Discrete Comput Geom 30:607–622 (2003)
DOI: 10.1007/s00454-003-0770-2

Discrete & Computational
Geometry


# Expressing the Box Cone Radius in the Relational Calculus with Real Polynomial Constraints

Floris Geerts*

Department WNI,
University of Limburg (LUC),
B-3590, Diepenbeek, Belgium

**Abstract.** We show that there is a query expressible in first-order logic over the reals that returns, on any given semi-algebraic set $A$, for every point, a radius around which $A$ is conical in every small enough box. We obtain this result by combining results from differential topology and real algebraic geometry, with recent algorithmic results by Rannou.

## 1. Introduction

The framework of constraint databases, introduced by Kanellakis et al. [13], provides a nice theoretical model for spatial databases [19]. A spatial dataset is modeled using real polynomial inequality constraints; such sets are also known as semi-algebraic sets [1], [4]. The relational calculus (first-order logic) with real polynomial constraints then serves as a basic query language, denoted here by FO.

The study of the expressive power of query languages for constraint databases is an active domain of research [18]. One of the problems in particular that received attention in recent years is that of determining the exact power of FO in expressing topological properties of spatial databases [3], [8], [11], [16], [21]. One such well-known property, which is central in this research, is that inside a small enough ball around each point, a semi-algebraic set has the topology of a cone. A radius at which this behavior shows is called a *cone radius* in the point for the set.

In this paper we prove a stronger property of semi-algebraic sets. We prove that inside *any* box located around the center of and completely included in a small enough ball around each point, a semi-algebraic set has the topology of a cone. A radius at which

* Current address: Helsinki Institute for Information Technology, Basic Research Unit, PO Box 26, University of Helsinki, FIN-00014, Finland. floris.geerts@cs.helsinki.fi.

this behavior shows is called a *box cone radius* in the point for the set. The existence of such a radius was already proven for semi-algebraic sets in $\mathbb{R}^2$, but the methods of the proof do not generalize to arbitrary dimensions [16].

Accordingly, a (box) cone radius query is a query that returns, for a semi-algebraic set $A$ in $n$-dimensional space $\mathbb{R}^n$, a set of pairs $(\vec{p}, r)$ giving for every point $\vec{p}$ a (box) cone radius $r$ in $\vec{p}$ for $A$. In this paper we show that there exists an FO formula expressing a (box) cone radius query. Again, the (box) cone radius query was shown to be expressible in FO for semi-algebraic sets in $\mathbb{R}^2$, but the methods of the proof do not generalize to arbitrary dimensions [8].

Expressibility of the (box) cone radius, apart from being a natural question in itself, also has applications. Indeed, the expressive power of FO in expressing topological properties is rather limited. For example, topological connectivity is not expressible in FO [2]. Therefore, recursive extensions of FO have been studied in order to express more queries [12], [15], [17], [14], [7], [6]. In particular, the question whether topological connectivity is expressible in FO + TC, the extension of FO with a transitive closure operator [15], was raised. This question was first answered affirmatively for linear spatial databases in $\mathbb{R}^n$ [15]. Later, this result was extended to quadratic spatial databases in $\mathbb{R}^2$ [17], and then to arbitrary closed spatial databases in $\mathbb{R}^2$ [8]. This last result was obtained by expressing the cone radius query for closed spatial databases in FO. In our companion paper [9] it is shown that for arbitrary spatial databases in $\mathbb{R}^n$, the expressibility of the box cone radius query implies that a piecewise linear spatial database can be constructed in FO + TC, which has the same topological properties as the original database. Hence, the question whether a spatial databases is connected, can be reduced to the question whether a linear spatial database is connected. Since this last question is expressible in FO + TC, topological connectivity of spatial databases in $\mathbb{R}^n$ is also expressible in FO + TC. More generally, using the expressibility of the box cone radius query, one can show that any computable topological query can be expressed in FO + TCS, a variant of FO + TC in which one can control the termination behavior of the transitive closure operator [9]. Again, this generalizes the result obtained for spatial databases in $\mathbb{R}^2$ [7].

