On the Computability of Region-Based Euclidean Logics

Yavor Nenov and Ian Pratt-Hartmann

School of Computer Science, University of Manchester, UK
{nenovy,ipratt}@cs.man.ac.uk

Abstract. By a *Euclidean logic*, we understand a formal language whose variables range over subsets of Euclidean space, of some fixed dimension, and whose non-logical primitives have fixed meanings as geometrical properties, relations and operations involving those sets. In this paper, we consider first-order Euclidean logics with primitives for the properties of *connectedness* and *convexity*, the binary relation of *contact* and the ternary relation of *being closer-than*. We investigate the computational properties of the corresponding first-order theories when variables are taken to range over various collections of subsets of 1-, 2- and 3-dimensional space. We show that the theories based on Euclidean spaces of dimension greater than 1 can all encode either first- or second-order arithmetic, and hence are undecidable. We show that, for logics able to express the *closer-than* relation, the theories of structures based on 1-dimensional Euclidean space have the same complexities as their higher-dimensional counterparts. By contrast, in the absence of the *closer-than* predicate, all of the theories based on 1-dimensional Euclidean space considered here are decidable, but non-elementary.

1 Introduction

By a *Euclidean logic*, we understand a formal language whose variables range over subsets of $\mathbb{R}^n$ for some fixed n, and whose non-logical primitives have fixed meanings as geometrical properties, relations and operations. The motivation for studying such languages comes primarily from the field of Artificial Intelligence—the idea being that an agent's qualitative knowledge of the space it inhabits can be understood as its access to the validities of such a logic.

Euclidean logics trace their ancestry back to the region-based spatial theories developed by Whitehead [23] and de Laguna [14], later taken up within AI by Randall, Cui and Cohn [19], Egenhoffer [7] and others. The fundamental idea behind such logics is that the the natural domain of quantification for theories representing an agent's spatial knowledge is the collection of *regions* potentially occupied by physical objects, rather than the set of points constituting the space in question. Conformably, the non-logical primitives should take such regions, not points, as their relata. Whitehead, for instance, built his logic on a single spatial primitive which he referred to as *extensive connection*, where, intuitively, two regions stand in this relation if they either overlap or touch at their boundaries. Latterly, such logics have been subject to a more rigorous, model-theoretic reconstruction, in which the spatial regions in question are identified with certain subsets of some standard model of the space under investigation, and the primitive non-logical relations and operations interpreted accordingly. Thus, for example, it is now customary to reconstruct Whitehead's relation of extensive connection between regions simply as the set-theoretic relation of having a non-empty intersection. We remark that this relation is nowadays generally referred to as *contact*, to avoid confusion with the separate notion of *connectedness* encountered in topology.

One of the first issues to resolve in the context of these logics is *which* subsets of the space in question their variables should range over. By far the most popular choice in the literature is the collection of *regular closed* sets—that is, those sets which are equal to the topological closures of their *interiors*. Taking variables to range only over the regular closed sets provides a convenient way of finessing the issue of whether spatial regions are open, closed or semi-open; at the same time, since the regular closed sets of a topological space always form a Boolean algebra under inclusion, such 'regions' can be combined in a natural way. The logics obtained by interpreting Whitehead's language over (dense subalgebras of) the regular closed algebras of various classes of topological spaces were investigated by Roeper [20], Düntsch and Winter [6] and Dimov and Vakarelov [4]. In each case, the authors provided a complete axiomatization of the theory corresponding to a particular class of topological spaces. We remark that some of these results can be found in the earlier work of de Vries [3], which was motivated by purely mathematical considerations.

However, the regular closed subsets of $\mathbb{R}^2$ or $\mathbb{R}^3$ still include many pathological sets, and thus are arguably poor candidates to represent the regions of space occupied by physical objects (Pratt-Hartmann [17] , Kontchakov et al. [11]). To avoid such 'abnormal' regions, it has been proposed that Euclidean logics should instead restrict themselves to collections of "tame" regular closed sets. Candidates include: $RCS(\mathbb{R}^n)$—the regular closed *semi-algebraic* sets; $RCP(\mathbb{R}^n)$—the regular closed *semi-linear* sets (*polytopes*); $RCP_{\mathbb{A}}(\mathbb{R}^n)$—the *algebraic polytopes* (i.e. polytopes whose vertices have algebraic coordinates); and $RCP_{\mathbb{Q}}(\mathbb{R}^n)$—the *rational polytopes* (i.e. polytopes whose vertices have rational coordinates). One natural question that arises in this context is whether, for a fixed set of non-logical primitives, such restrictions actually make a difference to the resulting logics (see, e.g. Kontchakov *et al.* [12]).

Logics interpreted over Euclidean spaces may employ primitives representing non-topological notions, of course. For example, Tarski [22] considers second-order logic with variables ranging over subsets of $\mathbb{R}^3$ and predicates expressing the parthood relation and the property of being *spherical*. More recently, there has been some interest within AI in logics featuring a primitive expressing the property of *convexity* (Davis [2]) and the ternary relation of being *closer than* (Sheremet *et al.* [21]). Thus, given any combination of non-logical primitives expressing *contact*, *convexity* and *relative closeness*, and any of the domains of quantification $RCS(\mathbb{R}^n)$, $RCP(\mathbb{R}^n)$, $RCP_{\mathbb{A}}(\mathbb{R}^n)$ and $RCP_{\mathbb{Q}}(\mathbb{R}^n)$, we obtain a particular first-order theory. The aim of this paper is to establish the complexity of these theories.

One of the earliest results of this kind was proved by Grzegorczyk [8]. Taking $\mathcal{L}_C$ to be the first-order language whose only non-logical primitive is the binary predicate C, interpreted as the *contact* relation, Grzegorczyk showed that the $\mathcal{L}_C$-theories of the regular closed sets of a large class of topological spaces can encode *first-order arithmetic*, and thus are *undecidable*. In [5], Dornheim showed that the $\mathcal{L}_C$-theory of $RCP(\mathbb{R}^2)$ is *r.e.-hard*. In [2], Davis considered the Euclidean logics $\mathcal{L}_{closer}$ and $\mathcal{L}_{conv}$ with primitives expressing the relation of being *closer-than* and the property of being *convex*, respectively. He showed that the $\mathcal{L}_{closer}$-theories and the $\mathcal{L}_{conv}$-theories of $RCP_{\mathbb{Q}}(\mathbb{R}^n)$ and $RCP_{\mathbb{A}}(\mathbb{R}^n)$ can encode first-order arithmetic, and that the $\mathcal{L}_{closer}$-theories and the $\mathcal{L}_{conv}$-theories of $RCP(\mathbb{R}^n)$, $RCS(\mathbb{R}^n)$ and $RC(\mathbb{R}^n)$ can encode second-order arithmetic.

In this paper, we provide a systematic overview of results of this kind, filling in some of the gaps left by the literature. We show that, for $n = 1$, the $\mathcal{L}_C$-theories of $RC(\mathbb{R}^n)$, $RCS(\mathbb{R}^n)$, $RCP(\mathbb{R}^n)$, $RCP_{\mathbb{A}}(\mathbb{R}^n)$ and $RCP_{\mathbb{Q}}(\mathbb{R}^n)$ are decidable but not *elementary*. For $n > 1$, we show that, the corresponding $\mathcal{L}_C$-theories can encode first-order arithmetic, and that the $\mathcal{L}_C$-theory of $RC(\mathbb{R}^n)$ can encode second-order arithmetic. For $n > 0$, we establish upper complexity bounds for the $\mathcal{L}_C$-theories, the $\mathcal{L}_{closer}$-theories and the $\mathcal{L}_{conv}$-theories of $RC(\mathbb{R}^n)$, $RCS(\mathbb{R}^n)$, $RCP(\mathbb{R}^n)$, $RCP_{\mathbb{A}}(\mathbb{R}^n)$ and $RCP_{\mathbb{Q}}(\mathbb{R}^n)$. These results are summarized in Table 1.

