Tractability conditions for numeric CSPs[☆]

Peter Jonsson

Department of Computer and Information Science, Linköpings universitet
SE-58183 Linköping, Sweden

Johan Thapper[*]

Laboratoire d'Informatique Gaspard-Monge, Université Paris-Est Marne-la-Vallée
5 boulevard Descartes, F-77420 Champs-sur-Marne, France

Abstract

The computational complexity of the constraint satisfaction problem (CSP) with semilinear relations over the reals has gained recent attraction. As a result, its complexity is known for all finite sets of semilinear relations containing the relation $R_+ = \{(x, y, z) \in \mathbb{R}^3 \mid x + y = z\}$. We consider larger and more expressive classes of relations such as *semialgebraic* and *o-minimal* relations. We present a general result for characterising computationally hard fragments and, under certain side conditions, this result implies that polynomial-time solvable fragments are only to be found within two limited families of sets of relations. In the setting of semialgebraic relation, our result takes on a simplified form and we provide a full complexity classification for constraint languages that consist of algebraic varieties. Full classifications like the one obtained here for algebraic varieties or the one for semilinear relations appear to be rare and we discuss several barriers for obtaining further such results. These barriers have strong connections with well-known open problems concerning the complexity of various restrictions of convex programming.

Keywords: Constraint satisfaction problem, Semialgebraic constraint, o-minimality, Algebraic variety, Computational complexity

1. Introduction

1.1. The constraint satisfaction problem

The *constraint satisfaction problem* (CSP) is an important computational problem in many areas of computer science and mathematics. In this problem, we are given a set of variables that take their values from a (finite or infinite) *domain*. The assignments to the variables are further subjected to a set of *constraints*. A constraint is defined by requiring that a tuple of variables belongs to some specified relation. The question is whether the variables can be assigned values such that all constraints are satisfied. Since even the general finite-domain CSP is NP-hard, the complexity of CSPs is often studied by introducing an additional parameter, a set Γ of allowed relations, known as a *constraint language* (or *template*). This leads to a problem $\text{CSP}(\Gamma)$ where the relations of all constraints in the input are required to come from Γ. This way of parameterizing constraint satisfaction problems has proved to be very fruitful for both finite and infinite domains. In the sequel, when we talk about a CSP, we will mean a problem $\text{CSP}(\Gamma)$ for some fixed Γ.

The complexity of finite-domain CSPs has been extensively investigated, beginning with Schaefer [31]. This has lead to the development of a set of standard tools, including the powerful universal-algebraic approach [12]. Much of this effort has been devoted to the *Feder-Vardi Dichotomy Conjecture* [14] which posits that every finite-domain CSP is either polynomial-time solvable or NP-complete. This conjecture has recently been resolved in the positive in two independent works [11, 35].

Infinite-domain CSPs, on the other hand, constitute a much more diverse set of problems: *Every* computational problem is polynomial-time equivalent to some infinite-domain CSP [4]. Obtaining a full understanding of their computational complexity is thus too ambitious. Instead, restricted classes of problems are studied. For example, one can consider CSPs over numeric domains, such as the reals, the rationals, or the integers, with constraints derived from arithmetic operations and the natural order on these domains. We refer the reader to [10] for a recent survey of such CSPs.

In this article, we study CSPs over the reals with constraints based on *o-minimal* and in particular *semialgebraic* relations. The point of departure of our investigation is recent progress on the project of classifying *semilinear* CSPs. We describe this next.

1.2. Semilinear CSPs

A relation is semilinear if it can be written as a finite union of finite intersections of open and closed half-spaces over, for instance, the reals, the rationals or the integers. Let $SL_X[Y]$ denote the set of semilinear relations

with domain X and coefficients in Y. We will mainly consider $SL_\mathbb{R}[\mathbb{Q}]$; this set of relations equals the set of first-order definable relations over $\{+, \leq, \{1\}\}$ [15].

Characterising the polynomial-time solvable cases of $SL_\mathbb{R}[\mathbb{Q}]$ is a challenging task. The construction presented in [21, Section 6.3] proves the following: for every finite constraint language Γ over a finite domain, there exists a finite $\Gamma' \subseteq SL_\mathbb{R}[\mathbb{Q}]$ such that $\mathrm{CSP}(\Gamma)$ and $\mathrm{CSP}(\Gamma')$ are polynomial-time equivalent problems. Hence, the classification task of semilinear CSPs is inextricably linked to the classification task of finite-domain CSPs, and thereby to the Feder-Vardi dichotomy conjecture. We also observe that the complexity of every finite *temporal* constraint language would be determined as a by-product of a full classification of $SL_\mathbb{R}[\mathbb{Q}]$. A temporal constraint language is a constraint language that is first-order definable in $(\mathbb{Q}; <)$. The complexity of temporal constraint languages has been fully determined [7] and the polynomial-time solvable cases fall into nine different categories. The proof is complex and makes heavy use of the universal-algebraic approach.

One way of obtaining classes of semilinear constraint languages that are more manageable is to restrict attention to expansions of certain natural sets of relations. One such choice is the set $\Gamma_{lin} = \{R_+, \leq, \{1\}\}$, where $R_+ = \{(x, y, z) \in \mathbb{R}^3 \mid x + y = z\}$. Studying the computational complexity of expansions of Γ_{lin} is well motivated by the fact that $\mathrm{CSP}(\Gamma_{lin})$ is polynomial-time many-one equivalent to the linear programming feasibility problem [5]. This direction has been pursued in a number of recent works [5, 6, 21, 22]. A complete classification for semilinear expansions of Γ_{lin} was obtained in [5]. This was generalised to semilinear expansions of $\{R_+\}$ by Jonsson and Thapper [23]. Here, we describe an intermediate classification for semilinear expansions of $\{R_+, \{1\}\}$ [22].

A relation is *primitive positive* (pp) definable from a constraint language Γ if it can be expressed using existential quantification over conjunctions of atoms. The importance of pp-definability is explained by Lemma 4 below.

We say that a relation $R \subseteq \mathbb{R}^k$ is *essentially convex* if for all $p, q \in R$ there are only finitely many points on the line segment between p and q that are not in R. A BNU (for *bounded*, *non-constant*, and *unary*) is a bounded unary relation that contains more than one point.

Theorem 1 (Jonsson and Thapper [22])**.** Let $\{R_+, \{1\}\} \subseteq \Gamma \subseteq SL_\mathbb{R}[\mathbb{Q}]$ be a finite constraint language. If

1. Γ contains a relation that is not essentially convex, and
2. Γ can primitive positively define a BNU relation,

then $\mathrm{CSP}(\Gamma)$ is NP-hard. Otherwise, $\mathrm{CSP}(\Gamma)$ is tractable.

Our goal will be to extend Theorem 1 as far as possible to *semialgebraic* and *o-minimal* constraint languages. We say that a relation $R \subseteq \mathbb{R}^k$ is *semialgebraic* if it can be first-order defined in $\{+, \cdot, \leq\}$ with parameters in $\mathbb{R}$. The classical Tarski-Seidenberg theorem [33] implies that semialgebraic constraint languages have a clear geometric interpretation: every semialgebraic relation can be written as a finite union of solution sets of strict and non-strict polynomial inequalities. Semialgebraic relations appear in many different contexts within mathematics and computer science (cf. the textbook by Basu, Pollack, and Roy [3]).

A constraint language Γ that contains a total ordering of the domain $\mathbb{R}$ is called *o-minimal* if every first-order definable set (with parameters from $\mathbb{R}$) can be represented as a union of finitely many intervals and points. Structures that are o-minimal have been studied thoroughly in model theory (cf. van Den Dries [34] and Macpherson [26]). A well-known class of relations that give rise to o-minimal but not semialgebraic constraint languages is the set of *Semi-Pfaffian* relations, cf. Khovanskii [24]. A concrete example of a semi-Pfaffian constraint language that is not semialgebraic is $\mathrm{SA}_{\mathbb{R}}[\mathbb{R}] \cup \{(x, y) \in \mathbb{R}^2 \mid y = e^x\}$.

1.3. Contributions

Let $LE_X[Y]$ be the set of relations defined by linear equations over X with coefficients in Y. In Section 3, we prove the following theorem.

Theorem 2. Let $\mathbb{F}$ be any subfield of $\mathbb{R}$. Let $LE_{\mathbb{R}}[\mathbb{F}] \subseteq \Gamma$ be a constraint language over $\mathbb{R}$ such that every primitive positively definable unary relation over Γ can be written as a finite union of intervals with endpoints in $\mathbb{F}$. Assume that the following hold:

1. Γ contains a relation that is not essentially convex and this is witnessed by points with coordinates in $\mathbb{F}$, and
2. Γ can primitive positively define a BNU relation.

Then, CSP(Γ) is NP-hard.

Theorem 2 narrows the possible tractable fragments dramatically. In fact, if we are considering constraint languages that contain a sufficiently large set of linear equations and behave well with respect to pp-definable unary relations, then one should only search for tractable fragments in the following two cases.

1. Γ only contains relations that are essentially convex, or
2. Γ cannot primitive positively define a BNU relation.

These two cases are analogous to the tractable cases of Theorem 1 for semilinear relations. Hence, we cannot expect to identify any profoundly new tractable classes when turning our attention from semilinear relations to more general classes of o-minimal relations. Furthermore, even though we know where candidates for tractable fragments are to be found, we do not expect it to be easy to actually identify concrete tractable constraint languages. Consider the first case: it includes many problems that are concerned with convex objects. This is a very well-studied topic, especially in connection with optimisation. Despite this, the complexity of many such problems remains a mystery: an important and well-known example is the feasibility problem for semidefinite programming. The second case is less well-studied; as far as we know, the only algorithm exploiting this property works with semilinear constraints and involves computing affine hulls [22]. This is tractable in the semilinear setting, but it is not clear that the same approach could be applied to non-linear relations. A more thorough discussion concerning these topics can be found in Section 5.