## 2. Preliminaries

### 2.1. *Spatial Databases and Queries*

A *semi-algebraic set in* $\mathbb{R}^n$ is a finite union of sets definable by conditions of the form

$$f_1(\vec{x}) = f_2(\vec{x}) = \cdots = f_k(\vec{x}) = 0, \qquad g_1(\vec{x}) > 0,\ g_2(\vec{x}) > 0, \ldots,\ g_\ell(\vec{x}) > 0,$$

with $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, and where $f_1(\vec{x}), \ldots, f_k(\vec{x}), g_1(\vec{x}), \ldots, g_\ell(\vec{x})$ are multivariate polynomials in the variables $x_1, \ldots, x_n$ with real coefficients. A *database schema* $\mathcal{S}$ is a finite set of relation names, each with a given arity. A *database* over $\mathcal{S}$ assigns to each $S \in \mathcal{S}$ a semi-algebraic set $S^D$ in $\mathbb{R}^n$ if $n$ is the arity of $S$. A *k-ary query over* $\mathcal{S}$ is a function mapping each database over $\mathcal{S}$ to a semi-algebraic set in $\mathbb{R}^k$.

As query language we use *first-order logic* (FO) over the vocabulary $(+, \cdot, 0, 1, <)$ expanded with the relation names in $\mathcal{S}$ [18]. Let $\mathbf{R}$ be the model-theoretic structure

$\langle \mathbb{R}, +, \cdot, 0, 1, < \rangle$. A formula $\varphi(x_1, \ldots, x_k)$ expresses the $k$-ary query defined by

$$\varphi(D) := \{(a_1, \ldots, a_k) \in \mathbb{R}^k \mid \langle \mathbf{R}, D\rangle \models \varphi(a_1, \ldots, a_k)\},$$

for any database $D$, where $\models$ denotes the model-theoretic satisfaction relation. Note that $\varphi(D)$ is always semi-algebraic because all relations in $D$ are; indeed, by Tarski's theorem [24], the relations that are first-order definable on the real ordered field are precisely the semi-algebraic sets.

**Example 1.** The interior query is expressible in FO: Let $\mathcal{S}$ be a schema containing the relation name $S$. Consider the FO formula

$$\varphi_{\text{int}}(\vec{x}) := (\exists \varepsilon > 0)(\forall x_1') \cdots (\forall x_n')(\|\vec{x} - \vec{x}'\| < \varepsilon \rightarrow S(x_1', \ldots, x_n')).$$

For any database $D$, $\varphi_{\text{int}}(D)$ equals the interior of $S^D$.

However, not every query is first-order expressible: the query which asks whether a set is connected is not expressible in FO. This result and other results related to constraint databases have recently been collected in a single volume [18].

### 2.2. *Cones*

Let $A \subseteq \mathbb{R}^n$ be a semi-algebraic set and let $\vec{p} \in \mathbb{R}^n$ be a point not in $A$. We define the *cone with base $A$ and top $\vec{p}$* as the union of all closed line segments between $\vec{p}$ and the points in $A$. We denote this set by $\text{Cone}(A, \vec{p}) := \{t\vec{b} + (1-t)\vec{p} \mid \vec{b} \in A,\ 0 \leq t \leq 1\}$. For a point $\vec{p} \in \mathbb{R}^n$ and $\varepsilon > 0$, denote the closed ball centered at $\vec{p}$ with radius $\varepsilon$ by $B^n(\vec{p}, \varepsilon)$, and denote the sphere centered at $\vec{p}$ with radius $\varepsilon$, by $S^{n-1}(\vec{p}, \varepsilon)$.

We use the following notation: Let $\mathbb{R}^n$ be equipped with the standard Euclidean topology. Let $X \subseteq Y \subseteq \mathbb{R}^n$, the closure of $X$ with respect to the induced topology on $Y$ is denoted by $\text{cl}_Y(X)$, and $\text{int}_Y(X)$ indicates the interior of $X$ with respect to the induced topology on $Y$. When the ambient space $Y$ is $\mathbb{R}^n$, we omit the subscript $Y$. We denote $\text{cl}(X)\backslash\text{int}(X)$, the boundary of $X$, by $\text{bd}(X)$.

Let $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$. A function $h$: $X \rightarrow Y$ is called a *homeomorphism* if it is a bijection and both $h$ and $h^{-1}$ are continuous with respect to induced topology on $X$ and $Y$.

The following well-known theorem says that, locally around each point of $A$, a semi-algebraic set $A$ has the topology of a cone.

**Theorem** [1], [4]. *Let $A \subseteq \mathbb{R}^n$ be a semi-algebraic set and let $\vec{p}$ be a point of $A$. Then there is a real number $\varepsilon > 0$ such that there exists a homeomorphism*

$$h\colon A \cap B^n(\vec{p}, \varepsilon) \rightarrow \text{Cone}(A \cap S^{n-1}(\vec{p}, \varepsilon), \vec{p}).$$

Any real number $\varepsilon > 0$ as in the lemma is called a *cone radius* of $A$ in $\vec{p}$.