		Signatures		
		$\langle C\rangle, \langle C, \leq\rangle, \langle c, \leq\rangle$	$\langle conv, \leq\rangle, \langle C, conv\rangle$	$\langle closer\rangle$
Domains	$RC(\mathbb{R})$	Decidable, NONELEMENTARY		Δ^1_ω-complete
	$RCS(\mathbb{R})$	Decidable, NONELEMENTARY		Δ^1_ω-complete
	$RCP(\mathbb{R})$			Δ^1_ω-complete
	$RCP_{\mathbb{A}}(\mathbb{R})$			Δ^0_ω-complete
	$RCP_{\mathbb{Q}}(\mathbb{R})$			Δ^0_ω-complete
	$RC(\mathbb{R}^n), n > 1$	Δ^1_ω-complete	Δ^1_ω-complete	Δ^1_ω-complete
	$RCS(\mathbb{R}^n), n > 1$	Δ^0_ω-hard	Δ^1_ω-complete	Δ^1_ω-complete
	$RCP(\mathbb{R}^n), n > 1$	Δ^0_ω-hard	Δ^1_ω-complete	Δ^1_ω-complete
	$RCP_{\mathbb{A}}(\mathbb{R}^n), n > 1$	Δ^0_ω-complete	Δ^0_ω-complete	Δ^0_ω-complete
	$RCP_{\mathbb{Q}}(\mathbb{R}^n), n > 1$	Δ^0_ω-complete	Δ^0_ω-complete	Δ^0_ω-complete

Table 1. A complexity map of the first-order region-based Euclidean spatial logics.

We obtain a surprising model-theoretic result from the established complexity bounds. Pratt [16] observed that the $\mathcal{L}_{conv}$-theories of $RCP(\mathbb{R}^2)$ and $RCP_{\mathbb{Q}}(\mathbb{R}^2)$ are different. The observation is based on a simple geometrical figure allowing the construction, in $RCP(\mathbb{R}^2)$, of square roots of arbitrary lengths. Because all real numbers constructable in this way are algebraic, one might be tempted to think that the $\mathcal{L}_{conv}$-theories of $RCP_{\mathbb{A}}(\mathbb{R}^2)$ and $RCP(\mathbb{R}^2)$ are the same. This, however, turns out to be false, because the two theories are shown to have different complexities.

2 Preliminaries

For basic model theoretic definitions and results, we refer to [10] ([10, p. 212] for the definition of *interpretation*). Let σ be a signature, $\mathcal{M}$ a σ-structure and $\psi(x_1, \ldots, x_n)$ a σ-formula. We define $\psi(\mathcal{M}) := \{\langle a_1, \ldots, a_n\rangle \in M \mid \mathcal{M} \models \psi[a_1, \ldots, a_n]\}$. Let σ and τ be signatures and $\mathcal{A}$ and $\mathcal{B}$ be two structures over these signatures. An *interpretation* Γ of $\mathcal{A}$ in $\mathcal{B}$ consists of:

1. for each sort i in $\mathcal{A}$, τ-formulas $\psi_{s_i}(\bar{x})$ and $\psi_{s_i \sim}(\bar{x}, \bar{y})$;
2. for each unnested atomic σ-formula $\phi(x_1^{i_1}, \ldots, x_k^{i_k})$, a τ-formula $\phi_\Gamma(\bar{x}_1, \ldots, \bar{x}_k)$;
3. a surjective mapping $f_\Gamma^i : \psi_{s_i}(\mathcal{B}) \to A_i$ for each sort i in $\mathcal{A}$, where A_i is the ith universe of $\mathcal{A}$,

such that for each unnested atomic σ-formula $\phi(x_1^{i_1}, \ldots, x_k^{i_k})$, $j = 1, \ldots, k$ and every $a_j^{i_j} \in \psi_{s_j}(\mathcal{B})$, $\mathcal{B} \models \phi_\Gamma[a_1^{i_1}, \ldots, a_k^{i_k}]$ iff $\mathcal{A} \models \phi[f_\Gamma^{i_1}(a_1^{i_1}), \ldots, f_\Gamma^{i_k}(a_k^{i_k})]$. If there exists an interpretation Γ of $\mathcal{A}$ in $\mathcal{B}$ such that $\phi \mapsto \phi_\Gamma$ is computable in polynomial time, then the theory of $\mathcal{A}$ is *polynomial-time many-one reducible* to the theory of $\mathcal{B}$ (see [10, pp. 214-215]). In this case we write $\mathcal{A} \leq_m^p \mathcal{B}$, and if in addition $\mathcal{B} \leq_m^p \mathcal{A}$, we write $\mathcal{A} \equiv_m^p \mathcal{B}$.

Let $\mathcal{X} = \langle X, \tau\rangle$ be a topological space with $\cdot^-$ and $\cdot^\circ$ the closure and interior operations in $\mathcal{X}$. A subset A of X is called *regular closed* in $\mathcal{X}$, if it equals the closure of its interior, i.e. $A = A^{\circ -}$. The set of all regular closed sets in $\mathcal{X}$ is denoted by $RC(\mathcal{X})$. $\langle RC(\mathcal{X}), +, -, \cdot, 0, 1, \leq\rangle$ is a Boolean algebra (see e.g. [13, pp. 25-28]), where, for $a, b \in RC(\mathcal{X})$, $a + b := a \cup b$, $a \cdot b := (a \cap b)^{\circ -}$, $-a := (X \setminus a)^-$ and $a \leq b$ iff $a \subseteq b$. A Boolean sub-algebra M of $RC(\mathcal{X})$ is called a *mereotopology* over $\mathcal{X}$ iff the domain of M is a closed basis for $\mathcal{X}$ (see [17, Definition 2.5]). A mereotopology over a Euclidean topological space is called a *Euclidean mereotopology.* We refer to the elements of a mereotopology as *regions.* The maximal connected subsets of $A \subseteq X$ are called *connected components* of A. A mereotopology M over $\mathcal{X}$ *respects components* iff the connected components of every $a \in M$ are also in M. M is *finitely decomposable* iff every $a \in M$ has only finitely many connected components.

In this paper, for a given Euclidean topological space, we consider different Euclidean mereotopologies. Recall that a real number is *algebraic* if it is a root of a non-zero polynomial in one variable with rational coefficients. The collection of algebraic numbers is denoted by $\mathbb{A}$. A subset of $\mathbb{R}^n$, for $n > 0$, which can be obtained by a Boolean combination of a finite number of polynomial equations and inequalities is called a *semi-algebraic* set (see e.g. [1]). If the polynomial equations and inequalities are linear, then the semi-algebraic set is called *semi-linear* or a *polytope.* A polytope whose polynomial equations and inequalities are with algebraic coefficients is called an *algebraic polytope.* Similarly, a polytope whose polynomial equations and inequalities are with rational coefficients is called a *rational polytope.* For $n > 0$, we denote by $RCS(\mathbb{R}^n)$, $RCP(\mathbb{R}^n)$, $RCP_{\mathbb{A}}(\mathbb{R}^n)$ and

$RCP_{\mathbb{Q}}(\mathbb{R}^n)$ the regular closed semi-algebraic sets, the regular closed polytopes, the regular closed algebraic polytopes and the regular closed rational polytopes. It is easy to see that $RCS(\mathbb{R}^n)$, $RCP(\mathbb{R}^n)$, $RCP_{\mathbb{A}}(\mathbb{R}^n)$ and $RCP_{\mathbb{Q}}(\mathbb{R}^n)$ are all dense Boolean subalgebras of $RC(\mathbb{R}^n)$.

For a Euclidean mereotopology M we are interested in languages able to express the property of being *connected* (denoted by c), the binary *contact* relation (denoted by C), the property of being *convex* (denoted by *conv*) and the ternary relation *closer-than* (denoted by *closer*). For $n > 0$, denote the Euclidean distance between the points $p, q \in \mathbb{R}^n$ by $d(p, q)$ and define:

$$\begin{aligned} c &:= \{a \in M \mid a \text{ is connected}\}; \quad & closer := \big\{\langle a,b,c\rangle \in M^3 \mid \\ C &:= \{\langle a,b\rangle \in M^2 \mid a \cap b \neq \emptyset\}; & glb(\{d(p,q) \mid p \in a, q \in b\}) \leq \\ conv &:= \{a \in M \mid a \text{ is convex}\}; & glb(\{d(p,q) \mid p \in a, q \in c\})\big\}. \end{aligned}$$

The following lemma characterizes the relative complexity of these properties and relations.

Lemma 1. *Let $n > 0$ and M any of $RC(\mathbb{R}^n)$, $RCS(\mathbb{R}^n)$, $RCP(\mathbb{R}^n)$, $RCP_{\mathbb{A}}(\mathbb{R}^n)$ and $RCP_{\mathbb{Q}}(\mathbb{R}^n)$. Then:*

$$\langle M, \leq, C\rangle \equiv^p_m \langle M, C\rangle; \qquad \langle M, \leq, closer\rangle \equiv^p_m \langle M, closer\rangle;$$
$$\langle M, \leq, c\rangle \equiv^p_m \langle M, C\rangle \leq^p_m \langle M, \leq, conv\rangle \leq^p_m \langle M, closer\rangle.$$

Dimov and Vakarelov [4] showed that the relation $\leq$ is definable in terms of the relation C; Davis [2] showed that the property *conv* is definable in terms of the relation *closer*; and Pratt-Hartmann [16] showed that, if M is finitely decomposable, the relation C is definable in terms of *conv* and $\leq$. It is not hard to show that the other statements also hold.