We exemplify the use of Theorem 2 on semialgebraic constraint languages by showing that semialgebraic constraint languages containing the set $LE_{\mathbb{R}}[\mathbb{A}]$ (where $\mathbb{A}$ denotes the set of algebraic numbers) satisfy the preconditions of Theorem 2. We can thus conclude that there are only two families of constraint languages, analogous to the cases (1) and (2) above, that may contain tractable languages, and that exactly identifying these tractable languages may be difficult. However, there are natural subclasses of semialgebraic relations where full complexity classifications can indeed be obtained by using available techniques. We examine such a class in Section 4. An *algebraic variety* is a relation R that can be defined as the zero-locus of a set, $p_1, \ldots, p_m$, of polynomials in k real variables with coefficients from $\mathbb{A}$:

$$R = \{x \in \mathbb{R}^k \mid p_1(x) = 0 \wedge \cdots \wedge p_m(x) = 0\}.$$

Algebraic varieties are the fundamental objects studied in algebraic geometry, and are consequently of great general interest. Using the hardness result for semialgebraic constraint languages from Section 3 and an algorithm for linear programming with algebraic coefficients by Adler and Beling [1], we present a complete complexity classification of CSP(Γ), when $LE_{\mathbb{R}}^0[\mathbb{A}] \subseteq \Gamma$ and Γ contains algebraic varieties only. Here, $LE_{\mathbb{R}}^0[\mathbb{A}]$ denotes the set of homogeneous linear equations with coefficients in $\mathbb{A}$.

2. Preliminaries

This section provides the definitions and basic results that are needed in the rest of the article. It is divided into three parts. We are concerned

with constraint languages, CSPs, and pp-definitions in the first part, with relations over an ordered structure in general in the second part, and with semialgebraic relations in particular in the last part.

2.1. Constraint satisfaction problems

Let Γ be a set of finitary relations over some domain D (which will usually be infinite). We refer to Γ as a *constraint language*. A first-order formula is called *primitive positive* if it is of the form $\exists x_1, \ldots, x_n. \psi_1 \wedge \cdots \wedge \psi_m$, where ψ_i are formulas of the form $x = y$ or $R(x_{i_1}, \ldots, x_{i_k})$ with R the relation symbol for a k-ary relation from Γ. We call such a formula a *pp-formula*. The conjuncts in a pp-formula Φ are also called the *constraints* of Φ. The *constraint satisfaction problem for* Γ (CSP(Γ) for short) is the computational problem to decide whether a given primitive positive sentence Φ is true in Γ or not.

Definition 3. We say that CSP(Γ) is *globally NP-hard* if CSP(Γ) is NP-hard. We say that CSP(Γ) is *locally NP-hard* if CSP(Γ') is NP-hard for *some* finite $\Gamma' \subseteq \Gamma$. Analogously, we say that CSP(Γ) is *globally tractable* if CSP(Γ) is polynomial-time solvable, and that CSP(Γ) is *locally tractable* if CSP(Γ') is polynomial-time solvable for *all* finite $\Gamma' \subseteq \Gamma$.

We will exclusively prove local NP-hardness results in this article so we equate "NP-hard" with "locally NP-hard". The actual representation of relations is not important when proving local NP-hardness while the choice of representation may be crucial when proving global NP-hardness results.[1]

A relation R is said to be *pp-definable from* a constraint language Γ if there exists a primitive positive formula φ over Γ such that

$$R = \{(x_1, \ldots, x_k) \in D^k \mid \Gamma \models \varphi(x_1, \ldots, x_k)\}.$$

In this case, we will also write $R(x_1, \ldots, x_k) \equiv \varphi(x_1, \ldots, x_k)$. Let $\langle \Gamma \rangle$ denote the set of all relations that are pp-definable in Γ. The following straightforward but important result explains the importance of primitive positive definability. We will use it extensively in the sequel without making explicit references to it.

Lemma 4 (Jeavons [20])**.** Let Γ be a constraint language and R a relation pp-definable from Γ. Then CSP($\Gamma \cup \{R\}$) polynomial-time reduces to CSP(Γ).

---

[1]One example is the problem MONOTONE TVPI INTEGER PROGRAMMING which is in P with a unary representation [19], but NP-complete with a binary representation [25].

We will make extensive use of constructions involving linear equations. Recall that $LE_X[Y]$ denotes the set of relations defined by linear equations over the set X with coefficients in Y. We similarly let $LE_X^0[Y]$ denote the set of relations defined by *homogeneous* linear equations over the set X with parameters from Y.

Lemma 5 (Bodirsky, Jonsson, and von Oertzen [5, Lemma 2.11]). *Let $r_0, r_1, \ldots, r_k \in \mathbb{Q}$. The relation*

$$\{(x_1, \ldots, x_k) \in \mathbb{R}^k \mid r_1 x_1 + \ldots + r_k x_k = r_0\}$$

is pp-definable in $\{R_+, \{1\}\}$. If $r_0 = 0$, then the relation is pp-definable in $\{R_+\}$. The pp-definition can in both cases be computed in polynomial time.

From Lemma 5, we conclude that $LE_{\mathbb{R}}[\mathbb{Q}] = \langle\{R_+, \{1\}\}\rangle$ and $LE_{\mathbb{R}}^0[\mathbb{Q}] = \langle\{R_+\}\rangle$.

Given a relation $R \subseteq \mathbb{R}^k$ and two distinct points $a, b \in \mathbb{R}^k$, we define the relation $\mathcal{L}_{R,a,b}$ as the parameterisation of the intersection between the relation R and a line through the points a and b. Formally,

$$\mathcal{L}_{R,a,b}(y) \equiv \exists x_1, \ldots, x_k. R(x_1, \ldots, x_k) \wedge \bigwedge_{i=1}^k x_i = (1 - y) \cdot a_i + y \cdot b_i.$$

Note that $\mathcal{L}_{R,a,b}$ is pp-definable in $LE_{\mathbb{R}}[\mathbb{R}] \cup \{R\}$. Furthermore, if $\mathbb{F}$ is a subfield of $\mathbb{R}$, and $a, b \in \mathbb{F}^k$, then $\mathcal{L}_{R,a,b}$ is pp-definable in $LE_{\mathbb{R}}[\mathbb{F}] \cup \{R\}$. In particular, if $a, b \in \mathbb{Q}^k$, then $\mathcal{L}_{R,a,b}$ is pp-definable in $LE_{\mathbb{R}}[\mathbb{Q}] \cup \{R\}$ and, consequently, in $\{R_+, \{1\}, R\}$ by Lemma 5. Also note that if $a = (0, \ldots, 0)$, then $\mathcal{L}_{R,a,b}$ is pp-definable in $LE_{\mathbb{R}}^0[\cdot] \cup \{R\}$.

2.2. Relations over a totally ordered domain

Let $(\mathbb{R}; \leq, \ldots)$ be a totally ordered structure. By an *interval* we mean either an open, half-open, or closed interval (with respect to the total order) containing at least one point. Given elements $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$, $a \leq b$, we let $[\![a, b]\!]$ denote either of the intervals $[a, b], (a, b], [a, b)$, or (a, b) if $a < b$, and $\{a\}$ if $a = b$.

For a subfield $\mathbb{F} \subseteq \mathbb{R}$, we let $\mathcal{O}_{\mathbb{R},\mathbb{F}}$ be the set of all unary relations $U \subseteq \mathbb{R}$ such that there are finitely many intervals $I_1, \ldots, I_m$ with endpoints in $\mathbb{R} \cup \{-\infty, \infty\}$ such that $U = I_1 \cup \cdots \cup I_m$. When $\mathbb{F} = \mathbb{R}$, we simply write $\mathcal{O}_{\mathbb{R}}$. The structure $(\mathbb{R}; \leq, \ldots)$ is o-*minimal* (see, for instance, Marker [27]) if every unary relation that is first-order definable in $(\mathbb{R}; \leq, \ldots)$ (with parameters from $\mathbb{R}$) is a member of $\mathcal{O}_{\mathbb{R}}$.

A relation $R \subseteq \mathbb{R}^k$ is *convex* if for all $p, q \in R$, R contains all points on the line segment between p and q. We say that R is *essentially convex* if for

all $p, q \in R$, there are only finitely many points on the line segment between p and q that are not in R. We say that R *excludes an interval* if there are $p, q \in R$ and real numbers $0 < \delta_1 < \delta_2 < 1$ such that $(1 - y) \cdot p + y \cdot q \notin R$ whenever $\delta_1 \leq y \leq \delta_2$.

The following lemma connects essential convexity and the exclusion of intervals.

Lemma 6. Let $\mathbb{F}$ be a subfield of $\mathbb{R}$, and let R be a relation over $\mathbb{R}$ such that every unary relation in $\langle LE_{\mathbb{R}}[\mathbb{F}] \cup \{R\} \rangle$ is a member of $\mathcal{O}_{\mathbb{R},\mathbb{F}}$. Let $p, q \in R$ be distinct points with coefficients in $\mathbb{F}$. The following are equivalent:

1. p and q witness that R is not essentially convex.
2. p and q witness that R excludes an interval.