In this paper we prove the following theorem:

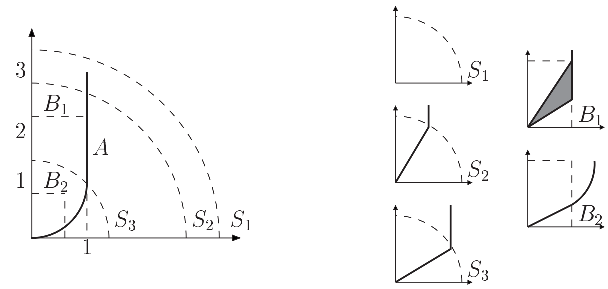


**Fig. 1.** The semi-algebraic set $A$ of Example 2 (left) and the cones corresponding to $S_1$, $S_2$, $S_3$, $B_1$, and $B_2$ (right).

**Theorem 1.** *Let $A \subseteq \mathbb{R}^n$ be a semi-algebraic set and let $\vec{p}$ be a point of $A$. Then there is a real number $\varepsilon > 0$ such that for any box* $\mathrm{Box} = [p_1 - b_1, p_1 + a_1] \times \cdots \times [p_n - b_n, p_n + a_n] \subseteq B^n(\vec{p}, \varepsilon)$, *there exists a homeomorphism*

$$h\colon\ A \cap \mathrm{Box} \to \mathrm{Cone}(A \cap \mathrm{bd}(\mathrm{Box}), \vec{p}).$$

Any real number $\varepsilon > 0$ as in the theorem is called a *box cone radius* of $A$ in $\vec{p}$.

In Section 3 we show that every box cone radius is a cone radius. However, the converse does not necessary hold as the following example shows:

**Example 2.** Consider $A = \{(x, y) \in \mathbb{R}^2 \mid (0 \leq x \leq 1 \wedge 0 \leq y \leq 1 \wedge x^2 + (y-1)^2 = 1) \vee (x = 1 \wedge 1 \leq y \leq 3)\}$ and let $\vec{p} = (0, 0)$. On the left of Fig. 1 the set $A$ together with three circles $S_1$, $S_2$, and $S_3$ and two rectangles $B_1$ and $B_2$ are shown. On the right of Fig. 1 the corresponding cones can be seen. It is clear that the radius of $S_1$ is not a cone radius of $A$ in $\vec{p}$ (the cone is the empty set) and that the radius of both $S_2$ and $S_3$ is a cone radius of $A$ in $\vec{p}$. However, looking at the cone corresponding to $B_1$, the radius of $S_2$ is not a box cone radius of $A$ in $\vec{p}$ (the cone is a two-dimensional set). The radius of $S_3$ is a box radius of $A$ in $\vec{p}$ as can be seen from the cone corresponding to $B_2$.

Let $\mathcal{S}$ be a schema containing a relation name $S$ of arity $n$. A (*box*) *cone radius query* $Q_{\mathrm{radius}}$ is a query which maps any database $D$ over $\mathcal{S}$ to a set of pairs $(\vec{p}, r) \in \mathbb{R}^n \times \mathbb{R}$ such that for every point $\vec{p} \in S^D$ there exists at least one pair $(\vec{p}, r) \in Q_{\mathrm{radius}}(D)$, and for every $(\vec{p}, r) \in Q_{\mathrm{radius}}(D)$, $r$ is a (box) cone radius in $\vec{p}$ for $S^D$.

Our second result is the following:

**Theorem 2.** *There exists an FO-expressible box cone radius query.*

### 2.3. *$C^1$-Whitney Decomposition*

In this section we construct a $C^1$-Whitney decomposition of a semi-algebraic set $A$ and show that this construction is expressible in FO. The construction consists of several steps. Firstly, $A$ is decomposed in parts which are $C^1$-smooth, resulting in the $C^1$-decomposition

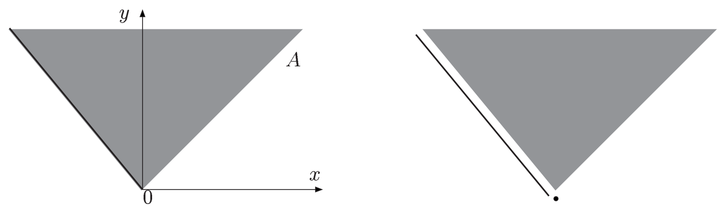


**Fig. 2.** The semi-algebraic set $A$ of Example 3 (left) and its $C^1$-decomposition of $A$ (right).

of $A$. Secondly, the decomposition is refined such that all points in a single part have the same local topological type. This gives the $C^1$-Whitney decomposition of $A$. Finally, this decomposition is made compatible with a finite number of specified sets.