As we see in the following lemma, the regions in every Euclidean mereotopology M are determined by the rational points which they contain.

Lemma 2. *For $a, b, c \in RC(\mathbb{R}^n)$,*

$$\begin{aligned} &1) \quad a \leq b \textit{ iff } a \cap \mathbb{Q}^n \subseteq b \cap \mathbb{Q}^n;\; a = b \textit{ iff } a \cap \mathbb{Q}^n = b \cap \mathbb{Q}^n; \\ &2) \; closer(a,b,c) \textit{ iff } glb(\{d(p,q) \mid p \in a^\circ \cap \mathbb{Q}^n, q \in b^\circ \cap \mathbb{Q}^n\}) \leq \\ &\qquad\qquad glb(\{d(p,q) \mid p \in a^\circ \cap \mathbb{Q}^n, q \in c^\circ \cap \mathbb{Q}^n\}). \end{aligned}$$

The *first-order arithmetic* (FOA) and the *second-order arithmetic* (SOA) are the first- and second-order languages of the signature $\upsilon = \langle \leq, +, \cdot, 0, 1\rangle$. The *arithmetical* hierarchy Δ^0_ω and the *analytical* hierarchy Δ^1_ω comprise the sets of natural numbers which are definable in $\langle \mathbb{N}, \upsilon\rangle$ using FOA and SOA, respectively ([9]). Let T be a theory. T is in Δ^0_ω if it is many-one reducible to FOA of $\langle \mathbb{N}, \upsilon\rangle$; T is Δ^0_ω-hard if FOA of $\langle \mathbb{N}, \upsilon\rangle$ is many-one reducible to T; and T is Δ^0_ω-complete if it is in Δ^0_ω and it is Δ^0_ω-hard. The same notions are defined for Δ^1_ω.

It is a standard result that using FOA one can interpret $\langle \mathbb{Z}, \upsilon\rangle$ and $\langle \mathbb{Q}, \upsilon\rangle$ in $\langle \mathbb{N}, \upsilon\rangle$. The same can be shown for $\langle \mathbb{A}, \upsilon\rangle$. Consider the signature $\tau = \langle \leq, +, \cdot, 0, 1, \pi, [\,], N\rangle$. Let the structures $\langle \mathbb{Q}, \tau\rangle$, $\langle \mathbb{A}, \tau\rangle$ and $\langle \mathbb{R}, \tau\rangle$ be such that, $N := \mathbb{N}$; and the predicate $\pi(x)$ and the function $x[y]$ are means of encoding finite sequences of numbers. I.e. if $\pi(q)$ holds, $q[0]$ is the length of the sequence encoded by q and for every natural $1 \leq n \leq q[0]$, $q[n]$ is the nth element of the sequence encoded by q. It is not hard to see that:

Lemma 3.

1. *The first-order theory of* $\langle \mathbb{N}, +, \cdot \rangle$ *is* Δ^0_ω*-complete. The first-order theory of the two-sorted structure* $\langle \mathbb{N}, \wp(\mathbb{N}); +, \cdot, \in \rangle$ *is* Δ^1_ω*-complete.*
2. *The first-order theories of* $\langle \mathbb{Q}, \tau \rangle$ *and* $\langle \mathbb{A}, \tau \rangle$ *are* Δ^0_ω*-complete.*
3. *The first-order theory of* $\langle \mathbb{R}, \tau \rangle$ *is* Δ^1_ω*-complete.*

3 Decidable Theories

We now show that the structures $\langle RC(\mathbb{R}), \sigma \rangle$, $\langle RCS(\mathbb{R}), \sigma \rangle$, $\langle RCP(\mathbb{R}), \sigma \rangle$, $\langle RCP_{\mathbb{A}}(\mathbb{R}), \sigma \rangle$ and $\langle RCP_{\mathbb{Q}}(\mathbb{R}), \sigma \rangle$, with σ being either $\langle C \rangle$ or $\langle conv, \leq \rangle$, have decidable theories which are not elementary. Note that a non-empty regular closed set $a \in RC(\mathbb{R})$ is convex iff it is connected. So WLOG we fix $\sigma = \langle C \rangle$.

We write $\mathcal{A} \preceq \mathcal{B}$ when $\mathcal{A}$ is an elementary substructure of $\mathcal{B}$. We have that:

$$\langle RCP_{\mathbb{Q}}(\mathbb{R}), \sigma \rangle \prec \langle RCP_{\mathbb{A}}(\mathbb{R}), \sigma \rangle \prec \langle RCP(\mathbb{R}), \sigma \rangle = \langle RCS(\mathbb{R}), \sigma \rangle \leq^p_m \langle RC(\mathbb{R}), \sigma \rangle.$$

To se that the structure $\langle RCP(\mathbb{R}), \sigma \rangle$ can be interpreted in $\langle RC(\mathbb{R}), \sigma \rangle$, we only need to provide a formula $\psi_{RCP}(x)$ for which $\psi_{RCP}(RC(\mathbb{R})) = RCP(\mathbb{R})$. A regular closed set a is in $RCP(\mathbb{R})$ exactly when it has finitely many frontier points. In other words, $a \in RCP(\mathbb{R})$ just in case the set of endpoints of its connected components lie in a bounded interval, and have no accumulation points. We identify a point in $r \in \mathbb{R}$ with the unbounded connected $a \in RC(\mathbb{R})$ whose single endpoint is r. Let the formula $\psi_{cc}(x, y)$ define the pairs $\langle a, b \rangle \in RC(\mathbb{R})^2$ for which a is a connected component of b; the formula $\psi_{\perp}(x)$ define the regular closed sets encoding the points in $\mathbb{R}$; and the formula $\psi_{\odot}(x, y)$ define the pairs $\langle a, b \rangle \in RC(\mathbb{R})^2$ for which a represents a real number which is in the interior of the connected b. Then the formula $\psi_{EP}(l, x) := \psi_{\perp}(l) \wedge \exists y(\psi_{cc}(y, x) \wedge C(y, l) \wedge C(y, -l) \wedge (y \leq l \vee y \leq -l))$ defines the pairs $\langle a, b \rangle \in RC(\mathbb{R})^2$ for which a represents an endpoint of b. The formula $\psi_{Iso}(x) := \forall l\big(\psi_{\perp}(l) \rightarrow \exists u\big(\psi_{\odot}(l, u) \wedge \forall t(\psi_{EP}(t, x) \wedge \psi_{\odot}(t, u) \rightarrow l = t \vee l = -t)\big)\big)$ is satisfied by those $a \in RC(\mathbb{R})$ whose endpoints lack accumulation points. And the formula $\psi_{Bnd}(x) := \exists u\big(c(u) \wedge \neg c(-u) \wedge \forall l(\psi_{EP}(l, x) \rightarrow \psi_{\odot}(l, u))\big)$ is satisfied by those $a \in RC(\mathbb{R})$ whose endpoints are bounded. Finally, the formula $\psi_{RCP}(x) := \psi_{Iso}(x) \wedge \psi_{Bnd}(x)$ is satisfied by those $a \in RC(\mathbb{R})$ which are in $RCP(\mathbb{R})$. We have thus shown the following.

Lemma 4. $\langle RCP(\mathbb{R}), \sigma \rangle \leq^p_m \langle RC(\mathbb{R}), \sigma \rangle$

We now show that $\langle RCP_{\mathbb{Q}}(\mathbb{R}), \sigma \rangle \prec \langle RCP_{\mathbb{A}}(\mathbb{R}), \sigma \rangle \prec \langle RCP(\mathbb{R}), \sigma \rangle$. For a function $f : A \rightarrow B$ denote by $f^+ : \wp(A) \rightarrow \wp(B)$ the function defined by $f^+(a) = \{f(x) \mid x \in a\}$. Let $\mathcal{X} = \langle X, \tau \rangle$ be a topological space. Two n-tuples $\bar{a}$ and $\bar{b}$ of subsets of $\mathcal{X}$ are *similarly situated*, denoted by $\bar{a} \sim \bar{b}$, iff there is a homeomorphism $f : \mathcal{X} \rightarrow \mathcal{X}$ such that $f^+ : \bar{a} \mapsto \bar{b}$. One can easily check that for every homeomorphism $f : \mathbb{R} \to \mathbb{R}$, the function f^+ is an automorphism for the structure $\langle RCP(\mathbb{R}), C\rangle$. For every $F \subseteq \mathbb{R}$, denote by $RCP_F(\mathbb{R})$ the set of all elements in $RCP(\mathbb{R})$ with endpoints in F.