Proof. We prove the non-trivial implication (1) $\implies$ (2). Let k be the arity of the relation R. Let $p, q \in R \cap \mathbb{F}^k$ be such that an infinite number of points on the line segment between p and q are not in R. Define the unary relation $U = \mathcal{L}_{R,p,q} \in \langle LE_{\mathbb{R}}[\mathbb{F}] \cup \{R\} \rangle$. Since every unary relation in $\langle LE_{\mathbb{R}}[\mathbb{F}] \cup \{R\} \rangle$ is a member of $\mathcal{O}_{\mathbb{R},\mathbb{F}}$, we know that U can be written as $\bigcup_{i=1}^{m} I_i$, where $I_i = [\![s_i, t_i]\!]$, for $1 \leq i \leq m$, and $-\infty < s_1 \leq t_1 \leq s_2 \leq t_2 \leq \cdots \leq s_m \leq t_m < \infty$. Clearly, $0, 1 \in U$ but an infinite number of points on the line segment between 0 and 1 are not in U, so there exists $1 \leq j < m$ such that $0 < t_j < s_{j+1} < 1$. Let $\delta = s_{j+1} - t_j > 0$. Then, $p, q \in R$ and for $\delta_1 = t_i + \delta/3$ and $\delta_2 = s_{i+1} - \delta/3$, we have that $\delta_1 \leq y \leq \delta_2$ implies $y \notin U$ which in turn implies $(1-y) \cdot p + y \cdot q \notin R$. It follows that p and q witness an excluded interval in R. $\qquad \square$

2.3. Semialgebraic relations

We consider first-order formulas in the language L of ordered fields, that is, the set of first-order formulas with atoms being polynomial equations and inequalities with coefficients from a given field. We refer to such a formula as an *L-formula*.

A relation $R \subseteq X^k$ is a member of $SA_X[Y]$ if and only if $R = \{(x_1, \ldots, x_k) \in X^k \mid \varphi(x_1, \ldots, x_k)\}$ where φ is an L-formula with free variables $x_1, \ldots, x_k$ and with coefficients taken from the set Y.

Tarski [33] has shown the following important result.

Theorem 7 ([17, Theorem 33.1.1])**.** Let φ be an L-formula interpreted over the real numbers and with real coefficients. There is a quantifier-free formula ψ logically equivalent to φ. If φ involves only rational coefficients, then so does the formula ψ.

The following well-known result is a consequence of Theorem 7.

Theorem 8. $SA_\mathbb{R}[\mathbb{R}]$ is o-minimal.

Lemma 9. $SA_\mathbb{R}[\mathbb{A}] = SA_\mathbb{R}[\mathbb{Q}]$.

Proof. Arbitrarily choose $a \in \mathbb{A}$. By definition, a is the kth smallest root of some non-zero polynomial $p(x)$ with rational coefficients. One may thus define an L-formula $\psi_a(x)$ with coefficients in $\mathbb{Q}$ such that $\psi_a(x) \equiv x = a$. Thus, every L-formula with coefficients in $\mathbb{A}$ can be rewritten as an equivalent L-formula with coefficients in $\mathbb{Q}$. This implies that $SA_\mathbb{R}[\mathbb{A}] = SA_\mathbb{R}[\mathbb{Q}]$. $\qquad\square$

Lemma 10. Every unary relation in $SA_\mathbb{R}[\mathbb{A}]$ is a member of $\mathcal{O}_{\mathbb{R},\mathbb{A}}$.

Proof. Let U be an arbitrary unary relation in $SA_\mathbb{R}[\mathbb{A}]$. The relation U is a member of $SA_\mathbb{R}[\mathbb{Q}]$ by Lemma 9, and there exists a quantifier-free L-formula $\psi(x)$ in one variable with coefficients in $\mathbb{Q}$ such that $U = \{x \in \mathbb{R} \mid \psi(x)\}$ by Theorem 7. We may without loss of generality assume that $\psi(x)$ is given in disjunctive normal form. It is easy to verify that each clause holds on a finite union of intervals $[\![a, b]\!]$ with $a, b \in \mathbb{A} \cup \{-\infty, \infty\}$. Thus, the same holds for the formula ψ, so $U \in \mathcal{O}_{\mathbb{R},\mathbb{A}}$. $\qquad\square$

Note that Theorem 8 combined with Lemma 6 implies that a relation in $SA_\mathbb{R}[\mathbb{R}]$ excludes an interval if and only if it is not essentially convex. Furthermore, Lemma 10 combined with Lemma 6 implies that a relation in $SA_\mathbb{R}[\mathbb{A}]$ excludes an interval with witnesses in $\mathbb{A}$ if and only if it not essentially convex with witnesses in $\mathbb{A}$.

3. The main hardness result

In this section, we prove our main hardness result in Section 3.1 and exemplify its use on semialgebraic constraint languages in Section 3.2.

3.1. Proof of hardness result

We begin by providing two auxiliary results.

Lemma 11 (Jonsson and Thapper [23, Lemma 10])**.** Let $\{-1, 1\} \subseteq T \subseteq (-\frac{3}{2}, -\frac{1}{2}) \cup (\frac{1}{2}, \frac{3}{2})$. Then, $\mathrm{CSP}(\{R_+, T\})$ is NP-hard.

Lemma 12. Let $\mathbb{F}$ be a subfield of $\mathbb{R}$ and $S \subseteq \mathbb{R}$ be a BNU in $\mathcal{O}_{\mathbb{R},\mathbb{F}}$ that excludes an interval. Then, $\mathrm{CSP}(LE_\mathbb{R}[\mathbb{F}] \cup \{S\})$ is NP-hard.

Proof. The proof reduces to an application of Lemma 11 by intersecting two appropriately affinely transformed copies of S. The details follow.

Let $S = \bigcup_{i=1}^{m} I_i$, where $I_i = [\![s_i, t_i]\!]$, for $1 \leq i \leq m$, and $-\infty < s_1 \leq t_1 \leq s_2 \leq t_2 \leq \cdots \leq s_m \leq t_m < \infty$.

The relation S excludes an interval, so there is an index $1 \leq j < m$ such that $t_j < s_{j+1}$. Let $\delta_1 = s_{j+1} - t_j > 0$ and let $a, b \in S \cap \mathbb{F}$ be such that $a \in s_j$, $b \in s_{j+1}$, and $\delta_1 \leq b - a < \frac{5}{4}\delta_1$. Note that $[a + \frac{\delta_1}{4}, b - \frac{\delta_1}{4}] \cap S = \varnothing$.

Let $\delta_2 = t_m - s_1$ and let $c, d \in S \cap \mathbb{F}$ be such that $c \in s_1$, $d \in s_m$, and $\frac{4}{5}\delta_2 < d - c \leq \delta_2$. Note that $S \subseteq (c - \frac{\delta_2}{5}, d + \frac{\delta_2}{5})$.

Define the following unary relations in $LE_{\mathbb{R}}[\mathbb{F}] \cup \{S\}$:

$$T_1(x) \equiv \exists y. S(y) \wedge y = \frac{b-a}{2}x + \frac{b+a}{2} \tag{1}$$

$$T_2(x) \equiv \exists y. S(y) \wedge y = \frac{d-c}{2}x + \frac{d+c}{2} \tag{2}$$

The relations T_1 and T_2 satisfy the following properties:

- $\{-1, 1\} \subseteq T_1 \cap T_2$;

- $[-\frac{1}{2}, \frac{1}{2}] \cap T_1 = \varnothing$, since $|x| \leq \frac{1}{2}$ in (1) implies $y = \frac{b-a}{2}x + \frac{b+a}{2} \in [-\frac{b-a}{4} + \frac{b+a}{2}, \frac{b-a}{4} + \frac{b+a}{2}] = [a + \frac{b-a}{4}, b - \frac{b-a}{4}] \subseteq [a + \frac{\delta_1}{4}, b - \frac{\delta_1}{4}]$, so $y \notin S$.

- $T_2 \subseteq (-\frac{3}{2}, \frac{3}{2})$, since $y \in S$ implies $y \in (c - \frac{\delta_2}{5}, d + \frac{\delta_2}{5})$ in (2) , we have $\frac{2}{d-c}\left(c - \frac{\delta_2}{5} - \frac{d+c}{2}\right) < x < \frac{2}{d-c}\left(d + \frac{\delta_2}{5} - \frac{d+c}{2}\right)$. Using $\delta_2 < \frac{5}{4}(d - c)$, we have $x \in (-\frac{3}{2}, \frac{3}{2})$.

Define $T(x) \equiv T_1(x) \wedge T_2(x)$. It follows from Lemma 11 that $\mathrm{CSP}(LE_{\mathbb{R}}[\mathbb{Q}] \cup \{T\})$ is NP-hard. Since T is pp-definable in $LE_{\mathbb{R}}[\mathbb{F}] \cup \{S\}$, it follows that $\mathrm{CSP}(LE_{\mathbb{R}}[\mathbb{F}] \cup \{S\})$ is NP-hard as well. $\qquad\square$

We are now ready to prove Theorem 2 which we restate here for the reader's convenience.

Theorem 2. Let $\mathbb{F}$ be any subfield of $\mathbb{R}$. Let $LE_{\mathbb{R}}[\mathbb{F}] \subseteq \Gamma$ be a constraint language over $\mathbb{R}$ such that every primitive positively definable unary relation over Γ can be written as a finite union of intervals with endpoints in $\mathbb{F}$. Assume that the following hold:

1. Γ contains a relation that is not essentially convex and this is witnessed by points with coordinates in $\mathbb{F}$, and
2. Γ can primitive positively define a BNU relation.