Before giving the formal definition of a $C^1$-decomposition, we give an example.

**Example 3.** Consider the semi-algebraic set $A = \{(x, y) \in \mathbb{R}^2 \mid (y > |x|) \vee (y = -x \wedge y \geq 0)\}$ which is depicted in Fig. 2. We first look at those points where $A$ has no tangent space. It is clear that $A$ has no tangent space in the origin. However, if we decompose $A$ into $\{(x, y) \in \mathbb{R}^2 \mid (y > |x|) \vee (y = -x \wedge y > 0)\}$ and $\{(0, 0)\}$, then both parts have a tangent space in any of its points. We then decompose these parts according to their dimension: let $A_2$ be the two-dimensional part $\{(x, y) \in \mathbb{R}^2 \mid y > |x|\}$, let $A_1$ be the one-dimensional part $\{(x, y) \in \mathbb{R}^2 \mid (y = -x \wedge y > 0)\}$, and finally let $A_0$ be the zero-dimensional part $\{(0, 0)\}$. The sets $A_0, A_1, A_2$ will form the $C^1$-decomposition of $A$.

The following definitions are taken from [20]. Let $A$ be a semi-algebraic set in $\mathbb{R}^n$. The *secants limit set* of $A$ in a point $\vec{p} \in A$ is defined as the set

$$\operatorname{limsec}_{\vec{p}} A := \bigcap_{\eta > 0} \operatorname{cl}(\{\lambda(\vec{u} - \vec{v}) \in \mathbb{R}^n \mid \lambda \in \mathbb{R} \text{ and } \vec{u}, \vec{v} \in A \cap B^n(\vec{p}, \eta)\}).$$

If $\operatorname{limsec}_{\vec{p}} A$ is a vector space (this is true when for all $\vec{s}, \vec{t} \in \operatorname{limsec}_{\vec{p}} A$, the sum $\vec{s} + \vec{t}$ is also an element of $\operatorname{limsec}_{\vec{p}} A$), then we define the *tangent space of $A$ in $\vec{p}$* as $\mathrm{T}_{\vec{p}} A := \vec{p} + \operatorname{limsec}_{\vec{p}} A$. If $\operatorname{limsec}_{\vec{p}} A$ is not a vector space, the tangent space of $A$ in $\vec{p}$ is undefined. The set $A$ is *$C^1$-smooth in $\vec{p}$* if and only if $\mathrm{T}_{\vec{p}} A$ exists and there exist a neighborhood $U$ of $\vec{p}$ such that the orthogonal projection of $A \cap U$ on $\mathrm{T}_{\vec{p}} A$ is bijective. A set is *$C^1$-smooth* if it is *$C^1$-smooth* in all its points. We define the set

$$\mathrm{Smooth}_k(A) = \{\vec{x} \in \mathbb{R}^n \mid A \text{ is } C^1\text{-smooth in } \vec{x} \text{ and of dimension } k\}.$$

We can now decompose [26], [22] $A$ into at most $n + 1$ nonempty $C^1$-smooth parts $A_0, \ldots, A_n$ as follows: Define $A_n = \mathrm{Smooth}_n(A)$. Suppose that $A_n, \ldots, A_{k+1}$ is already constructed. Then define

$$A_k = \mathrm{Smooth}_k\left(A \setminus \bigcup_{i=k+1}^{n} A_i\right). \tag{1}$$

The sets $A_0, \ldots, A_n$ are called the *$C^1$-decomposition of $A$*.

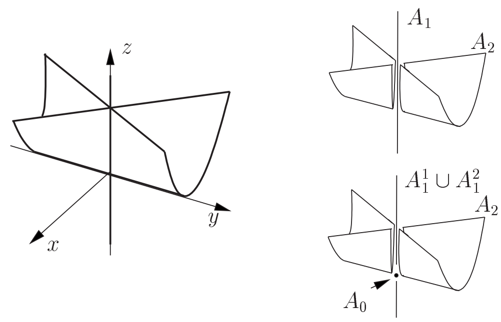


**Fig. 3.** The Whitney umbrella (left), its $C^1$-decomposition (top right), and its $C^1$-Whitney decomposition (bottom right), as described in Example 4.