Lemma 5. *Let F be a dense subset of $\mathbb{R}$, $b \in RCP(\mathbb{R})$ and $\bar{a}$ be an n-tuple of elements in $RCP_F(\mathbb{R})$. Then there is an $a \in RCP_F(\mathbb{R})$ such that $\bar{a}a \sim \bar{a}b$.*

Corollary 1. *Let F be a dense subset of $\mathbb{R}$, $\psi(x, \bar{y})$ be a $\langle C\rangle$-formula, $\bar{a} \in RCP_F(\mathbb{R})^n$ and $b \in RCP(\mathbb{R})$. Then $RCP(\mathbb{R}) \models \psi[b, \bar{a}]$ iff there exists $a \in RCP_F(\mathbb{R})$ such that $RCP(\mathbb{R}) \models \psi[a, \bar{a}]$.*

Lemma 6. *[Tarski-Vaught Test] Let $\mathcal{A}$ be a substructure of $\mathcal{B}$. Then $\mathcal{A} \prec \mathcal{B}$ iff for every formula $\phi(x, \bar{y})$ and $\bar{a} \in A$, if there is $b \in B$ such that $\mathcal{B} \models \phi[b, \bar{a}]$ then there is $a \in A$ such that $\mathcal{B} \models \phi[a, \bar{a}]$. (See e.g. [10, Theorem 2.5.1, p. 55].)*

Corollary 2. $\langle RCP_{\mathbb{Q}}(\mathbb{R}), C\rangle \prec \langle RCP_{\mathbb{A}}(\mathbb{R}), C\rangle \prec \langle RCP(\mathbb{R}), C\rangle$.

3.1 Upper bounds

We now show that the first-order theory of $\langle RC(\mathbb{R}), C\rangle$ is decidable. We provide an interpretation of $\langle RC(\mathbb{R}), C\rangle$ in the monadic second-order theory of $\langle \mathbb{Q}, <\rangle$, which Rabin showed decidable in [18]. In the monadic second-order theory of $\langle \mathbb{Q}, <\rangle$, we may evidently define the constants $\emptyset$ and $\mathbb{Q}$, the relation $\subseteq$ and the operations $\cup$ and $\cap$.

Denote by $\tau_{\mathbb{R}}$ the set of open sets in $\mathbb{R}$ and by $\beta_{\mathbb{R}}$ the set of open connected sets in $\mathbb{R}$. Let the mapping $f : \wp(\mathbb{R}) \to \wp(\mathbb{Q})$ be defined by $f(a) = a \cap \mathbb{Q}$. We identify every region $a \in RC(\mathbb{R})$ with the set of rational points that it contains, i.e. with $f(a)$. By Lemma 2 $f \restriction RC(\mathbb{R})$ has an inverse. Note that a set $A \subseteq \mathbb{Q}$ is an f-image of some $a \in RC(\mathbb{R})$ iff A consists of exactly those $q \in \mathbb{Q}$ which are dense in A (i.e. not isolated from A).

The formula $\psi_i(X) := \forall x \forall y(X(x) \wedge X(y) \wedge x < y \to \forall z(x < z \wedge z < y \to X(z)))$ defines the f-images of *intervals*, and the formula $\psi_o = \psi_i(X) \wedge \forall x(X(x) \to \exists y \exists z(X(y) \wedge X(z) \wedge y < x \wedge x < z))$ defines the f-images of *open intervals*. The pairs $\langle q, A\rangle \in \mathbb{Q} \times \wp(\mathbb{Q})$ with q being isolated from A are defined by $\psi_{iso}(x, X) := \exists Y_1(Y_1(x) \wedge \psi_o(Y_1) \wedge \forall Y_2(\psi_o(Y_2) \wedge Y_2 \subseteq X \to Y_1 \cap Y_2 = \emptyset))$. As a result, the formula $\psi_{RC}(X) := \forall x(X(x) \leftrightarrow \neg\psi_{iso}(x, X))$ defines the set of f-images of regular closed sets in $\mathbb{R}$ (see Lemma 7).

To define the contact relation, we encode a real number r by the pair $\langle L, R\rangle \in \wp(\mathbb{Q})^2$, where $L = \{q \in \mathbb{Q} \mid q \leq r\}$ and $R = \{q \in \mathbb{Q} \mid q \geq r\}$. Clearly, a pair $\langle L, R\rangle \in \wp(\mathbb{Q})^2$ encodes a real number iff it satisfies the formula $\psi_{\mathbb{R}}(L, R) := L \cup R = \mathbb{Q} \wedge L \neq \emptyset \wedge R \neq \emptyset \wedge \forall x \forall y(L(x) \wedge R(y) \to x \leq y)$. The formula $\psi_{\odot}(L, R, X) := \psi_{\mathbb{R}}(L, R) \wedge \psi_o(X) \wedge L \cap X \neq \emptyset \wedge R \cap X \neq \emptyset$ defines the tuples $\langle A, B, D\rangle \in \wp(\mathbb{Q})^3$, such that $\langle A, B\rangle$ represent a some $r \in \mathbb{R}$ and D represents a connected open neighborhood of r. The tuples $\langle A, B, D\rangle \in \wp(\mathbb{Q})^3$ such that $\langle A, B\rangle$ represents some $r \in \mathbb{R}$ that is in the closure of D are defined by the formula $\psi_{\in}(L, R, X) := \psi_{\mathbb{R}}(L, R) \wedge \forall N(\psi_{\odot}(L, R, N) \to N \cap X \neq \emptyset)$. Finally, two regular closed sets are in contact iff their f-images satisfy the formula $\psi_C(X, Y) := \exists L \exists R(\psi_{\in}(L, R, X) \wedge \psi_{\in}(L, R, Y))$.

Lemma 7. *For $t, u \subseteq \mathbb{Q}$ we have that:*

(i) $\mathbb{Q} \models \psi_{RC}[t]$ *iff there is some* $a \in RC(\mathbb{R})$*, such that* $f(a) = t$*;*
(ii) if $\mathbb{Q} \models \psi_{RC}[t] \wedge \psi_{RC}[u]$*, then* $\mathbb{Q} \models \psi_C[t, u]$ *iff* $f^{-1}(t) C f^{-1}(u)$*.*

Lemma 8. *Γ is an interpretation of $\langle RC(\mathbb{R}), C\rangle$ in the monadic second-order theory of $\langle \mathbb{Q}, <\rangle$, where Γ consists of:*

1. *$\psi_{RC}(X)$ as a formula defining the domain;*
2. *the inverse function of $f \restriction RC(\mathbb{R})$;*
3. *$\psi_C(X, Y)$ as the formula defining the contact relation.*

Theorem 1. *[18] The monadic second-order theory of $\langle \mathbb{Q}, <\rangle$ is decidable.*

Corollary 3. *For $M \in \{RC(\mathbb{R}), RCS(\mathbb{R}), RCP(\mathbb{R}), RCP_{\mathbb{A}}(\mathbb{R}), RCP_{\mathbb{Q}}(\mathbb{R})\}$, the theories of $\langle M, C\rangle$ and $\langle M, conv, \leq\rangle$ are decidable.*

3.2 Lower bounds

We show that the theory of $\langle RCP(\mathbb{R}), \sigma\rangle$ is not elementary by introducing a polynomial reduction from the weak monadic second-order theory of one successor, denoted by $WS1S$, to the the theory of $\langle RCP(\mathbb{R}), \sigma\rangle$. Denote by $\mathcal{L}^{Mon}_{S1S}$ the monadic second-order language of the structure $\langle \mathbb{N}, S\rangle$, where $S = \{\langle n, n+1\rangle \mid n \in \mathbb{N}\}$. $WS1S$ is shown to be non-elementary by Meyer [15].

For the rest of the section, we abbreviate $M := RCP(\mathbb{R})$ and $\mathcal{M} := \langle M, \sigma\rangle$. For $n \in \mathbb{N}$, we encode the initial segment $\{0, \ldots, n\}$ of $\mathbb{N}$ by the pairs of disconnected regions $\langle a, b\rangle \in M^2$ (later defined by the formula $\psi_{\vdash}(x, y)$) such that the connected components of $a + b$ are bounded, a is non-empty and all the connected components of b are on the same side of a. A natural number $k \leq n$ is represented by the $(k + 1)$st connected component of b closest to a. A set $A \subseteq \{0, \ldots, n\}$ is represented by the sum (in $\mathcal{M}$) of the representatives of its members.