Then, $\mathrm{CSP}(\Gamma)$ is NP-hard.

Proof. Let $R \in \Gamma$ be a k-ary relation that is not essentially convex and assume that this is witnessed by $p, q \in \mathbb{F}^k$. By Lemma 6, p and q also witness the exclusion of an interval in R. Define $R' = \mathcal{L}_{R,p,q} \in \langle LE_{\mathbb{R}}[\mathbb{F}] \cup \{R\} \rangle \subseteq \langle \Gamma \rangle$.

The unary relation R' is a member of $\mathcal{O}_{\mathbb{R},\mathbb{F}}$ and excludes an interval between 0 and 1.

Let $T \in \langle \Gamma \rangle$ be a BNU. Since T is a member of $\mathcal{O}_{\mathbb{R},\mathbb{F}}$, it follows that it contains at least two distinct points $u, v \in \mathbb{F}$. Define $T' = \mathcal{L}_{T,u,v} \in \langle LE_{\mathbb{R}}[\mathbb{F}] \cup \{T\} \rangle \subseteq \langle \Gamma \rangle$. Note that T' is a BNU containing 0 and 1.

Finally, define $U(x) \equiv R'(x) \wedge T'(x)$. We conclude that U is a bounded unary relation in $\mathcal{O}_{\mathbb{R},\mathbb{F}}$ that contains 0 and 1 and excludes an interval. It follows from Lemma 12 that $CSP(LE_{\mathbb{R}}[\mathbb{F}] \cup \{U\})$ is NP-hard. Since U is pp-definable in Γ, and $LE_{\mathbb{R}}[\mathbb{F}] \subseteq \Gamma$, it follows that $CSP(\Gamma)$ is NP-hard as well. $\qquad\square$

Let us consider a few immediate applications of Theorem 2.

Example 13. Let Γ be an o-minimal constraint language over $\mathbb{R}$ such that $LE_{\mathbb{R}}[\mathbb{R}] \subseteq \Gamma$. If Γ contains a relation that is not essentially convex and $\langle \Gamma \rangle$ contains a BNU relation, then $CSP(\Gamma)$ is NP-hard by Theorem 2 with $\mathbb{F} = \mathbb{R}$.

It may appear problematic that we require the constraint language Γ in Example 13 to contain an uncountable set $LE_{\mathbb{R}}[\mathbb{R}]$. Formally, this is no problem since we will always identify a finite subset of relations that yield NP-hardness. Naturally, there are many cases when we do not need to consider uncountable sets of relations. One such case is provided by an earlier hardness result for semialgebraic relations.

Example 14. Consider the countable constraint language $\Gamma = LE_{\mathbb{R}}[\mathbb{Q}] \cup \{\leq\}$. Note that Γ pp-defines the BNU $U = \{x \in \mathbb{R} \mid -1 \leq x \wedge x \leq 1\}$ and that U contains an infinite number of rational points. Hence, Theorem 2 with $\mathbb{F} = \mathbb{Q}$ is applicable. The set $LE_{\mathbb{R}}[\mathbb{Q}]$ can be replaced with the finite set $\{R_+, \{1\}\}$ (cf. Lemma 4 and 5), so we recover the following result.

Lemma 15 (Bodirsky, Jonsson, and von Oertzen [5, Lemma 3.5]). *Let $R \in \mathbb{R}^k$ be a semialgebraic relation that is not essentially convex, and this is witnessed by two rational points. Then, $CSP(\Gamma_{lin} \cup \{R\})$ is NP-hard where $\Gamma_{lin} = \{R_+, \leq, \{1\}\}$.*

Example 16. Theorem 2 also implies the hardness-part of Theorem 1. In this case, the exclusion of intervals is always witnessed by rational points [5, Lemma 3.9]. Every non-empty unary relation that is semilinear contains at least one rational point [5, Lemma 3.7] so, in particular, every semilinear BNU contains at least one rational point. If a semilinear BNU U contains exactly one rational point p, then it contains no intervals so $U = \{p, q_1, \ldots, q_k\}$ where $q_1, \ldots, q_k \notin \mathbb{Q}$. This implies that one can pp-define a non-rational constant in $\Gamma_{lin} \cup \{U\}$ and this, as we have already noted, is impossible. Thus, Theorem 2 is applicable by setting $\mathbb{F} = \mathbb{Q}$.

11

3.2. Semialgebraic constraint languages

In this section, we study constraint languages $\Gamma \subseteq SA_\mathbb{R}[\mathbb{Q}]$. For this application, we need to understand which unary relations that are pp-definable in $LE_\mathbb{R}[\mathbb{A}] \cup \Gamma$. First, we show (in Lemma 17) that every non-empty relation in $SA_\mathbb{R}[\mathbb{Q}]$ contains an algebraic point. Second, we use this result to show (in Lemma 18), that a non-essentially convex semialgebraic relation $R \in SA_\mathbb{R}[\mathbb{A}]$ has algebraic witnesses. Finally, by combining this with Lemma 10, we conclude that Theorem 2 is applicable.

It follows that if Γ is a semialgebraic constraint language and $\mathrm{CSP}(LE_\mathbb{R}[\mathbb{A}] \cup \Gamma)$ is in P, then either every relation in Γ is essentially convex or $\langle \Gamma \rangle$ does not contain any BNU. The important question of whether the reverse implication holds (as is the case in Theorem 1) is discussed in Section 5.

Lemma 17. Every non-empty relation $R \in SA_\mathbb{R}[\mathbb{A}]$ contains an algebraic point.

Proof. The reals and the algebraic numbers are both real closed fields (see, for instance, Marker [27]). The theory of real closed fields is complete (this is one of the consequences of Theorem 7). This implies that any two models (e.g., the reals and the algebraic numbers) are *elementarily equivalent*: every L-sentence is true in one if and only if it is true in the other.

Let R be defined by an L-formula $\varphi(x)$. Since R is non-empty, the sentence $\exists x.\varphi(x)$ is true over the reals. Furthermore, $\exists x.\varphi(x)$ is true over the reals if and only if it is true over the algebraic numbers, so R contains an algebraic point. $\square$

Lemma 18. Let $R \in SA_\mathbb{R}[\mathbb{A}]$ be a relation that is not essentially convex. Then, this is witnessed by algebraic points $p, q \in R \cap \mathbb{A}$.

Proof. From Lemma 6 with $\mathbb{F} = \mathbb{R}$ and Theorem 8, it follows that R excludes an interval. Consider the following relation:

$$\begin{aligned}
Z(x_1, \ldots, x_m, y_1, \ldots, y_m) \;\equiv\; & R(x) \wedge R(y) \wedge \\
& \exists \delta_1 \delta_2. 0 < \delta_1 < \delta_2 < 1 \wedge \\
& \forall t.(\delta_1 \le t \le \delta_2) \to \neg R(t \cdot x + (1 - t) \cdot y),
\end{aligned}$$

where $x = (x_1, \ldots, x_m)$ and $y = (x_1, \ldots, y_m)$. Clearly, $Z \in SA_\mathbb{R}[\mathbb{A}]$.

We see that $Z(x, y)$ holds if and only if x, y witness the exclusion of an interval in R. By Lemma 17, $Z = \varnothing$ or $Z \cap \mathbb{A}^{2m} \neq \varnothing$. The case $Z = \varnothing$ is impossible since R excludes an interval. Hence, there exist witnesses $p, q \in R \cap \mathbb{A}^m$. These also witness that S is not essentially convex. $\square$

Theorem 19. Let $LE_\mathbb{R}[\mathbb{A}] \subseteq \Gamma \subseteq SA_\mathbb{R}[\mathbb{A}]$. Assume that the following hold:

1. Γ contains a relation that is not essentially convex, and

2. $\langle\Gamma\rangle$ contains a BNU relation.

Then, $\mathrm{CSP}(\Gamma)$ is NP-hard.

Proof. By Lemma 10, every unary relation in $SA_{\mathbb{R}}[\mathbb{A}]$ is a member of $\mathcal{O}_{\mathbb{R},\mathbb{A}}$. Let $R \in \Gamma$ be a relation that is not essentially convex. By Lemma 18, this is witnessed by algebraic points in R. It follows that Theorem 2 is applicable with $\mathbb{F} = \mathbb{A}$. $\qquad\square$

4. Algebraic varieties

We will now turn our attention to constraint languages based on algebraic varieties. A set $V \subseteq \mathbb{R}^n$ is called an *(algebraic) variety* if it can be defined by $V = \{x \in \mathbb{R}^n \mid p_1(x) = 0 \land \cdots \land p_m(x) = 0\}$ where $p_1, \ldots, p_m$ are polynomials with algebraic coefficients. Let $\mathcal{V}$ denote the set of all algebraic varieties. Clearly, $\mathcal{V} \subseteq SA_{\mathbb{R}}[\mathbb{A}]$. We will prove (in Theorem 25) that for every finite set $\Gamma \subseteq \mathcal{V}$, the problem $\mathrm{CSP}(LE_{\mathbb{R}}^0[\mathbb{A}] \cup \Gamma)$ is either locally tractable or locally NP-hard. The proof is based on Theorem 19 but it also makes use of a number of auxiliary results that we will present below. Observe that Theorem 25 is based on exploiting the set of *homogeneous* linear equations and not the full set of equations as in the proof of Theorem 19.

The following result (which is well-known and easy to prove) will be used numerous times in the sequel.

Lemma 20. Let $V \subseteq \mathbb{R}^n$ be a variety and let $L \subseteq \mathbb{R}^n$ be a line. If infinitely many points of L are in V, then $L \subseteq V$.