Let $\mathcal{S}$ be a database schema containing a relation name $S$ of arity $n$. For each $k \geq 0$, define the query $Q_{k\text{-smooth}}$ as $Q_{k\text{-smooth}}(D) := \text{Smooth}_k(S^D)$ for any database $D$ over $\mathcal{S}$. By the constructions given by Rannou [20], the following is readily verified:

**Proposition 1.** *Let $\mathcal{S}$ be a database schema containing a relation name $S$ of arity $n$. For each $0 \leq k \leq n$ the query $Q_{k\text{-smooth}}$ is expressible in FO.*

The following example motivates the construction of the $C^1$-Whitney decomposition.

**Example 4.** Let $A = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 - zy^2 = 0\}$. This set is known as the *Whitney umbrella* and is depicted in Fig. 3. The $C^1$-decomposition of $A$ consists of two nonempty sets $A_2 = \text{Smooth}_2(A) = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 - zy^2 = 0 \wedge \neg(x = 0 \wedge y = 0)\}$ and $A_1 = \text{Smooth}_1(A) = \{(x, y, z) \in \mathbb{R}^3 \mid x = 0 \wedge y = 0\}$. However, the local topological type changes when one looks at points in $A_1$ with $z < 0$, $z = 0$, and $z > 0$. On the contrary, any two points in $A_2$ have the same local topological type. For this reason, one agrees to split up the set $A_1$ into $A_1^1 = \{(x, y, z) \in \mathbb{R}^3 \mid x = 0 \wedge y = 0 \wedge z < 0\}$, $A_1^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x = 0 \wedge y = 0 \wedge z > 0\}$, and $A_0 = \{(x, y, z) \in \mathbb{R}^3 \mid x = 0 \wedge y = 0 \wedge z = 0\}$. Then the sets $A_0$, $A_1^1 \cup A_1^2$, and $A_2$ form a decomposition of $A$ which is called the $C^1$-Whitney decomposition of $A$.

To avoid such situations as that in Example 4, Whitney [27] introduced the following condition for $C^1$-smooth semi-algebraic sets $X, Y \subseteq \mathbb{R}^n$ and a point $\vec{x} \in X$.

One says that the triple $(X, \vec{x}, Y)$ has the *Whitney property* when the following holds: if $(\vec{x}_i)$, $(\vec{y}_i)$ are sequences in $X$, $Y$, respectively, both converging to $\vec{x}$, if the sequence of tangent spaces $(\mathrm{T}_{\vec{y}_i} Y)$ converges to a subspace $T \subseteq \mathbb{R}^n$, and if the sequence $(\overrightarrow{x_i\, y_i})$ of lines containing $\vec{x}_i - \vec{y}_i$ converges to a line $L \subseteq \mathbb{R}^n$, then $L \subseteq T$. One says that $(X, Y)$ has the Whitney property if $(X, \vec{x}, Y)$ has the Whitney property for any point $\vec{x} \in X$. We define the set $\text{Whitney}(X, Y) = \{\vec{x} \in \mathbb{R}^n \mid X, Y$ are $C^1$-smooth, $\vec{x} \in X$ and $(X, \vec{x}, Y)$ has the Whitney property$\}$.

Let $\mathcal{S}$ be a database schema containing two relation names $S_1$ and $S_2$ of arity $n$. Define the $n$-ary query, defined as $Q_{\text{Whitney}}(D) := \text{Whitney}(S_1^D, S_2^D)$ for any database $D$ over $\mathcal{S}$.

Again, by constructions given by Rannou [20] the following is readily verified:

**Proposition 2.** *Let $\mathcal{S}$ be a database schema containing two relation names $S_1$ and $S_2$ of arity $n$. The $n$-ary query $Q_{\text{Whitney}}$ is expressible in FO.*

We can decompose $A$ into $n+1$ $C^1$-smooth sets $A_0, \ldots, A_n$ such that $(A_i, A_j)$ has the Whitney property for every $i < j$. Indeed, let $A_n := \text{Smooth}_n(A)$. Now suppose $A_n, \ldots, A_{k+1}$ have already been constructed. Part $A_k$ is then constructed as follows:

$$R_k := \text{Smooth}_k\left(A \setminus \bigcup_{i=k+1}^{n} A_i\right), \tag{2}$$

$$A_k := \bigcap_{i=k+1}^{n} \text{int}_{R_k}(\text{Whitney}(R_k, R_i)). \tag{3}$$

In (3) the interior is taken relative to the set $R_k$. As a result, the set $A_k$ is still $C^1$-smooth because it is an open subset of the $C^1$-smooth set $R_k$. The sets $A_0, \ldots, A_n$ are called the *$C^1$-Whitney decomposition of $A$*.

A $C^1$-Whitney decomposition $A_0, \ldots, A_n$ of $A$ is called *compatible* with a finite set of semi-algebraic sets $\{B_1, \ldots, B_k\}$ if for any of the $B_i$'s, each connected component of the $A_i$'s is either included in or disjoint with $B_i$.