Let $\psi_{cc}(x, y)$ define the pairs $\langle a, b\rangle \in M^2$ such that a is a connected component of b. The formula

$$\psi_{ord}(x, y, z) := c(x) \wedge c(y) \wedge c(z) \wedge \neg C(y, x + z) \wedge \forall t(c(t) \wedge x + z \leq t \rightarrow y \leq t),$$

defines the tuples $\langle a, b, c\rangle \in M^3$ of connected regions such that the endpoints of b are between all of the endpoints of a and all of the endpoints of c. The formula $\psi_{\subseteq}(x, y) := \forall x'(\psi_{cc}(x', x) \rightarrow \psi_{cc}(x', y))$ defines the pairs $\langle a, b\rangle \in M^2$ such that every connected component of a is a connected component of b. The formula

$$\psi_{\vdash}(x, y) := c(x) \wedge \neg c(-x) \wedge \neg C(x, y) \wedge \forall z(\psi_{cc}(z, y) \rightarrow \neg c(-z)) \wedge \\ \forall z \forall t(\psi_{cc}(z, y) \wedge \psi_{cc}(t, y) \rightarrow \neg \psi_{ord}(z, x, t))$$

defines the pairs $\langle a, b\rangle \in M^2$ encoding an initial segment of natural numbers. The formula

$$\psi_{\leq}(x, y, z) := \psi_{\vdash}(x, y) \wedge \psi_{\vdash}(x, z) \wedge \psi_{\subseteq}(y, z) \wedge \\ \forall c \forall d(\psi_{cc}(c, y) \wedge \psi_{cc}(d, z) \wedge \psi_{ord}(x, d, c) \rightarrow \psi_{cc}(d, y))$$

defines the tuples $\langle a,b,c\rangle \in M^3$ such that $\langle a,b\rangle$ and $\langle a,c\rangle$ encode initial segments of $\mathbb{N}$, say $\{0,\ldots,n\}$ and $\{0,\ldots,m\}$, with $n \leq m$, and $\langle a,b\rangle$ and $\langle a,c\rangle$ are compatible in the sense that every $k \in \{0,\ldots,n\}$ is represented by the same region in M with respect to $\langle a,b\rangle$ and $\langle a,c\rangle$.

Let the pair $\langle a,b\rangle \in M^2$ encode an initial segment $\{0,\ldots,n\}$ of $\mathbb{N}$. Let $c,d \in M$ represent numbers $k,l \in \{0,\ldots,n\}$ and $e \in M$ represent a finite $A \subseteq \{0,\ldots,n\}$ all with respect to $\langle a,b\rangle$. Then $k \in A$ iff $\langle c,e\rangle$ satisfies the formula $\psi_\in(x,y) := \psi_{cc}(x,y)$, and $k+1=l$ iff $\langle c,d,a,b\rangle$ satisfies the formula

$$\begin{aligned}\psi_S(x,y;x_0,x_1) := \psi_{cc}(x,x_1) \wedge \psi_{cc}(y,x_1) \wedge \psi_{ord}(x_0,x,y) \wedge x \neq y \wedge \\ \forall z(\psi_{cc}(z,x_1) \wedge \psi_{ord}(x_0,z,y) \wedge z \neq y \rightarrow \psi_{ord}(x_0,z,x)).\end{aligned}$$

For every $\phi \in \mathcal{L}^{Mon}_{S1S}$, denote by $\delta(\phi)$ the *quantifier depth* of ϕ.

Definition 1. *We now define a translation* $(\cdot)_\Gamma : \mathcal{L}^{PMon}_{S1S} \to \mathcal{L}_\sigma$*, where* $\mathcal{L}^{PMon}_{S1S}$ *is the set of all formulas in the language* $\mathcal{L}^{Mon}_{S1S}$ *that are in prenex normal form. We use the special variables* $r, s_0, \ldots, s_{\delta(\phi)}$*.*

$$\begin{aligned}
(x_n = x_m)_\Gamma &:= p_n = p_m;\ (X_n = X_m)_\Gamma := q_n = q_m;\\
(S(x_n,x_m))_\Gamma &:= \psi_S(p_n,p_m;\ r,s_0); (x_n \in X_m)_\Gamma := \psi_\in(p_n,q_m);\\
(\neg\psi)_\Gamma &:= \neg\psi_\Gamma;\ (\psi' \wedge \psi'')_\Gamma := \psi'_\Gamma \wedge \psi''_\Gamma;\\
(\exists x_n\psi)_\Gamma &:= \exists p_n \exists s_{\delta(\psi)} \big(\psi_\leq(r,s_{\delta(\psi)+1},s_{\delta(\psi)}) \wedge \psi_{cc}(p_n,s_{\delta(\psi)}) \wedge \psi_\Gamma\big),\\
(\forall x_n\psi)_\Gamma &:= \forall p_n \forall s_{\delta(\psi)} \big(\psi_\leq(r,s_{\delta(\psi)+1},s_{\delta(\psi)}) \wedge \psi_{cc}(p_n,s_{\delta(\psi)}) \to \psi_\Gamma\big),\\
(\exists X_n\psi)_\Gamma &:= \exists q_n \exists s_{\delta(\psi)} \big(\psi_\leq(r,s_{\delta(\psi)+1},s_{\delta(\psi)}) \wedge \psi_\subseteq(q_n,s_{\delta(\psi)}) \wedge \psi_\Gamma\big),\\
(\forall X_n\psi)_\Gamma &:= \forall q_n \forall s_{\delta(\psi)} \big(\psi_\leq(r,s_{\delta(\psi)+1},s_{\delta(\psi)}) \wedge \psi_\subseteq(q_n,s_{\delta(\psi)}) \to \psi_\Gamma\big).
\end{aligned}$$

Lemma 9. *For every* $\phi \in \mathcal{L}^{PMon}_{S1S}$*,* $\mathbb{N} \models \phi$ *iff* $\mathcal{M} \models \forall r \forall s_{\delta(\phi)} \psi_\vdash(r,s_{\delta(\phi)}) \to \psi_\Gamma$*.*

Theorem 2. *[15] WS1S is not elementary.*

Corollary 4. *The first-order theory of* $\langle RCP(\mathbb{R}), C\rangle$ *is not elementary.*

4 Undecidable Theories

In this section we establish upper and lower bounds on the complexities of some undecidable region-based theories of space. In particular, we show that the theory of $\langle RC(\mathbb{R}^n),\sigma\rangle$ with $n>1$ and σ any of $\langle C\rangle$, $\langle conv,\leq\rangle$ and $\langle closer\rangle$ is Δ^1_ω-complete. Further, the theories of $RCP_{\mathbb{A}}(\mathbb{R}^n)$ and $RCP_{\mathbb{Q}}(\mathbb{R}^n)$, with the same non-logical primitives, are all Δ^0_ω-complete. Although we show that the theories of $\langle RCS(\mathbb{R}^n),\sigma\rangle$ and $\langle RCP(\mathbb{R}^n),\sigma\rangle$ are Δ^0_ω-hard, their precise complexity remains open. In contrast to the decidable one-dimensional region-based theories of space considered in Section 3, we shall show that the theories of $\langle RC(\mathbb{R}), closer\rangle$, $\langle RCS(\mathbb{R}), closer\rangle$ and $\langle RCP(\mathbb{R}), closer\rangle$ are Δ^1_ω-complete and that the theories of $\langle RCP_{\mathbb{A}}(\mathbb{R}), closer\rangle$ and $\langle RCP_{\mathbb{Q}}(\mathbb{R}), closer\rangle$ are Δ^0_ω-complete.

4.1 Lower Bounds

Combining ideas from [2] and [8], we show that, for $\sigma = \langle C, c, +, \cdot, -, \leq \rangle$ and $n > 1$, the theory of every extension of $\langle RCP_{\mathbb{Q}}(\mathbb{R}^n), \sigma \rangle$ is Δ^0_ω-hard, and that the theory of $\langle RC(\mathbb{R}^n), \sigma \rangle$ is Δ^1_ω-hard.

Fix an extension $\mathcal{M} = \langle M, \sigma \rangle$ of $\langle RCP_{\mathbb{Q}}(\mathbb{R}^n), \sigma \rangle$, for some $n > 1$. For every $a \in M$, denote by $|a|$ the number of connected components of a. We interpret the structure $\langle \mathbb{N}, +, \cdot \rangle$ in $\mathcal{M}$, by encoding any natural number k as a region $a \in M$ having k connected components, i.e. with $|a| = k$.