4.1. Tractability

We begin by proving a tractability result that is strictly stronger than what we need for proving Theorem 25. Let $LI_{\mathbb{R}}[\mathbb{A}]$ denote the set of linear inequalities with coefficients in $\mathbb{A}$. We prove that if Γ contains only convex varieties, then $\mathrm{CSP}(LI_{\mathbb{R}}[\mathbb{A}] \cup \Gamma)$ is locally tractable. This immediately implies that $\mathrm{CSP}(LE_{\mathbb{R}}^0[\mathbb{A}] \cup \Gamma)$ is locally tractable, too.

For a subset $R \subseteq \mathbb{R}^n$, let $\mathrm{aff}(R)$ denote the *affine hull of R* in $\mathbb{R}^n$: $\mathrm{aff}(R) = \{\sum_{i=1}^k c_i x_i \mid k \geq 1, \ c_i \in \mathbb{R}, \ x_i \in R, \ \sum_{i=1}^k c_i = 1\}$. The set R is said to be *affine* if $R = \mathrm{aff}(R)$.

The points $x_1, \ldots, x_k \in \mathbb{R}^n$ are said to be *affinely independent* if $\sum_{i=1}^k a_i x_i = 0$ with $\sum_{i=1}^k a_i = 0$ implies $a_1 = \cdots = a_k = 0$. The dimension, $\dim(R)$, of $R \subseteq \mathbb{R}^n$ is defined to be one less than the maximum number of affinely independent points in R.

If R is affine and $\dim(R) = d$, then R can be written as the affine hull of any $d+1$ affinely independent points, and alternatively as the set of solutions to a (finite) linear system of equations. Furthermore, if these points are members of $\mathbb{A}^n$, then the linear system of equations can be chosen to have coefficients in $\mathbb{A}$ only.

We can now characterise the convex algebraic varieties.

Lemma 21. Let $V \subseteq \mathbb{R}^n$ be a non-empty algebraic variety. If V is convex, then V is the solution to a finite linear system of equations with coefficients in $\mathbb{A}$.

Proof. We first show that V is affine. To this end, let $x_1, \ldots, x_k \in V$ and let $c_1, \ldots, c_k \in \mathbb{R}$ be such that $\sum_{i=1}^k c_i = 1$. We want to show that $y = \sum_{i=1}^k c_i x_i$ is also in V. Since V is convex, the point $x = \frac{1}{k} \sum_{i=1}^k x_i$ lies in V. But then, for some $\varepsilon > 0$, the line segment $\{(1-p) \cdot x + p \cdot y \mid 0 \leq p \leq \varepsilon\}$ lies in V. Therefore, by Lemma 20, the entire line $\{(1-p) \cdot x + p \cdot y \mid p \in \mathbb{R}\}$ lies in V. It follows that $y \in V$, so V is affine.

Next, we show that V can be defined by a linear system of equations with coefficients in $\mathbb{A}$. Since $V \neq \varnothing$, we have $\dim(V) = d \geq 0$. Define a relation Z as follows:

$$Z(x_0, \ldots, x_d) \equiv V(x_0) \wedge \cdots \wedge V(x_d) \wedge$$

$$\forall a_0, \ldots, a_d \Big(\sum_{i=0}^d a_i = 0 \wedge \neg(a_0 = \cdots = a_d = 0) \Big) \to \sum_{i=0}^d a_i x_i = 0$$

The relation Z contains all $(d+1)$-tuples of affinely independent points in V. Since Z is non-empty, it follows from Lemma 17 that it contains an algebraic point. Thus, there exist $d+1$ affinely independent algebraic points $x_0, \ldots, x_d \in V$. Since V is affine, it follows that $V = \text{aff}(\{x_0, \ldots, x_d\})$ can be defined by a linear system of equations with coefficients in $\mathbb{A}$. $\square$

Thus, a convex algebraic variety can always be viewed as a conjunction of linear equations with coefficients in $\mathbb{A}$. This enables us to use a result by Adler and Beling [1]. They prove that the complexity of linear programming with algebraic coefficients depends on certain parameters that can be deduced from the coefficients. However, if only a finite number of algebraic coefficients are used, then these parameters are bounded and linear programming can be performed in polynomial time. We get the following result using our terminology.

Theorem 22 (Adler and Beling [1]). $\text{CSP}(LI_{\mathbb{R}}[\mathbb{A}])$ is locally tractable and globally tractable if the coefficients are represented by their minimal polynomials.

Corollary 23. Let Γ be a set of convex algebraic varieties. Then, $\mathrm{CSP}(LI_{\mathbb{R}}[\mathbb{A}] \cup \Gamma)$ is locally tractable.

4.2. Complexity classification

We begin by presenting an NP-hardness result. For arbitrary $c \in \mathbb{A}$ and unary relation $U \subseteq \mathbb{R}$, let $c \cdot U = \{c \cdot x \mid x \in U\}$. The relation $T_c = \{(x, y) \in \mathbb{R}^2 \mid c \cdot x = y\}$ is a member of $LE_{\mathbb{R}}^0[\mathbb{A}]$ for every $c \in \mathbb{A}$. Hence, we can pp-define $c \cdot U$ in $LE_{\mathbb{R}}^0[\mathbb{A}] \cup \{U\}$ via

$$(c \cdot U)(y) \equiv \exists x.U(x) \wedge T_c(x, y).$$

Lemma 24. If $U \subseteq \mathbb{A} \setminus \{0\}$ is finite and $|U| > 1$, then $\mathrm{CSP}(LE_{\mathbb{R}}^0[\mathbb{A}] \cup \{U\})$ is NP-hard.

Proof. First, suppose that $U = -U$. This implies that

$$U = \{-a_k, \ldots, -a_1, a_1, \ldots, a_k\}$$

where $0 < a_1 < \cdots < a_k$. Thus, $\{-1, 1\} = (a_1^{-1} \cdot U) \cap (a_k^{-1} \cdot U) \in \langle LE_{\mathbb{R}}^0[\mathbb{A}] \cup \{U\}\rangle$, so $\mathrm{CSP}(LE_{\mathbb{R}}^0[\mathbb{A}] \cup \{U\})$ is NP-hard by Lemma 11.

Now, suppose instead that $U \neq -U$. If $U \cap (-U) \neq \varnothing$, then we can apply Case 1 on the relation $U \cap -U$. Otherwise, let $a = \min U$ and $b = \max U$. Then, $\{1\} = (a^{-1} \cdot U) \cap (b^{-1} \cdot U)$ is a member of $\langle LE_{\mathbb{R}}^0[\mathbb{A}] \cup \{U\}\rangle$. Now, every equation $b_1 x_1 + \cdots + b_n x_n = b$ (with $b, b_1, \ldots, b_n \in \mathbb{A}$) can be pp-defined in $LE_{\mathbb{R}}^0[\mathbb{A}] \cup \{\{1\}\}$:

$$b_1 x_1 + \cdots + b_n x_n = b \Leftrightarrow \exists y, z.b_1 x_1 + \cdots + b_n x_n = y \wedge \{1\}z \wedge bz = y.$$

It follows that $LE_{\mathbb{R}}[\mathbb{A}] \subseteq \langle LE_{\mathbb{R}}^0[\mathbb{A}] \cup \{U\}\rangle$.

The relation U is a BNU that excludes an interval, so $\mathrm{CSP}(LE_{\mathbb{R}}[\mathbb{A}] \cup \{U\})$ is (locally) NP-hard by Lemma 12. It follows from Lemma 4 that $\mathrm{CSP}(LE_{\mathbb{R}}^0[\mathbb{A}] \cup \{U\})$ is (locally) NP-hard as well. □

We are now ready to prove the main result of this section. We say that a relation $R \subseteq \mathbb{R}^k$ is *0-valid* if $(0, \ldots, 0) \in R$. Similarly, we say that a constraint language Γ is 0-valid if every relation in it is 0-valid. If Γ is a 0-valid constraint language, then $\mathrm{CSP}(\Gamma)$ is *trivially tractable*: we do not even need to look at the instance to say that it is satisfied by the assignment that gives every variable the value 0. We have not encountered 0-valid constraint languages earlier in this article since (for instance) the constant relation $\{1\}$ has always been a member of the languages under consideration.

Theorem 25. Let $LE_{\mathbb{R}}^0[\mathbb{A}] \subseteq \Gamma \subseteq \mathcal{V}$. Then, one of the following holds:

1. Γ is 0-valid and CSP(Γ) is trivially tractable,
2. Γ contains convex relations only and CSP(Γ) is locally tractable, or
3. CSP(Γ) is NP-hard.

Proof. If Γ is 0-valid, then CSP(Γ) is trivially tractable. Thus, we assume that Γ contains a (non-empty) relation S that is not 0-valid.

If the relations in Γ are convex, then CSP(Γ) is locally tractable by Corollary 23. We can henceforth assume that Γ contains at least one non-convex relation R.

Assume R is essentially convex and has arity n. Since R is non-convex, there exist two distinct points $p, q \in R$ such that the line segment M between them intersected with R excludes some finite number of points, i.e. $M \setminus R$ is a finite set. Now note that the line $L \subseteq \mathbb{R}^n$ through p and q intersects R in an infinite number of points. Consequently, $L \subseteq R$ by Lemma 20. This contradicts that the line segment between p and q excludes points. We conclude that R is not essentially convex.