For reasons that will become clear in Section 3, we now construct a $C^1$-Whitney decomposition of $\text{cl}(A)$ which is compatible with $A$ and $\{\vec{x} \in \mathbb{R}^n \mid x_1 a_1 p_1, \ldots, x_n a_n p_n\}$ for each tuple $(a_1, \ldots, a_n) \in \{=, \neq\}^n$.

**Example 5.** Consider again the semi-algebraic set $A = \{(x, y) \in \mathbb{R}^2 \mid (y > |x|) \vee (y = -x \wedge y \geq 0)\}$. Let $A_2 = \{(x, y) \in \mathbb{R}^2 \mid y > |x| > 0\}$, $A_1 = \{(x, y) \in \mathbb{R}^2 \mid y = -x \wedge y > 0\} \cup \{(x, y) \in \mathbb{R}^2 \mid y = x \wedge y > 0\} \cup \{(x, y) \in \mathbb{R}^2 \mid x = 0 \wedge y > 0\}$, and $A_0 = \{(0, 0)\}$. Then $A_0, A_1, A_2$ is a $C^1$-Whitney decomposition of $\text{cl}(A)$ compatible with $A$, $\{(0, 0)\}$, $\{(x, y) \in \mathbb{R}^2 \mid x = 0\}$, and $\{(x, y) \in \mathbb{R}^2 \mid y = 0\}$. This example is illustrated in Fig. 4.

We now decompose $\text{cl}(A)$ into $n+1$ $C^1$-smooth parts $A_0, \ldots, A_n$ such that $(A_i, A_j)$ has the Whitney property for every $i < j$, and such that this decomposition is compatible with $A$ and $\{\vec{x} \in \mathbb{R}^n \mid x_1 a_1 p_1, \ldots, x_n a_n p_n\}$ for each tuple $(a_1, \ldots, a_n) \in \{=, \neq\}^n$.

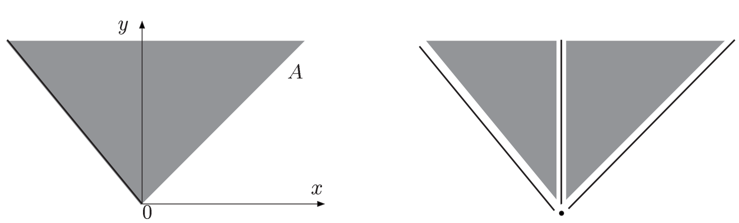


**Fig. 4.** Set $A$ of Example 5 (left) and its $C^1$-Whitney decomposition of $\text{cl}(A)$ compatible with $A$, $\{(0, 0)\}$, $\{(x, y) \in \mathbb{R}^2 \mid x = 0\}$, and $\{(x, y) \in \mathbb{R}^2 \mid y = 0\}$ (right).

The following construction is an adaptation of the construction given by Shiota [22, Lemma I.2.2].

We define $A_n = \mathrm{Smooth}_n(A)$. Now suppose that the parts $A_n, \ldots, A_{k+1}$ have already been constructed. Then part $A_k$ is constructed as follows. Define $B_0 = A$, $B_1 = \mathrm{cl}(A)\backslash A$ and for $i = 0, 1$ define

$$B_i^{\sigma_1\cdots\sigma_n} := \{\vec{x} \in B_i \mid x_1\sigma_1 p_1, \ldots, x_n\sigma_n p_n\},$$

with $\sigma_1, \ldots, \sigma_n \in \{<, =, >\}$. For each tuple $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n) \in \{<, =, >\}^n$ and $i = 0, 1$ construct

$$R_{i,k}^{\boldsymbol{\sigma}} := \mathrm{Smooth}_k\left(B_i^{\boldsymbol{\sigma}}\backslash \bigcup_{j=k+1}^{n} A_j\right), \tag{4}$$

$$W_{i,k}^{\boldsymbol{\sigma}} := \bigcap_{j=k+1}^{n} \mathrm{int}_{R_{i,k}^{\boldsymbol{\sigma}}}(\mathrm{Whitney}(R_{i,k}^{\boldsymbol{\sigma}}, A_j)), \tag{5}$$

$$A_{i,k}^{\boldsymbol{\sigma}} := W_{i,k}^{\boldsymbol{\sigma}}\backslash \mathrm{cl}\left(W_{1-i,k}^{\boldsymbol{\sigma}} \cup \bigcup_{\substack{\boldsymbol{\sigma}'\in\{<,=,>\}^n\\ \boldsymbol{\sigma}'\neq\boldsymbol{\sigma}}} (W_{0,k}^{\boldsymbol{\sigma}'} \cup W_{1,k}^{\boldsymbol{\sigma}'})\right). \tag{6}$$

