac - b^2 < 0\ \land\ -a < 0\ \land\ -b < 0\ \land\ c = 0)\ \lor \\
(b < 0\ \land\ a < 0\ \land\ 4ac - b^2 < 0\ \land\ -c < 0) \lor (b < 0\ \land\ 4ac - b^2 < 0\ \land\ a = 0\ \land\ -c < 0)\ \lor \\
(b < 0\ \land\ 4ac - b^2 < 0\ \land\ -a < 0\ \land\ -c < 0) \lor (b < 0\ \land\ -4ac + b^2 = 0\ \land\ -a < 0\ \land\ -c < 0)\ \lor \\
(a < 0\ \land\ 4ac - b^2 < 0\ \land\ b = 0\ \land\ -c < 0) \lor (a < 0\ \land\ 4ac - b^2 < 0\ \land\ -b < 0\ \land\ -c < 0)\ \lor \\
(4ac - b^2 < 0\ \land\ a = 0\ \land\ -b < 0\ \land\ -c < 0) \lor (4ac - b^2 < 0\ \land\ -a < 0\ \land\ -b < 0\ \land\ -c < 0)\ \lor \\
(-4ac + b^2 = 0\ \land\ -a < 0\ \land\ -b < 0\ \land\ -c < 0)
\end{array}$$

Whitney umbrella

The quantified formula considered by the whitney umbrella problem [6] is shown below.

$$\exists u \; \exists v \; (-x + uv = 0 \;\wedge\; -v + y = 0 \;\wedge\; -z + u^2 = 0)$$

The equivalent quantifier-free formula constructed by the implemented quantifier elimination method is shown below. The formula was constructed in 1 second.

$$(z = 0 \;\wedge\; -x^2 + y^2 z = 0) \vee (-z < 0 \;\wedge\; -x^2 + y^2 z = 0)$$

The atomic formulas $z = 0$ and $-z < 0$ could be merged into $-z \leq 0$ in order to obtain a formula similar to the equivalent quantifier-free formula constructed by QEPCAD shown below.

$$z \geq 0 \;\wedge\; y^2 z - x^2 = 0$$

The equivalent quantifier-free formula constructed by the original quantifier elimination method is shown below.

$$\begin{aligned}
&(z = 0 \;\wedge\; y < 0 \;\wedge\; x = 0 \;\wedge\; -x^2 + y^2 z = 0) \vee (z = 0 \;\wedge\; x = 0 \;\wedge\; -x^2 + y^2 z = 0 \;\wedge\; y = 0) \vee \\
&(z = 0 \;\wedge\; x = 0 \;\wedge\; -x^2 + y^2 z = 0 \;\wedge\; -y < 0) \vee (y < 0 \;\wedge\; -x^2 + y^2 z = 0 \;\wedge\; -z < 0 \;\wedge\; x < 0) \vee \\
&(y < 0 \;\wedge\; -x^2 + y^2 z = 0 \;\wedge\; -z < 0 \;\wedge\; -x < 0) \vee (x = 0 \;\wedge\; -x^2 + y^2 z = 0 \;\wedge\; y = 0 \;\wedge\; -z < 0) \vee \\
&(-x^2 + y^2 z = 0 \;\wedge\; -y < 0 \;\wedge\; -z < 0 \;\wedge\; x < 0) \vee (-x^2 + y^2 z = 0 \;\wedge\; -y < 0 \;\wedge\; -z < 0 \;\wedge\; -x < 0)
\end{aligned}$$

Davenport and Heintz

The quantified formula considered by the Davenport and Heintz problem [8] is given below.

$$\begin{aligned}
\exists c \; \forall b \; \forall a \; ((-a + d = 0 \;\wedge\; -b + c = 0) \vee \\
((-a + c = 0 \;\wedge\; -1 + b = 0) \Rightarrow (-b + a^2 = 0)))
\end{aligned}$$

The formula was determined to be equivalent to $\top$. The result was obtained in 3 seconds. The formula constructed by QEPCAD is the same.