Arbitrarily choose an algebraic point $a \in S$; such a point exists due to Lemma 17. Let k be the arity of S and consider the line $L \subseteq \mathbb{R}^k$ through $(0, \ldots, 0)$ and a. If $L \cap S$ is infinite, then $L \subseteq S$ by Lemma 20 and this contradicts the fact that S is not 0-valid. Hence, $L \cap S$ is a finite set and $(0, \ldots, 0) \notin L \cap S$. Let $T = \mathcal{L}_{S,(0,\ldots,0),a}$ and note that T is a unary relation containing a finite number of points, $0 \notin T$, and $T \subseteq \mathbb{A}$ by Lemma 10.

Assume first that $|T| > 1$. By Lemma 24, CSP($LE_{\mathbb{R}}^0[\mathbb{A}] \cup \{T\}$) is NP-hard.

Assume instead that $|T| = 1$ so that $T = \{a\}$ for some $a \neq 0$, $a \in \mathbb{A}$ by Lemma 17. It follows that $LE_{\mathbb{R}}[\mathbb{A}] \subseteq \langle \Gamma \rangle$. The relation R is not essentially convex, and by Lemma 18, this is witnessed by algebraic points $p, q \in R \cap \mathbb{A}^n$. Let $L \subseteq \mathbb{R}^n$ be the line through p and q. If $L \cap R$ is an infinite set, then $L \subseteq R$ by Lemma 20. This contradicts that p and q are witnesses of R being not essentially convex. Hence, $L \cap R$ is a finite set and $|L \cap R| > 1$ since p, q are distinct. It follows that $\langle \Gamma \rangle$ contains the BNU $\mathcal{L}_{R,p,q}$ together with the relation R that is not essentially convex. NP-hardness of CSP(Γ) follows from Theorem 19. $\square$

Constraint languages that are 0-valid are, in a sense, trivial from a CSP point of view. In certain cases, it makes sense to verify if there is a solution that does not assign zero to all variables. This problem and several variants of it have been studied in the literature, cf. the survey [8]. Finding non-zero solutions to CSPs may increase the computational complexity considerably, and this holds also when we restrict ourselves to algebraic varieties. Define $U = \{x \in \mathbb{R} \mid x^3 - x = 0\}$ and consider the constraint language $\Gamma = LE_{\mathbb{R}}^0[\mathbb{Q}] \cup \{U\}$. We note that U is an algebraic variety and that $U = \{-1, 0, 1\}$. Furthermore,

the language Γ is 0-valid so CSP(Γ) is in P. The following problem (which is sometimes referred to as WEAK PARTITION) is known to be NP-hard [32].

Input. A finite vector $(a_1, \ldots, a_n)$ of integers.
Question. Is the equation $\sum_{i=1}^{n} a_i x_i = 0$ solvable with all $x_i \in \{-1, 0, -1\}$ and with at least one $x_j \neq 0$?

By using the relation U and Lemma 5, it is easy to see that WEAK PARTITION is polynomial-time reducible to the problem of finding a non-zero solution to CSP(Γ).

5. Discussion

We finish this article with a discussion of a number of interesting and non-trivial questions that have materialised along the way.

Q1. What is the computational complexity of CSP(Γ) when Γ is essentially convex or when $\langle \Gamma \rangle$ does not contain any BNU?
Q2. What is the computational complexity of CSP(Γ) when $LE_\mathbb{R}[\mathbb{Q}] \subseteq \Gamma \subseteq SA_\mathbb{R}[\mathbb{Q}]$?
Q3. What is the computational complexity of CSP(Γ) when $LE_\mathbb{R}^0[\mathbb{A}] \subseteq \Gamma \subseteq SA_\mathbb{R}[\mathbb{A}]$?

5.1. Q1: Tractable numeric CSPs

In this section, we will discuss a number of apparent barriers to determining the complexity of the fragments remaining after applying Theorem 2 and Theorem 19.

We note that there are some recent results by Bodirsky and Mamino [9] concerning the complexity of CSP(Γ) when Γ is semilinear, convex and additionally invariant under the max operation [9]: here, a relation R is invariant under a function f if it is preserved by applying f componentwise on arbitrary tuples in R. Such CSPs have important connections with, for example, mean payoff games. It is easy to verify that R_+ is not invariant under the max operation so the approach by Bodirsky and Mamino cannot immediately be used for refining results like Theorem 2 and Theorem 19.

5.1.1. Essentially convex languages

In order to see the difficulties in classifying essentially convex languages, it suffices to consider the convex case. Take the innocuous-looking semialgebraic constraint language $\Gamma = \{R_+, \leq, \{1\}, \{(x, y) \in \mathbb{R}^2 \mid y \geq x^2\}\}$. It is easy to verify that this constraint language contains convex relations only. We will first compare CSP(Γ) with the following well-known problem.

17

SUMS OF SQUARE ROOTS
Input: $a_0, \ldots, a_n \in \mathbb{N}$ (represented in binary).
Question: Is $\sum_{i=1}^n \sqrt{a_i} \geq a_0$.

The complexity of this problem is unknown and it has been discussed (at least) since 1976 when Garey, Graham, and Johnson [16] observed that its complexity may be difficult to determine; in fact, it is not even known if the problem is in NP. It is, however, known to be in PSPACE and in the counting hierarchy [2]. The second inclusion point in the direction that the problem is not PSPACE-complete if we believe that the counting hierarchy does not collapse to a finite level. It can be verified that

$$\sum_{i=1}^n \sqrt{a_i} \geq a_0 \Leftrightarrow \exists b_1, \ldots, b_n. \left(\bigwedge_{i=1}^n a_i \geq b_i^2 \wedge b_1 + \ldots + b_n \geq a_0 \right).$$

This implies that SUMS OF SQUARE ROOTS is polynomial-time reducible to CSP(Γ) and that CSP(Γ) is at least as hard SUMS OF SQUARE ROOTS[2].

We continue our analysis of CSP(Γ) by relating it to the following problem.

FEASIBILITY OF CONVEX POLYNOMIAL INEQUALITIES
Input: A set of variables V and a set of polynomial inequalities each of which defining a convex set; the coefficients of the polynomials are rational numbers.
Question: Is there a point in $\mathbb{R}^{|V|}$ that satisfies all inequalities?

This problem has unknown complexity, too. It is a member of PSPACE [13] but it is probably not PSPACE-complete since the unrestricted FEASIBILITY OF POLYNOMIAL INEQUALITIES is not believed to be PSPACE-complete. In fact, the complexity class $\exists \mathbb{R}$ has been introduced with the purpose of capturing the complexity of this problem [29, 30]. Obviously, CSP(Γ) is polynomial-time reducible to FEASIBILITY OF CONVEX POLYNOMIAL INEQUALITIES which implies that CSP(Γ) is "sandwiched" between two problems of unknown complexity. It is thus conceivable that determining the complexity of CSP(Γ) is a non-trivial problem.

The difficulty of determining the complexity of CSP(Γ) is not an isolated phenomenon: the computational complexity of problems based on convex objects is to a large extent unknown. Even when considering extremely well-studied convex sets, there are large gaps in our understanding. The prime example is the feasibility problem for semidefinite programming: we only know that this problem is either in NP $\cap$ coNP or that it is not in NP $\cup$ coNP due to duality results by Ramana [28]. Semidefinite programming is of particular interest for understanding both convex and semialgebraic relations

---

[2]This example was suggested by Manuel Bodirsky and Marcello Mamino

since every semidefinite representable relation is simultaneously convex and semialgebraic. In fact, Helton and Nie [18] have showed that the converse statement is true surprisingly often and they conjecture that it is true in general.

One objection against the discussion above is that Γ may contain too few linear equations: what happens if we instead study $\Gamma' = LE_{\mathbb{R}}[\mathbb{A}] \cup \Gamma$? This is an interesting question but it does not seem to be much easier to answer. The relations in $LE_{\mathbb{R}}[\mathbb{A}]$ are all convex so the discussion concerning convexity is still valid in this case. In the comparison with Sums of Square Roots, we argue that Sums of Square Roots is polynomial-time reducible to $\mathrm{CSP}(\Gamma)$ and since the complexity of Sums of Square Roots is unknown, we do not get any concrete information concerning $\mathrm{CSP}(\Gamma)$. The very same holds for $\mathrm{CSP}(\Gamma')$ since $\mathrm{CSP}(\Gamma)$ is trivially polynomial-time reducible to $\mathrm{CSP}(\Gamma')$. Nevertheless, it may be a good idea to analyse $\mathrm{CSP}(\Gamma')$ instead of $\mathrm{CSP}(\Gamma)$ as long as one is interested in obtaining hardness results.

5.1.2. Languages where BNUs are not pp-definable

Let Γ be a constraint language and assume that $\langle \Gamma \rangle$ does not contain any BNU. What can then be said about the complexity of $\mathrm{CSP}(\Gamma)$? In [22], this problem is studied in the special case when $\{R_+, \{1\}\} \subseteq \Gamma$ and Γ is semilinear with coefficients in $\mathbb{Q}$. The inclusion of R_+ and $\{1\}$ (essentially $LE_{\mathbb{R}}[\mathbb{Q}]$) ensures that the intersection of any relation in $\langle \Gamma \rangle$ with a line is either empty, contains a single point, or is unbounded in both directions. The algorithm is based on the idea of shrinking an upper bound A on the affine hull of the solution space to the instance. The upper bound is stored as an affine subspace. It is repeatedly intersected by a single constraint, and adjusted to the affine hull of the intersection. When the procedure can no longer reduce the dimension of A, it terminates and answers "unsatisfiable" if A is empty, and "satisfiable" otherwise. This algorithm is called the affine consistency algorithm. If A is empty, then the answer given by the algorithm is clearly correct. However, if A is non-empty, then we cannot in general know whether the instance has a solution. It is shown in [22] that the affine consistency algorithm returns the correct answer for all instances of $\mathrm{CSP}(\Gamma)$ if and only if $\langle \Gamma \rangle$ does not contain any BNU. In fact, this result extends to our setting:

> if $LE_{\mathbb{R}}[\mathbb{A}] \subseteq \Gamma$ and every unary relation in $\langle \Gamma \rangle$ is a member of $\mathcal{O}_{\mathbb{R},\mathbb{A}}$, then the affine consistency algorithm returns the correct answer for all instances of $\mathrm{CSP}(\Gamma)$ if and only if $\langle \Gamma \rangle$ does not contain any BNU.