Then we define $A_k^{\boldsymbol{\sigma}} = A_{0,k}^{\boldsymbol{\sigma}} \cup A_{1,k}^{\boldsymbol{\sigma}}$ and $A_k := \bigcup_{\boldsymbol{\sigma}\in\{<,=,>\}^n} A_k^{\boldsymbol{\sigma}}$. Set $A_k$ indeed has the desired properties: by (4) it is $C^1$-smooth and of dimension $k$, (5) guarantees that for all points in $A_k$, and for any $j > k$, $(A_k, A_j)$ has the Whitney property, and (6) ensures that the connected components are either included in or disjoint with $A$ and $\{\vec{x} \in \mathbb{R}^n \mid x_1 a_1 p_1, \ldots, x_n a_n p_n\}$ for each tuple $(a_1, \ldots, a_n) \in \{=, \neq\}^n$.

It is well known [26], [22] that the dimension of $\mathrm{cl}(A)\backslash \bigcup_{j=k}^n A_j$ is strictly smaller than the dimension of $\mathrm{cl}(A)\backslash \bigcup_{j=k+1}^n A_j$ for $i = 0, 1$. Hence, the decomposition will consists of exactly $n + 1$ sets $A_k$, some of which may be empty.

Let $\mathcal{S}$ be a database schema containing a relation name $S$ of arity $n$. We define the $n$-ary query $Q_{k\text{-part}}(D) = (S^D)_k$, with $(S^D)_k$ the $k$th part of the decomposition constructed above for $A = S^D$.

A direct consequence of Propositions 1 and 2 is the following:

**Proposition 3.** *Let $\mathcal{S}$ be a database schema containing a relation name $S$ of arity $n$. For each $0 \leq k \leq n$, the $n$-ary query $Q_{k\text{-part}}$ is expressible in FO.*

## 3. Expressing the Box Cone Radius in FO

Before proving Theorem 1, we look again at Example 2. In this example it is clear what the local topology of $A$ in $\vec{p}$ is. Indeed, $A$ looks like a straight line locally around $\vec{p}$. Hence, a radius $\varepsilon$ will be a box cone radius if the boundary of any box around $\vec{p}$ in $B^n(\vec{p}, \varepsilon)$ intersects $A$ in a single point. For this reason the radius of $S_2$ is not a box cone radius (there is even a box whose boundary has a one-dimensional intersection!). Similarly, if $\varepsilon$ is large and you take a very large box, the intersection of the boundary of this box with $A$ will be empty. Also in this case, $\varepsilon$ will not be a box cone radius.

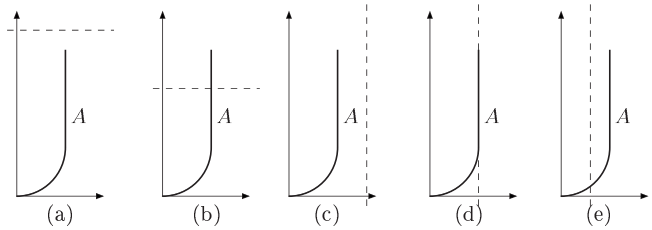


**Fig. 5.** Possible intersections of vertical and horizontal lines with $A$.

Since the boundary of a box in $\mathbb{R}^2$ consists of horizontal and vertical line segments, it is sufficient to look at the intersection of those segments with $A$. The possible intersections are shown in Fig. 5. Only cases (b) and (e) deliver a single intersection point. In the proof of Theorem 1 we will identify the points in which the intersection behavior changes for the vertical or horizontal line segments. In Example 2 the vertical intersections change from the empty set to one point, when crossing the point $(1, 3)$. Similarly, the horizontal intersections change from the empty set to a single point when crossing the horizontal line at $x = 1$. The points in which these changes occur will be the so-called critical points. These points correspond to zero-dimensional parts of the $C^1$-Whitney decomposition (like the point $(1, 3)$), or those points where there is a horizontal or vertical tangent space (like the points in $\{(x, y) \in \mathbb{R}^2 \mid x = 1 \wedge (1 \leq y < 3)\}$). A radius $\varepsilon$ will then be a box cone radius when $B^2(\vec{p}, \varepsilon)$ does not contain any critical points. The critical point closest to $\vec{p}$ is the point $(1, 1)$, hence any radius smaller than 1 will be a box cone radius.