The first step is to provide a formula $\psi_{\sim}(x, y)$ which is satisfied by the pairs $\langle a, b \rangle \in M$ having the same number of connected components. Let the formula $\psi_{cc}(x, y)$ define the pairs of regions $\langle a, b \rangle \in M^2$, with a being a connected component of b. For $a, b \in M$, we say that a is a *shrinking* of b if and only if every connected component of a is contained in a connected component of b and every connected component of b contains exactly one connected component of a. The formula $\psi_{shrink}(x, y) := x \leq y \wedge \forall y'(\psi_{cc}(y', y) \rightarrow \psi_{cc}(x \cdot y', x))$ defines the pairs of regions $\langle a, b \rangle \in M^2$ such that a is a shrinking of b.

For $a, b, c \in M$ we say that c is a *wrapping* of a and b if and only if $a + b \leq c$ and every connected component of c contains one connected component of a and b. The formula

$$\psi_{d\sim}(x, y) := \exists z(x + y \leq z \wedge \forall z'(\psi_{cc}(z', z) \rightarrow \psi_{cc}(x \cdot z', x) \wedge \psi_{cc}(y \cdot z', y))$$

defines the pairs of regions $\langle a, b \rangle \in M^2$ for which there exists a wrapping $c \in M$.

Lemma 10. *Let $a, b \in M$. Then $|a| = |b|$ if and only if there exist $a', b', c \in M$ such that a' and b' are shrinkings of a and b and c is a wrapping of a' and b'.*

So, the formula $\psi_{\sim}(x, y) := \exists x' \exists y'(\psi_{shrink}(x', x) \wedge \psi_{shrink}(y', y) \wedge \psi_{d\sim}(x', y'))$ defines the pairs $\langle a, b \rangle \in M^2$ such that $|a| = |b|$. The formula $\psi_S(x, y) := \exists x'(\psi_{cc}(x', x) \wedge \psi_{\sim}(x \cdot -x', y))$ defines the pairs of regions $\langle a, b \rangle \in M^2$ with $|a| = |b| + 1$ (taking $\aleph_0 + 1 = \aleph_0$), and the formula $\psi_{fin}(x) := \neg\psi_S(x, x)$ defines the regions having finitely many components. Clearly,

Lemma 11. *The function $f_\Gamma : \psi_{fin}(\mathcal{M}) \rightarrow \mathbb{N}$ defined by $f_\Gamma(a) = |a|$ is surjective.*

The following formulas define the arithmetical operations on numbers:

$\psi_+(x, y, z) := \exists x' \exists y'(\psi_{\sim}(x, x') \wedge \psi_{\sim}(y, y') \wedge \neg C(x', y') \wedge x' + y' = z)$ and
$\psi_\times(x, y, z) := \exists u \exists v[\psi_{shrink}(u, z) \wedge u \leq v \wedge \psi_{\sim}(v, y) \wedge \forall t(\psi_{cc}(t, v) \rightarrow \psi_{\sim}(t \cdot u, x))]$.

Lemma 12. *Γ is an interpretation of $\langle \mathbb{N}, +, \cdot \rangle$ in $\mathcal{M}$, where Γ consists of:*

1. *the σ-formulas $\psi_{fin}(x)$ and $\psi_{\sim}(x, y)$;*
2. *the σ-formulas $\psi_+(x, y, z)$ and $\psi_\times(x, y, z)$;*
3. *the surjective map: f_Γ.*

Corollary 5. *Let τ be any of $\langle C \rangle$, $\langle conv, \leq \rangle$ and $\langle closer \rangle$, $n > 1$ and M be any of $RC(\mathbb{R}^n)$, $RCS(\mathbb{R}^n)$, $RCP(\mathbb{R}^n)$, $RCP_{\mathbb{A}}(\mathbb{R}^n)$ and $RCP_{\mathbb{Q}}(\mathbb{R}^n)$. Then the theory of $\langle M, \tau \rangle$ is Δ^0_ω-hard.*

We now show that when $M = RC(\mathbb{R}^n)$, $\langle M, \sigma \rangle$ can interpret $\langle \mathbb{N}, \wp(\mathbb{N}); +, \cdot, \in \rangle$. We identify every set $A \subseteq \mathbb{N}$ with a pair of regions $\langle a, b \rangle \in M^2$ such that, for every $k \in \mathbb{N}$, $k \in A$ if and only if there exists a connected component a' of a with $|a' \cdot b| = k$. The collection of pairs $\langle a, b \rangle \in M^2$ that represent a set of natural numbers is thus defined by the formula $\psi_{set}(x, y) := \forall x'(\psi_{cc}(x', x) \to \psi_{fin}(x' \cdot y))$. The formula $\psi_{\in}(z; x, y) := \exists x'(\psi_{cc}(x', x) \wedge \psi_{\sim}(z, x' \cdot y))$ likewise defines a set of triples $\langle a, b, c \rangle \in M^3$ such that, if a represents the natural number k and $\langle b, c \rangle$ represents a set of natural numbers A, then $k \in A$.

Lemma 13. *The function $f'_\Gamma : \psi_{set}(\mathcal{M}) \to \wp(\mathbb{N})$ is surjective, where $f'_\Gamma(a, b) = \{|a' \cdot b| : a'$ a connected component of $a\}$.*

Lemma 14. *Let $a, b, c \in M$. If $\mathcal{M} \models \psi_{fin}[a]$ and $\mathcal{M} \models \psi_{set}[b, c]$ then $\mathcal{M} \models \psi_{\in}[a, b, c]$ iff $f_\Gamma(a) \in f'_\Gamma(b, c)$.*

Lemma 15. *Γ is an interpretation of $\langle \mathbb{N}, \wp(\mathbb{N}); +, \cdot, \in \rangle$ in $\mathcal{M}$, where Γ consists of:*

1. *the formulas $\psi_{fin}(x)$, $\psi_{set}(x, y)$, $\psi_{\sim}(x, y)$ and $\psi_{set\sim}(x, y, x', y')$;*
2. *the formulas $\psi_{+}(x, y, z), \psi_{\times}(x, y, z)$ and $\psi_{\in}(x, y, z)$;*
3. *the surjective maps: f_Γ, f'_Γ.*

Corollary 6. *Let σ be any of $\langle C \rangle$, $\langle covn, \leq \rangle$ and $\langle closer \rangle$ and $n > 1$. Then the theory of $\langle RC(\mathbb{R}^n), \sigma \rangle$ is Δ^1_ω-hard.*

Some of the theories from Corollary 5 and Corollary 6 are known to have even higher computational complexities. In particular:

Lemma 16. *[2] Let σ be any of $\langle conv, \leq \rangle$ and $\langle closer \rangle$, and let $n \geq 1$. Then*

- *$\langle RC(\mathbb{R}), closer \rangle$, $\langle RCS(\mathbb{R}), closer \rangle$ and $\langle RCP(\mathbb{R}), closer \rangle$ are Δ^1_ω-hard;*
- *$\langle RCP_{\mathbb{A}}(\mathbb{R}), closer \rangle$ and $\langle RCP_{\mathbb{Q}}(\mathbb{R}), closer \rangle$ are Δ^0_ω-hard;*
- *$\langle RCS(\mathbb{R}^n), \sigma \rangle$ and $\langle RCP(\mathbb{R}^n), \sigma \rangle$ are Δ^1_ω-hard*