Range of Lower Bounds

The quantified formula considered by the range of lower bounds problem [12] is given below.

$$\forall x\ \forall a\ \forall b\ \forall c\ \exists z\ ((-a < 0\ \wedge\ c+bz+az^2 \neq 0) \Rightarrow (-bx+y-c-ax^2 < 0))$$

The equivalent quantifier-free formula constructed by the implemented quantifier elimination method is shown below. The formula was constructed in 1 second. The formula constructed by the original quantifier elimination method is the same.

$$y = 0\ \vee\ y < 0$$

The equivalent quantifier-free formula constructed by QEPCAD is shown below.

$$y \leq 0$$

Collision

The quantified formula considered by the collision problem [8] is given below. Note that the formula shown below is a sentence.

$$\begin{gathered}\exists t\ \exists x\ \exists y\ (96-17t \leq 0\ \wedge\ -160+17t \leq 0\ \wedge\ -16+17t-16x \leq 0\ \wedge\\ -16-17t+16x \leq 0\ \wedge\ -144+17t-16y \leq 0\ \wedge\\ 112-17t+16y \leq 0\ \wedge\ -1+x^2+t^2-2tx+y^2 \leq 0)\end{gathered}$$

The sentence was determined to be *true*. The sentence was determined to be *true* in 23 seconds. QEPCAD determined the sentence to be *true* as well.

Hong-90

The quantified formula considered by the Hong-90 problem [14] is shown below.

$$\exists a\ \exists b\ (s+r+t = 0\ \wedge\ -a+rt+rs+st = 0\ \wedge\ -b+rst = 0)$$

The equivalent quantifier-free formula constructed by the implemented quantifier elimination method is shown below. The formula was constructed in less than a second. The formula constructed by QEPCAD is the same.

$$s + r + t = 0$$

The equivalent quantifier-free formula constructed by the original quantifier elimination method is shown below.

$$\begin{aligned}
&(s+r+t=0 \ \wedge \ t<0 \ \wedge \ s+t<0 \ \wedge \ s<0 \ \wedge \ -s^2+-st-t^2<0 \ \wedge \ rt+rs+st<0 \ \wedge \ -r<0)\vee\\
&(s+r+t=0 \ \wedge \ t<0 \ \wedge \ s+t<0 \ \wedge \ -s^2-st-t^2<0 \ \wedge \ rt+rs+st<0 \ \wedge \ -r<0 \ \wedge \ s=0)\vee\\
&(s+r+t=0 \ \wedge \ t<0 \ \wedge \ s+t<0 \ \wedge \ -s^2-st-t^2<0 \ \wedge \ rt+rs+st<0 \ \wedge \ -r<0 \ \wedge \ -s<0)\vee\\
&(s+r+t=0 \ \wedge \ t<0 \ \wedge \ -s^2-st-t^2<0 \ \wedge \ rt+rs+st<0 \ \wedge \ -s<0 \ \wedge \ s+t=0 \ \wedge \ r=0)\vee\\
&(s+r+t=0 \ \wedge \ t<0 \ \wedge \ -s^2-st-t^2<0 \ \wedge \ rt+rs+st<0 \ \wedge \ -s<0 \ \wedge \ -s-t<0 \ \wedge \ r<0)\vee\\
&(s+r+t=0 \ \wedge \ s+t<0 \ \wedge \ s<0 \ \wedge \ -s^2-st-t^2<0 \ \wedge \ rt+rs+st<0 \ \wedge \ -r<0 \ \wedge \ t=0)\vee\\
&(s+r+t=0 \ \wedge \ s=0 \ \wedge \ s+t=0 \ \wedge \ r=0 \ \wedge \ t=0 \ \wedge \ s^2+st+t^2=0 \ \wedge \ rt+rs+st=0)\vee\\
&(s+r+t=0 \ \wedge \ -s^2-st-t^2<0 \ \wedge \ rt+rs+st<0 \ \wedge \ -s<0 \ \wedge \ -s-t<0 \ \wedge \ r<0 \ \wedge \ t=0)\vee\\
&(s+r+t=0 \ \wedge \ s+t<0 \ \wedge \ s<0 \ \wedge \ -s^2-st-t^2<0 \ \wedge \ rt+rs+st<0 \ \wedge \ -r<0 \ \wedge \ -t<0)\vee\\
&(s+r+t=0 \ \wedge \ s<0 \ \wedge \ -s^2-st-t^2<0 \ \wedge \ rt+rs+st<0 \ \wedge \ s+t=0 \ \wedge \ r=0 \ \wedge \ -t<0)\vee\\
&(s+r+t=0 \ \wedge \ s<0 \ \wedge \ -s^2-st-t^2<0 \ \wedge \ rt+rs+st<0 \ \wedge \ -s-t<0 \ \wedge \ r<0 \ \wedge \ -t<0)\vee\\
&(s+r+t=0 \ \wedge \ -s^2-st-t^2<0 \ \wedge \ rt+rs+st<0 \ \wedge \ s=0 \ \wedge \ -s-t<0 \ \wedge \ r<0 \ \wedge \ -t<0)\vee\\
&(s+r+t=0 \ \wedge \ -s^2-st-t^2<0 \ \wedge \ rt+rs+st<0 \ \wedge \ -s<0 \ \wedge \ -s-t<0 \ \wedge \ r<0 \ \wedge \ -t<0)
\end{aligned}$$