In the semilinear setting, the affine consistency algorithm can be implemented by computing the affine hulls of a number of polyhedra in each step.

19

This, in turn, can be done in polynomial time using the ellipsoid method. In the general setting, it is not at all clear how to carry out such an operation in polynomial time. For example, for $\Gamma \subseteq SA_\mathbb{R}[\mathbb{A}]$, we need to compute the affine hull of the intersection of an affine subspace with a semialgebraic set. It should be noted that the semialgebraic sets appearing in these calculations are not arbitrary, since we are assuming that $\langle \Gamma \rangle$ does not contain any BNU. It is possible that one could leverage this additional information.

5.2. Q2: Semialgebraic languages with rational equations

Assume we want to strengthen Theorem 19 and only consider languages containing $LE_\mathbb{R}[\mathbb{Q}]$ instead of $LE_\mathbb{R}[\mathbb{A}]$. It is possible to show the following.

Proposition 26. Let $LE_\mathbb{R}[\mathbb{Q}] \subseteq \Gamma \subseteq SA_\mathbb{R}[\mathbb{Q}]$ and assume that the following hold:

1. Γ contains a relation that is not essentially convex and this is witnessed by rational points, and
2. $\langle \Gamma \rangle$ contains a BNU relation with at least two rational points.

Then, CSP(Γ) is NP-hard.

If we drop the requirements of rational witnesses in both (1) and (2), then things are getting more complicated. One may observe that if some BNU in $\langle \Gamma \rangle$ contains an interval, then it contains an infinite number of rational points. It thus seems reasonable to begin by studying BNUs that are unions of a finite number of points (and consequently are not essentially convex.) We illustrate this in the following three examples.

Example 27. Let $U_1 = \{-\sqrt{2}, \sqrt{2}\}$ and $\Gamma_1 = LE_\mathbb{R}[\mathbb{Q}] \cup \{U_1\}$. The relation U_1 is not essentially convex and this is only witnessed by irrational points. Furthermore, U_1 is a BNU that contains no rational points so CSP(Γ_1) could be a candidate for a polynomial-time solvable problem—it clearly avoids NP-hardness implied by Proposition 26. However, CSP(Γ_1) is NP-hard: Consider the relation

$$N = \{(x, y, z) \in \mathbb{R}^3 \mid \exists w. x + y + z = w \wedge T(x) \wedge T(y) \wedge T(z) \wedge T(w)\}.$$

This relation is pp-definable in Γ_1 and

$$N = \{-\sqrt{2}, \sqrt{2}\}^3 \setminus \{(-\sqrt{2}, -\sqrt{2}, -\sqrt{2}), (\sqrt{2}, \sqrt{2}, \sqrt{2})\}$$

so NP-hardness of CSP(Γ_1) follows by an immediate polynomial-time reduction from CSP($\{0, 1\}^3 \setminus \{(0, 0, 0), (1, 1, 1)\}$), which is an NP-hard problem [31].

Example 28. Let $U_2 = \{1, \sqrt{2}\}$. By Theorem 19, we know that $\text{CSP}(LE_{\mathbb{R}}[\mathbb{A}] \cup \{U_2\})$ is NP-hard. We will now investigate the complexity of $\text{CSP}(LE_{\mathbb{R}}[\mathbb{Q}] \cup \{U_2\})$ and prove that this problem is polynomial-time solvable.

We first recall a basic fact concerning linear equation systems. Let $Ax = b$ be a solvable linear system where $x = (x_1, \ldots, x_m)$ and all coefficients are rational. Define $S[Ax = b, x_i] = \{x_i \mid A(x_1, \ldots, x_m) = b\}$, $1 \leq i \leq m$, and note that $S[Ax = b, x_i]$ either equals $\mathbb{R}$ or $S[Ax = b, x_i] = \{a\}$ for some $a \in \mathbb{Q}$.

It is sufficient to study the complexity of $\text{CSP}(\{R_+, \{1\}, U_2\})$ since $LE_{\mathbb{R}}[\mathbb{Q}] \subseteq \langle \{R_+, \{1\}\} \rangle$ so we let (V, C) denote an arbitrary instance of $\text{CSP}(\{R_+, \{1\}, U_2\})$. Let (V, C') be the same instance as (V, C) but with every occurrence of U_2 replaced with $\{1\}$. We will show that (V, C) is satisfiable if and only if (V, C') is satisfiable, and this implies that $\text{CSP}(\{R_+, \{1\}, U_2\})$ is in P since $\text{CSP}(\{R_+, \{1\}\})$ is in P.

If (V, C) is not satisfiable, then (V, C') is not satisfiable either. Hence, we assume that (V, C) is satisfiable. If there is a solution s to (V, C) satisfying $\sqrt{2} \notin \{s(v) \mid v \in V\}$, then (V, C') is satisfiable, too. Assume to the contrary that no such solution exist and arbitrarily choose a solution $s : V \to \mathbb{R}$. Define the non-empty set of variables $V_{\sqrt{2}} = \{v \in V \mid s(v) = \sqrt{2}\}$. Now, construct a new instance (V_e, C_e) by equating $V_{\sqrt{2}}$ in (V, C) into a fresh variable y, i.e. we replace each occurrence of variables $v \in V_{\sqrt{2}}$ with y. Note that y *cannot* be assigned the value 1 by any satisfying assignment. If y is constrained by U_2, then remove this constraint. Replace each occurrence of U_2 with $\{1\}$ and let (V'_e, C'_e) denote the resulting instance. Note that (V'_e, C'_e) is still satisfiable and it can be viewed as a linear equation system $Ax = b$ (with $x = (x_1, \ldots, x_m, y)$) with rational coefficients since it does not contain any occurrence of the relation U_2. We conclude the following.

- If $S[Ax = b, y] = \{a\}$, then $a \in \mathbb{Q}$ by the observation above. However, the solution s shows that $\sqrt{2} \in S[Ax = b, y]$ which leads to a contradiction.

- If $S[Ax = b, y] = \mathbb{R}$, then we reach a contradiction since $1 \notin S[Ax = b, y]$.

Thus, there exists a solution to (V, C) that assigns the value 1 to every variable in $V_{\sqrt{2}}$ and, consequently, (V, C') has a solution.

Example 29. Let $U_3 = U_2 \cup \{1 + 2\sqrt{2}\} = \{1, \sqrt{2}, 1 + 2\sqrt{2}\}$ and $\Gamma_3 = LE_{\mathbb{R}}[\mathbb{Q}] \cup \{U_3\}$. We claim that $\text{CSP}(\Gamma_3)$ is NP-hard. One can verify that the relation

$$\{(x, y, z) \in \mathbb{R}^3 \mid \exists w. x + y + z = w \wedge U_3(x) \wedge U_3(y) \wedge U_3(z) \wedge U_3(w)\}$$

is pp-definable in Γ_3 and that it only contains the three tuples $(1, \sqrt{2}, \sqrt{2})$ $(\sqrt{2}, 1, \sqrt{2})$, and $(\sqrt{2}, \sqrt{2}, 1)$. This implies that CSP(Γ_3) is NP-hard by a straightforward polynomial-time reduction from CSP($\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$), which is an NP-hard problem [31].

We conclude from these examples that the borderline between tractable and NP-hard problems is not clear even if we restrict our attention to expansions of $LE_{\mathbb{R}}[\mathbb{Q}]$ with BNUs containing a finite number of points.

5.3. Q3: Semialgebraic languages with homogeneous equations

Assume we want to strengthen Theorem 19 and only consider languages containing $LE_{\mathbb{R}}^0[\mathbb{A}]$ instead of $LE_{\mathbb{R}}[\mathbb{A}]$. By inspecting the proof, we see that it breaks down when we want to apply the main hardness result Theorem 2: this result requires that we have access to $LE_{\mathbb{R}}[\mathbb{A}]$ and not merely $LE_{\mathbb{R}}^0[\mathbb{A}]$. The reason is two-fold:

1. we cannot assume that one of the witnesses of an excluded interval equals $(0, \ldots, 0)$, and
2. we cannot assume that there is a BNU that contains the point $(0, \ldots, 0)$.

We do not see a straightforward way of generalising Theorem 2. Unfortunately, the hardness part of Theorem 25 (where we only need the set $LE_{\mathbb{R}}^0[\mathbb{A}]$) does not seem to provide any guidance. Here, we crucially exploit Lemma 20, i.e., if a line L intersects an algebraic variety V in infinitely many points, then $L \subseteq V$.

A related problem is that it seems hard to generalise the basic NP-hardness result for BNUs (Lemma 12). The proof is based on the possibility to scale and translate unary relations and this is not possible when only having access to $LE_{\mathbb{R}}^0[\mathbb{A}]$. This is reflected in the fact that we need to take care of 0-valid relations when studying algebraic varieties. It is thus inevitable that an analogue of Theorem 2 for constraint languages that only contain $LE_{\mathbb{R}}^0[\mathbb{A}]$ has to deal with additional cases.