4.2 Upper Bounds

We now show that for $n > 0$ the theories of the structures $\langle RCP(\mathbb{R}^n), closer \rangle$, $\langle RCP_{\mathbb{A}}(\mathbb{R}^n), closer \rangle$ and $\langle RCP_{\mathbb{Q}}(\mathbb{R}^n), closer \rangle$ are interpretable in the structures $\langle \mathbb{R}, \tau \rangle$, $\langle \mathbb{A}, \tau \rangle$ and $\langle \mathbb{Q}, \tau \rangle$, respectively, where $\tau = \langle \leq, +, \cdot, 0, 1, \pi, [\ ], N \rangle$. In the sequel, we take R to range over the fields $\mathbb{R}$, $\mathbb{A}$ and $\mathbb{Q}$, writing $RCP(\mathbb{R}^n)$ alternatively as $RCP_{\mathbb{R}}(\mathbb{R}^n)$. We denote the structure $\langle R, \tau \rangle$ by $\mathcal{R}$. Note that the regions in $RCP_R(\mathbb{R}^n)$ are exactly the sums of finitely many products of finitely many half-spaces whose boundaries are $(n-1)$-dimensional hyperplanes definable by degree 1 polynomials in $R[X_1, \ldots, X_n]$. So, for every sequence of sequences $s = \langle \langle a_{11}, \ldots, a_{1m_1} \rangle, \ldots, \langle a_{m1}, \ldots, a_{mm_m} \rangle \rangle$ of half-spaces in $RCP_R(\mathbb{R}^n)$, there is an unique region $a \in RCP_R(\mathbb{R}^n)$ such that $a = \sum_{i=1}^{m} \prod_{j=1}^{m_i} a_{ij}$. And conversely, every region $a \in RCP_R(\mathbb{R}^n)$ is represented by some (in fact infinitely many) sequences of that form. Thus, we may encode elements of $RCP_R(\mathbb{R}^n)$ using sequences of sequences of half-spaces. Further, since each half-space is defined by a linear equation, we may encode it as the sequence of its coefficients, so that elements of $RCP_R(\mathbb{R}^n)$ may be encoded as sequences of sequences of sequences of real numbers. Of course, we may represent points in R^n by the sequence of real numbers in the obvious way. Let $\psi_\bullet(x)$, $\psi_{PT}(x)$, $\psi_{\in PT}(x,y)$ and $\psi_{\in PT^\circ}(x,y)$ be τ-formulas such that: $\psi_\bullet(x)$ defines those $r \in R$ that encode points in R^n; $\psi_{PT}(x)$ defines those $r \in R$ that encode regions in $RCP_R(\mathbb{R}^n)$; $\psi_{\in PT}(x,y)$ defines the pairs $\langle r,s\rangle \in R$ such that r encodes a point in the region encoded by s; and $\psi_{\in PT^\circ}(x,y)$ defines the pairs $\langle r,s\rangle$ such that r encodes a point in R^n in the interior of the region in $RCP_R(\mathbb{R}^n)$ encoded by s.

By Lemma 2, we get that the τ-formula $\psi_\leq(x,y) := \forall z(\psi_\bullet(z) \wedge \psi_{\in PT}(z,x) \rightarrow \psi_{\in PT}(z,y))$ defines the part-of relation, and that the τ-formula $\psi_{closer}(x,y,z)$ defines the relation closer-than, where:

$$\psi_{closer}(x,y,z) := \forall p \forall q \big(\psi_{\in PT^\circ}(p,x) \wedge \psi_{\in PT^\circ}(q,z) \rightarrow \exists r \exists s \big(\psi_{\in PT^\circ}(r,x) \wedge \psi_{\in PT^\circ}(s,y) \wedge \sum_{i=1}^{n}(r[i]-s[i])^2 \leq \sum_{i=1}^{n}(p[i]-q[i])^2\big)\big).$$

Lemma 17. *For $n > 0$, Γ is an interpretation of $\langle RCP_R(\mathbb{R}^n), closer\rangle$ in $\mathcal{R}$, where Γ consists of:*

1. *the formulas $\psi_{PT}(x)$ and $\psi_{PT\sim}(x,y) := \psi_\leq(x,y) \wedge \psi_\leq(y,x)$;*
2. *the formula $\psi_{closer}(x,y,z)$ corresponding to the closer relation;*
3. *the surjective map $f : \psi_{PT}(\mathcal{R}) \rightarrow RCP_R(\mathbb{R}^n)$ defined by:*

$$f(a) = \sum_{i=1}^{a[0]} \prod_{j=1}^{a[i][0]} \big\{\langle x_1,\ldots,x_n\rangle \in \mathbb{R}^n \big| \sum_{k=1}^{n} a[i][j][k]\cdot x_k + a[i][j][n+1] \leq 0\big\}.$$

Corollary 7. *Let $n > 0$ and σ be any of the signatures $\langle C\rangle$, $\langle conv, \leq\rangle$ and $\langle closer\rangle$. Then $\langle RCP_{\mathbb{Q}}(\mathbb{R}^n), \sigma\rangle$ and $\langle RCP_{\mathbb{A}}(\mathbb{R}^n), \sigma\rangle$ are in Δ^0_ω and $\langle RCP(\mathbb{R}^n), \sigma\rangle$ is in Δ^1_ω.*

Now, as we promised in the introduction, for $n > 1$ and $\sigma = \langle conv, \leq\rangle$, we obtain from the computational properties of $\langle RCP(\mathbb{R}^n), \sigma\rangle$ and $\langle RCP_{\mathbb{A}}(\mathbb{R}^n), \sigma\rangle$ a very interesting and surprising model-theoretic result.

Corollary 8. *The structure $\langle RCP_{\mathbb{A}}(\mathbb{R}), conv, \leq\rangle$ is not an elementary substructure of the structure $\langle RCP(\mathbb{R}), conv, \leq\rangle$.*

We now show that the structures $\langle RCS(\mathbb{R}^n), closer\rangle$, for $n > 0$, are definable in the structure $\mathcal{R} = \langle \mathbb{R}, \tau\rangle$. Fix a positive $n \in \mathbb{N}$. Recall that a subset of $\mathbb{R}^n$ is semi-algebraic if it is definable by a Boolean combination of finite number of polynomial equations and inequalities. Following essentially the same procedure as for semi-linear sets, it is routine to encode semi-algebraic sets as real numbers. All that then remains to do is to show that we can write a formula defining the real numbers that encode semi-algebraic sets which are also *regular closed*.

Let the formula $\psi_\bullet(x)$ define the real numbers encoding points in $\mathbb{R}^n$; let the formula $\psi_{SA}(x)$ define the set of real numbers encoding semi-algebraic sets in

$\mathbb{R}^n$; and let the formula $\psi_{\in SA}(x, y)$ define the pair $\langle r, s\rangle \in \mathbb{R}^2$ such that r encodes a point in $\mathbb{R}^n$ which is in a semi-algebraic set encoded by s. We use n-balls to determine if a semi-algebraic sets is regular closed. We identify each n-ball with the $n + 1$ coefficients of its inequality. A real number encodes an n-ball iff it satisfies the formula $\psi_{\bigcirc}(x) := \pi(x) \wedge x[0] = n+1$, and a point, encoded by some $p \in \mathbb{R}$, lies in the interior of an n-ball, encoded by some $o \in \mathbb{R}$, iff the pair $\langle p, o\rangle$ satisfies the formula $\psi_{\in\bigcirc}(x, y) := \sum_{i=1}^{n}(x[i] - y[i])^2 < x[n + 1]$.

A point, encoded by $p \in \mathbb{R}$, is isolated from a semi-algebraic set, encoded by $o \in \mathbb{R}$, iff $\langle p, o\rangle$ satisfies the formula:

$$\psi_{\odot}(x, y) := \exists z(\psi_{\bigcirc}(z) \wedge \psi_{\in\bigcirc}(x, z) \wedge \forall t(\psi_{\bullet}(t) \wedge \psi_{\in\bigcirc}(t, z) \rightarrow \neg\psi_{\in SA}(t, y))).$$

A set $A \subseteq \mathbb{R}^n$ is regular closed iff it contains exactly the points that are dense in it. So the formula $\psi_{RCS}(x) := \psi_{SA}(x) \wedge \forall y(\psi_{\in\bigcirc}(y, x) \leftrightarrow \neg\psi_{\odot}(y, x))$ defines exactly the codes of the regular closed semi-algebraic sets. Two numbers encode the same regular closed semi-algebraic sets if they satisfy the formula: $\psi_{RCS\sim}(x, y) := \forall t(\psi_{\bullet}(t) \rightarrow (\psi_{\in SAS}(t, x) \leftrightarrow \psi_{\in SA}(t, y)))$. One can easily write a formula $\psi_{closer}(x, y, z)$ defining the relation closer-than.