Simplified YangXia

The quantified formula considered by the simplified YangXia problem [5] is shown below.

$$\begin{aligned}
\exists b \ (&b \neq 0 \ \wedge \ -R<0 \ \wedge \ -b<0 \ \wedge \ -h<0 \ \wedge\\
&16h^4R^4 + a^4b^4 - 8h^2R^2b^4 - 2a^2b^6 - 8a^2h^2R^2b^2 + 4a^2h^2b^4 + b^8 = 0 \ \wedge\\
&2hRb - ab^2 - b^3 < 0 \ \wedge \ -hRb < 0 \ \wedge \ -2hRb - ab^2 + b^3 < 0 \ \wedge\\
&ab^2 - 2hRb - b^3 < 0)
\end{aligned}$$

The equivalent quantifier-free formula constructed by the implemented quantifier elimination method is shown below. The formula was constructed in 2

minutes and 8 seconds.

$$\begin{aligned}
&(-R < 0 \ \wedge \ -a < 0 \ \wedge \ a^2 + 4h^2 - 8hR < 0) \ \vee \\
&(-R < 0 \ \wedge \ -a < 0 \ \wedge \ -a^2 - 4h^2 + 8hR = 0) \ \vee \\
&(-R < 0 \ \wedge \ -h < 0 \ \wedge \ h - R < 0 \ \wedge \ -a + 2R = 0) \ \vee \\
&(-R < 0 \ \wedge \ -h < 0 \ \wedge \ h - R < 0 \ \wedge \ -a < 0 \ \wedge \ a - 2R < 0))
\end{aligned}$$

The equivalent quantifier-free formula constructed by QEPCAD is shown below.

$$a > 0 \ \wedge \ h > 0 \ \wedge \ 2R - a \geq 0 \ \wedge \ (2h - a < 0 \ \vee \ 8hR - 4h^2 - a^2 \geq 0)$$

The execution of the original quantifier elimination method timed out.

Cyclic-3

The quantified formula considered by the collision problem [19] is given below.

$$\exists b \ \exists a \ (b + a + c = 0 \ \wedge \ ac + ab + bc = 0 \ \wedge \ -1 + abc = 0)$$

The formula was determined to be equivalent to $\bot$. The result was obtained in 38 seconds. The formula constructed by QEPCAD is the same.

For 20 out of the 30 considered exemplary quantified formulas, the execution of the implemented quantifier elimination method timed out. An examination of these examples revealed, that large multivariate polynomials occurred in these exemplary formulas, which imply costly computations in the projection and lifting phase for the construction of the cylindrical algebraic decomposition defined by these multivariate polynomials.

Chapter 5

Conclusion

In this thesis, at first, the preliminaries for quantifier elimination by cylindrical algebraic decomposition were provided. The non-linear real arithmetic was introduced, the fragment of the first-order logic for which quantifier elimination was considered. The term cylindrical algebraic decomposition was defined, the concept the presented quantifier elimination method is based on. Furthermore, a method to construct a cylindrical algebraic decomposition for a given set of multivariate polynomials was described. The hitting set problem was presented, the concept used in several steps of the presented quantifier elimination method in the hope of constructing a simpler equivalent quantifier-free formula. Subsequently, a quantifier elimination method was described. A method to determine the truth values of a quantified formula, using the cylindrical algebraic decomposition defined by the multivariate polynomials occuring in the formula, was refined. Based on the assignment of truth values to cells, a method to simplify a cylindrical algebraic decomposition was presented. Furthermore, the concept of signatures was introduced as well as a method using the signatures and the truth values to test whether a cylindrical algebraic decomposition is projection-definable, a property necessary to construct an actual equivalent quantifier-free formula based on cylindrical algebraic decomposition. In addition, a method was refined to assure projection-definability. Finally, a method to construct an equivalent quantifier-free formula, using the cylindrical algebraic decomposition, defined by the multivariate polynomials occurring in a quantified formula, was described.