An interesting question is whether there are any provably polynomial-time solvable classes besides the 0-valid relations? Due to the triviality of 0-valid relations, one may hypothesise that there is some class of relations that include the 0-valid relations and has a polynomial-time solvable CSP problem. We will now demonstrate that this is not the case by showing a much stronger result. Let R denote the 0-valid relation $R = \{(x, y) \in \mathbb{R}^2 \mid x^2 = y\}$, let S be an arbitrary semialgebraic relation that is not 0-valid, and let $\Gamma = LE_{\mathbb{R}}^0[\mathbb{A}] \cup \{R, S\}$. We claim that CSP($\Gamma$) is NP-hard. We prove this by pp-defining $\{1\}$ in Γ. This implies that we can pp-define the relation $\{-1, 1\}(x)$ via $\exists y.\{1\}y \wedge R(x, y)$ and NP-hardness follows from Lemma 24.

Now, let k denote the arity of S and consider the following relations

$$S_1(x_1, x_2, x_3, \ldots, x_{k-1}, x_k) \equiv S(x_1, x_2, x_3, \ldots, x_{k-1}, x_k),$$
$$S_2(x_2, x_3, \ldots, x_{k-1}, x_k) \equiv S(0, x_2, x_3, \ldots, x_{k-1}, x_k),$$
$$S_3(x_3, \ldots, x_{k-1}, x_k) \equiv S(0, 0, x_3, \ldots, x_{k-1}, x_k),$$
$$\vdots \qquad\qquad \vdots$$
$$S_{k-1}(x_{k-1}, x_k) \equiv S(0, 0, 0, \ldots, 0, x_{k-1}, x_k),$$
$$S_k(x_k) \equiv S(0, 0, 0, \ldots, 0, 0, x_k).$$

The relations $S_1, \ldots, S_k$ are clearly pp-definable in Γ. Let $M = \{i \mid S_i$ is non-empty, $1 \leq i \leq k\}$ and note that M is non-empty since S_1 is non-empty. Let $m = \max M$. If $m = k$, then S_k is a unary relation such that $0 \notin S_k$ since S is not 0-valid. Arbitrarily choose $t \in S_k \cap \mathbb{A}$ — we know that $S_k \cap \mathbb{A} \neq \varnothing$ by Lemma 17. Consider the following relation.

$$U = \{x \in \mathbb{R} \mid \exists y.S_k(x) \wedge x^2 = y \wedge y = t \cdot x\}$$

and note that it is pp-definable in Γ. The equation $x^2 = t \cdot x$ has the two solutions 0 and t so $U = \{t\}$. It follows that

$$\{1\} = \{x \in \mathbb{N} \mid U(x) \wedge \exists y.U(y) \wedge x \cdot t = y\}$$

and $\{1\}$ is pp-definable in Γ.

Assume instead that $m < k$. Define the relation

$$U = \{x \in \mathbb{N} \mid \exists x_1, \ldots, x_{k-m}.R(\underbrace{0, \ldots, 0}_{m-1 \text{ zeroes}}, x, x_1, \ldots, x_{k-m}\}$$

and note that it is pp-definable in Γ, it is unary, and $0 \notin U$. The construction above can hence be used for pp-defining $\{1\}$.

References

[1] I. Adler and P. A. Beling. Polynomial algorithms for linear programming over the algebraic numbers. *Algorithmica*, 12(6):436–457, 1994.

[2] E. Allender, P. Bürgisser, J. Kjeldgaard-Pedersen, and P. B. Miltersen. On the complexity of numerical analysis. *SIAM J. Comput.*, 38(5):1987–2006, 2009.

[3] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in Real Algebraic Geometry*, volume 10 of *Algorithms and Computation in Mathematics*. Springer, 2nd edition, 2006.

[4] M. Bodirsky and M. Grohe. Non-dichotomies in constraint satisfaction complexity. In *Proceedings of the 35th International Colloquium on Automata, Languages and Programming (ICALP-2008)*, pages 184–196, 2008.

[5] M. Bodirsky, P. Jonsson, and T. von Oertzen. Essential convexity and complexity of semi-algebraic constraints. *Logical Methods in Computer Science*, 8(4), 2012.

[6] M. Bodirsky, P. Jonsson, and T. von Oertzen. Horn versus full first-order: Complexity dichotomies in algebraic constraint satisfaction. *J. Log. Comput.*, 22(3):643–660, 2012.

[7] M. Bodirsky and J. Kára. The complexity of temporal constraint satisfaction problems. *J. ACM*, 57(2), 2010.

[8] M. Bodirsky, J. Kára, and B. Martin. The complexity of surjective homomorphism problems - a survey. *Discrete Applied Mathematics*, 160(12):1680–1690, 2012.

[9] M. Bodirsky and M. Mamino. Max-closed semilinear constraint satisfaction. In *11th International Computer Science Symposium in Russia (CSR-2016)*, pages 88–101, 2016.

[10] M. Bodirsky and M. Mamino. Constraint satisfaction problems over numeric domains. In A. A. Krokhin and S. Živný, editors, *The Constraint Satisfaction Problem: Complexity and Approximability*, volume 7 of *Dagstuhl Follow-Ups*, pages 79–111. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.

[11] A. A. Bulatov. A dichotomy theorem for nonuniform CSPs. In *Proceedings of the 58th IEEE Symposium on Foundations of Computer Science (FOCS-2017)*, 2017. To appear.

[12] A. A. Bulatov, P. Jeavons, and A. A. Krokhin. Classifying the computational complexity of constraints using finite algebras. *SIAM J. Comput.*, 34(3):720–742, 2005.

[13] J. F. Canny. Some algebraic and geometric computations in PSPACE. In *Proc. 20th ACM Symposium on Theory of Computing (STOC-1988)*, pages 460–467, 1988.

[14] T. Feder and M. Y. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: a study through datalog and group theory. *SIAM J. Comput.*, 28(1):57–104, 1998.

[15] J. Ferrante and C. Rackoff. A decision procedure for the first order theory of real addition with order. *SIAM J. Comput.*, 4(1):69–76, 1975.

[16] M. R. Garey, R. L. Graham, and D. S. Johnson. Some NP-complete geometric problems. In *Proc. 8th Annual ACM Symposium on Theory of Computing (STOC-1976)*, pages 10–22, 1976.

[17] J. E. Goodman and J. O'Rourke, editors. *Handbook of Discrete and Computational Geometry*, volume 2. Chapman & Hall/CRC, 2nd edition, 2004.

[18] J. W. Helton and J. Nie. Sufficient and necessary conditions for semidefinite representability of convex hulls and sets. *SIAM Journal on Optimization*, 20(2):759–791, 2009.

[19] D. S. Hochbaum and J. Naor. Simple and fast algorithms for linear and integer programs with two variables per inequality. *SIAM J. Comput.*, 23(6):1179–1192, 1994.

[20] P. Jeavons. On the algebraic structure of combinatorial problems. *Theoretical Comput. Sci.*, 200(1–2):185–204, 1998.

[21] P. Jonsson and T. Lööw. Computational complexity of linear constraints over the integers. *Artif. Intell.*, 195:44–62, 2013.

[22] P. Jonsson and J. Thapper. Affine consistency and the complexity of semilinear constraints. In *Proceedings of 39th International Symposium on Mathematical Foundations of Computer Science (MFCS-2014)*, pages 420–431, 2014.

[23] P. Jonsson and J. Thapper. Constraint satisfaction and semilinear expansions of addition over the rationals and the reals. *J. Comput. Syst. Sci.*, 82(5):912–928, 2016.

[24] A. G. Khovanskii. *Fewnominals*, volume 88 of *AMS Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1991.

[25] J. C. Lagarias. The computational complexity of simultaneous diophantine approximation problems. *SIAM J. Comput.*, 14(1):196–209, 1985.

[26] D. Macpherson. Notes on *o*-minimality and variations. In *Model theory, algebra, and geometry*, volume 39 of *Math. Sciences Research Institute*

Publications, pages 97–130. Cambridge University Press, Cambridge, 2000.

[27] D. Marker. *Model Theory: An Introduction*, volume 217 of *Graduate Texts in Mathematics*. Springer, 2002.

[28] M. V. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Mathematical Programming*, 77(1):129–162, 1997.

[29] M. Schaefer. Complexity of some geometric and topological problems. In *Proc. 17th International Symposium on Graph Drawing (GD-2009)*, pages 334–344, 2009.

[30] M. Schaefer and D. Štefankovič. Fixed points, Nash equilibria, and the existential theory of the reals. *Theory Comput. Syst.*, 60(2):172–193, 2017.

[31] T. J. Schaefer. The complexity of satisfiability problems. In *Proceedings of the 10th ACM Symposium on Theory of Computing (STOC-1978)*, pages 216–226, 1978.

[32] A. Shamir. On the cryptocomplexity of knapsack systems. In *Proceedings of the 11th ACM Symposium on Theory of Computing (STOC-1979)*, pages 118–129, 1979.

[33] A. Tarski. *A Decision Method for Elementary Algebra and Geometry*. Rand Corporation, Berkeley and Los Angeles, 1951.

[34] L. van den Dries. *Tame Topology and o-minimal Structures*. Cambridge University Press, Cambridge, 1998.

[35] D. Zhuk. The proof of CSP dichotomy conjecture. In *Proceedings of the 58th IEEE Symposium on Foundations of Computer Science (FOCS-2017)*, 2017. To appear.