Lemma 18. *For $n > 0$, Γ is an interpretation of $\langle RCS(\mathbb{R}^n), closer\rangle$ in $\mathcal{R}$, where Γ consists of:*

1. *the formulas $\psi_{RCS}(x)$ and $\psi_{RCS\sim}(x, y)$;*
2. *the formula $\psi_{closer}(x, y, z)$;*
3. *the surjective map $f : \psi_{RCS}(\mathcal{R}) \to RCS(\mathbb{R}^n)$ defined by:*

$$f(a) = \{\langle k[1], \ldots, k[n]\rangle \mid \mathcal{R} \models \psi_{\bullet}[k], \mathcal{R} \models \psi_{\bullet\in SAS}[k, a]\}.$$

Corollary 9. *Let $n > 0$ and σ be any of the signatures $\langle C\rangle$, $\langle conv, \leq\rangle$ and $\langle closer\rangle$. Then the theory of $\langle RCS(\mathbb{R}^n), \sigma\rangle$ is in Δ^1_ω.*

We now show that, for $n > 0$, the structures $\langle RC(\mathbb{R}^n), closer\rangle$ can be interpreted in the second-order theory of $\mathcal{Q} = \langle \mathbb{Q}, \tau\rangle$. We identify every region in $RC(\mathbb{R}^n)$ with the set of rational points that it contains, and we make use of the fact that a subset of $\mathbb{R}^n$ is regular closed iff it contains exactly the points that are dense in it. A point $p \in \mathbb{R}^n$ is dense in $A \subseteq \mathbb{R}^n$ iff every open neighborhood of p intersects A. Define $f^n : \wp(\mathbb{R}^n) \to \wp(\mathbb{Q}^n)$ by $f^n(a) := a \cap \mathbb{Q}^n$.

Let $\psi_{\bullet}(x)$, $\psi_{\bigcirc}(x)$ and $\psi_{\in\bigcirc}(x, y)$ be as in the case of semi-algebraic sets. The formula $\psi_{\odot}(x, X) := \exists z(\psi_{\bigcirc}(z) \wedge \psi_{\in\bigcirc}(x, z) \wedge \forall t(\neg(\psi_{\in\bigcirc}(t, z) \wedge t \in X)))$ defines the pairs $\langle q, A\rangle \in \mathbb{Q} \times \wp(\mathbb{Q})$ such that q encodes a point which is isolated from the set of points encoded by the members of A.

A set A of rational points is an f^n-image of a regular closed set iff it contains exactly the rational points that are dense in it iff A satisfies

$$\psi_{RC}(X) := \forall x(x \in X \rightarrow \psi_{\bullet}(x)) \wedge \forall x(x \in X \leftrightarrow \neg\psi_{\odot}(x, X)).$$

One can easily find a formula $\psi_{closer}(X, Y, Z)$ defining the relation closer-than.

Lemma 19. *For $n > 0$, Γ is an interpretation of $\langle RC(\mathbb{R}^n), closer\rangle$ in $\mathcal{Q}$, where Γ consists of:*

1. *the formulas $\psi_{RC}(X)$ and $\psi_{RC\sim}(X, Y) := \forall x(x \in X \leftrightarrow x \in Y)$;*
2. *the formula $\psi_{closer}(X, Y, Z)$ corresponding to the closer relation;*
3. *the inverse of f^n as a surjective map.*

Corollary 10. *Let σ be any of $\langle C\rangle$, $\langle conv, \leq\rangle$ and $\langle closer\rangle$. Then the theory of $\langle RC(\mathbb{R}^n), \sigma\rangle$ is in Δ^1_ω.*

5 Conclusions

In this paper we examined the complexity of the first-order theories of some region-based Euclidean spatial logics. We showed that the spatial logic for expressing the *contact* relation is decidable but not elementary, when interpreted over $\mathbb{R}$. We showed that the same logic interpreted over $\mathbb{R}^n$, for $n > 1$, can encode first-order arithmetic, and when regions with infinitely many components are allowed, second-order arithmetic as well. These lower complexity bounds also hold for more expressive logics such as those able to express the property of *convexity* or the relation *closer-than*. It was shown in [2] that when polytopes with vertices having transcendental coordinates are allowed, these logics have complexities no less than that of second-order arithmetic. It also follows from [2] that, the complexities of the spatial logics which are able to express the *closer-than* relation are not influenced by the dimension of the space over which they are interpreted.

We showed that all structures with countable domains are definable in first-order arithmetic and that all others are definable in second-order arithmetic. This yields precise complexity bounds for all our structures but $\langle RCS(\mathbb{R}^n), C\rangle$ and $\langle RCP(\mathbb{R}^n), C\rangle$, where $n > 1$. For $n = 2$, the theories of these structures are the same as the theory of $\langle RCP_{\mathbb{Q}}(\mathbb{R}^2), C\rangle$ (see [17]), which makes them complete with respect to first-order arithmetic. However, for $n > 2$, the precise complexity the theories of $\langle RCS(\mathbb{R}^n), C\rangle$ and $\langle RCP(\mathbb{R}^n), C\rangle$ remains open.

From the established complexity bounds we obtain an interesting and surprising model-theoretic result — namely that $\langle RCP_{\mathbb{A}}(\mathbb{R}^2), conv, \leq\rangle$ is not an elementary substructure of $\langle RCP(\mathbb{R}^2), conv, \leq\rangle$.

Acknowledgements. The work on this paper was supported by the UK EPSRC research grants EP/P504724/1 and EP/E035248/1.

References

1. J. Bochnak, M. Coste, and M.F. Roy. *Real algebraic geometry.* Springer Verlag, 1998.
2. Ernest Davis. The expressivity of quantifying over regions. *Journal of Logic and Computation*, 16(6):891–916, 2006.
3. H. de Vries. *Compact Spaces and Compactifications: an algebraic approach.* Van Gorcum, Assen, 1962.

4. Georgi Dimov and Dimiter Vakarelov. Contact algebras and region-based theory of space: Proximity approach - I. *Fundamenta Mathematicae*, 74(2,3):209–249, 2006.
5. Christoph Dornheim. Undecidability of plane polygonal mereotopology. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixth International Conference*, pages 342–355, 1998.
6. Ivo Düntsch and Michael Winter. A representation theorem for boolean contact algebras. *Theor. Comput. Sci.*, 347(3):498–512, 2005.
7. M. J. Egenhofer. Reasoning about binary topological relations. In *Advances in Spatial databases, Proceedings (SSD '91)*, Lecture Notes in Computer Science 525, pages 143–160, Berlin, 1991. Springer.
8. Andrzej Grzegorczyk. Undecidability of some topological theories. *Fundamenta Mathematicae*, 38:137–152, 1951.
9. P.G. Hinman and M. Frontmatter. *Recursion-theoretic hierarchies.* New York, 1978.
10. W. Hodges. *Model theory.* Cambridge Univ Pr, 1993.
11. Roman Kontchakov, Ian Pratt-Hartmann, Frank Wolter, and Michael Zakharyaschev. On the computational complexity of spatial logics with connectedness constraints. In *LPAR '08: Proceedings of the 15th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning*, pages 574–589, Berlin, Heidelberg, 2008. Springer-Verlag.
12. Roman Kontchakov, Ian Pratt-Hartmann, and Michael Zakharyaschev. Interpreting topological logics over Euclidean spaces. In *Principles of Knowledge Representation and Reasoning: Proc. of the 12th International Conference*, pages 534–544. AAAI Press, 2010.
13. S. Koppelberg, R. Bonnet, and J.D. Monk. *Handbook of Boolean Algebras Vol. 1.* Elsevier, 1989.
14. Theodore de Laguna. Point, line, and surface, as sets of solids. *The Journal of Philosophy*, 19(17):449–461, 1922.
15. Albert R. Meyer. Weak monadic second order theory of succesor is not elementary-recursive. In *Logic Colloquium*, volume 453 of *Lecture Notes in Mathematics*, pages 132–154. Springer Berlin / Heidelberg, 1975.
16. Ian Pratt. First-order qualitative spatial representation languages with convexity. *Spatial Cognition and Computation*, 1(2):181–204, June 1999.
17. Ian Pratt-Hartmann. First-order mereotopology. In *Handbook of Spatial Logics*, pages 13–97. Springer Netherlands, 2007.
18. M.O. Rabin. Decidability of second-order theories and automata on infinite trees. *Transactions of the American Mathematical Society*, 141(1-35):4, 1969.
19. D. Randell, Z. Cui, and A. Cohn. A spatial logic based on regions and connection. In B. Nebel, C. Rich, and W. Swartout, editors, *Principles of Knowledge Representation and Reasoning, Proc. 3rd International Conference*, pages 165–176, San Franciso, 1992. Morgan Kaufmann.
20. P. Roeper. Region-Based Topology. *Journal of Philosophical Logic*, 26(3):251–309, 1997.
21. M. Sheremet, D. Tishkovsky, F. Wolter, and M. Zakharyaschev. From topology to metric: modal logic and quantification in metric spaces. In I. Hodkinson and Y. Venema, editors, *Advances in Modal Logic*, volume 6, pages 429–448, London, 2006. College Publications.
22. A. Tarski. Foundations of the geometry of solids. *Logic, semantics, metamathematics*, pages 24–29, 1956.
23. Alfred North Whitehead. *Process and reality : an essay in cosmology.* Cambridge University Press, Cambridge :, 1929.