For this thesis, the described quantifier elimination method was implemented in order to extend SMT-RAT by the capability to eliminate quantifiers. Some remarks on the implementation were made as well as experimental results were presented. Significant improvements in terms of the simplicity of the constructed equivalent quantifier-free formulas, compared to the formulas constructed by the original quantifier elimination method, were observed. However, the equivalent quantifier-free formulas constructed by the implemented quantifier elimination method were observed to be slightly inferior to those constructed by other modern implementations such as QEPCAD. For future work, additional simplifications on the constructed equivalent quantifier-free formulas could be considered in order to improve the simplicity of the formulas even more.

Bibliography

[1] C. Barrett, P. Fontaine, and C. Tinelli. The Satisfiability Modulo Theories Library (SMT-LIB). `www.SMT-LIB.org`, 2016.

[2] C. W. Brown. *Solution Formula Construction for Truth Invariant Cad's*. PhD thesis, Newark, DE, USA, 1999. AAI9927664.

[3] C. W. Brown. Improved projection for cylindrical algebraic decomposition. *J. Symb. Comput.*, 32(5):447–465, Nov. 2001.

[4] C. W. Brown. Qepcad b: A program for computing with semi-algebraic sets using cads. *SIGSAM Bull.*, 37(4):97–108, Dec. 2003.

[5] C. W. Brown and C. Gross. Efficient preprocessing methods for quantifier elimination. In V. G. Ganzha, E. W. Mayr, and E. V. Vorozhtsov, editors, *Computer Algebra in Scientific Computing*, pages 89–100, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[6] C. Chen, M. M. Maza, B. Xia, and L. Yang. Computing cylindrical algebraic decomposition via triangular decomposition. In *Proceedings of the 2009 International Symposium on Symbolic and Algebraic Computation*, ISSAC '09, pages 95–102, New York, NY, USA, 2009. ACM.

[7] G. E. Collins. Quantifier elimination for real closed fields by cylindrical algebraic decompostion. In H. Brakhage, editor, *Automata Theory and Formal Languages 2nd GI Conference Kaiserslautern, May 20–23, 1975*, pages 134–183, Berlin, Heidelberg, 1975. Springer Berlin Heidelberg.

[8] G. E. Collins and H. Hong. Partial cylindrical algebraic decomposition for quantifier elimination. *J. Symb. Comput.*, 12(3):299–328, Sept. 1991.

[9] F. Corzilius, G. Kremer, S. Junges, S. Schupp, and E. Ábrahám. Smt-rat: An open source c++ toolbox for strategic and parallel smt solving. In *SAT*, 2015.

[10] M. Coste and M. F. Roy. Thom's lemma, the coding of real algebraic numbers and the computation of the topology of semi-algebraic sets. *J. Symb. Comput.*, 5(1-2):121–129, Feb. 1988.

[11] J. H. Davenport and J. Heintz. Real quantifier elimination is doubly exponential. *J. Symb. Comput.*, 5(1-2):29–35, Feb. 1988.

[12] A. Dolzmann, A. Seidl, and T. Sturm. Efficient projection orders for cad. In *Proceedings of the 2004 International Symposium on Symbolic and Algebraic Computation*, ISSAC '04, pages 111–118, New York, NY, USA, 2004. ACM.

[13] H. Hong. Simple solution formula construction in cylindrical algebraic decomposition based quantifier elimination. In *Papers from the International Symposium on Symbolic and Algebraic Computation*, ISSAC '92, pages 177–188, New York, NY, USA, 1992. ACM.

[14] H. Hong. An improvement of the projection operator in cylindrical algebraic decomposition. In B. F. Caviness and J. R. Johnson, editors, *Quantifier Elimination and Cylindrical Algebraic Decomposition*, pages 166–173, Vienna, 1998. Springer Vienna.

[15] M. Jirstrand. *Cylindrical Algebraic Decomposition: An Introduction*. LiTH-ISY-R. Linköpings university, 1995.

[16] S. McCallum. An improved projection operation for cylindrical algebraic decomposition of three-dimensional space. *Journal of Symbolic Computation*, 5(1):141 – 161, 1988.

[17] A. Tarski. A decision method for elementary algebra and geometry. In B. F. Caviness and J. R. Johnson, editors, *Quantifier Elimination and Cylindrical Algebraic Decomposition*, pages 24–84, Vienna, 1998. Springer Vienna.

[18] T. Viehmann. *Comparing Different Projection Operators in the Cylindrical Algebraic Decomposition for SMT Solving*. PhD thesis, 2016.

[19] D. Wilson. Real geometry and connectedness via triangular description: Cad example bank, November 2012.