We will formalize the above intuitive example and prove the first result of this paper.

**Theorem 1.** *Let $A \subseteq \mathbb{R}^n$ be a semi-algebraic set and let $\vec{p}$ be a point of $A$. Then there is a real number $\varepsilon > 0$ such that for any box* $\mathrm{Box} = [p_1 - b_1, p_1 + a_1] \times \cdots \times [p_n - b_n, p_n + a_n] \subseteq B^n(\vec{p}, \varepsilon)$, *there exists a homeomorphism*

$$h\colon\ A \cap \mathrm{Box} \to \mathrm{Cone}(A \cap \mathrm{bd}(\mathrm{Box}), \vec{p}).$$

*Proof.* Let $A \subseteq \mathbb{R}^n$, $\vec{p} \in A$, and let $A_0, \ldots, A_n$ be the $C^1$-Whitney decomposition of $\mathrm{cl}(A)$ compatible with $A$ and $\{\vec{x} \in \mathbb{R}^n \mid x_1 a_1 p_1, \ldots, x_n a_n p_n\}$ for each tuple $(a_1, \ldots, a_n) \in \{=, \neq\}^n$. Let $\vec{x} \in \mathbb{R}^n$ and define $f_1(\vec{x}) = (p_1 - x_1)^2, \ldots, f_n(\vec{x}) = (p_n - x_n)^2$ and $f_{n+1}(\vec{x}) = (p_1 - x_1)^2 + \cdots + (p_n - x_n)^2$.

For each $k = 0, \ldots, n$, any connected component of $A_k$ is the disjoint union of connected components of $A_k^{\boldsymbol{\sigma}}$ with $\boldsymbol{\sigma} \in \{<, =, >\}^n$. Since the $C^1$-Whitney decomposition $A_0, \ldots, A_n$ is compatible with $\{\vec{x} \in \mathbb{R}^n \mid x_1 a_1 p_1, \ldots, x_n a_n p_n\}$ for each tuple $(a_1, \ldots, a_n) \in \{=, \neq\}^n$, any connected component of $A_k^{\boldsymbol{\sigma}}$ is either included in or disjoint with $f_j^{-1}(0)$ for any $j = 1, \ldots, n$.

For each $k = 1, \ldots, n$ and any $\boldsymbol{\sigma} \in \{<, =, >\}^n$, we define $J_k^{\boldsymbol{\sigma}} \subseteq \{1, \ldots, n+1\}$ as those indices $j$ such that $A_k^{\boldsymbol{\sigma}} \cap f_j^{-1}(0) = \emptyset$.

The theorem directly follows from the following two claims:

**Claim 1.** *If $\varepsilon$ is a positive real number such that*

$$(A_0 \backslash \{\vec{p}\}) \cap \mathrm{int}(B^n(\vec{p}, \varepsilon)) = \emptyset, \tag{7}$$

*and, for each $k = 1, \ldots, n$ and any $\boldsymbol{\sigma} \in \{<, =, >\}^n$ and $j \in J_k^{\boldsymbol{\sigma}}$, the restriction*

$$f_j \mid (A_k^{\boldsymbol{\sigma}} \cap \operatorname{int}(B^n(\vec{p}, \varepsilon))) \longrightarrow \mathbb{R} \text{ has no critical points,} \tag{8}$$

*then $\varepsilon$ is a box cone radius in $\vec{p}$ for $A$.*

Here, the *critical points* of $f_j \mid X$ for some subset $X \subseteq \mathbb{R}^n$, are the points $\vec{x} \in X$ where the differential mapping $d_{\vec{x}}(f_j \mid X)$: $\mathrm{T}_{\vec{x}}X \to \mathbb{R}$, defined by $d_{\vec{x}}(f_j \mid X)(v_1, \ldots, v_n) = ((\partial(f_j \mid X)/\partial x_1)(\vec{x}), \ldots, (\partial(f_j \mid X)/\partial x_n)(\vec{x})) \cdot (v_1, \ldots, v_n)$, is not surjective.

**Claim 2.** *There exists a positive real number $\varepsilon$ such that the conditions of the above claim hold.*

*Proof of Claim* 1. We equip each part $A_k$ with the standard Riemannian metric $\langle , \rangle$ induced from $\mathbb{R}^k$. Then for any $C^1$-function $f$ on $A_k$, we can define the *gradient vector field* $\operatorname{grad}(f \mid A_k)$. The value of $\operatorname{grad}(f \mid A_k)$ in a point $\vec{x}$ is the unique vector in $T_{\vec{x}}A_k$ with the property

$$\langle \operatorname{grad}(f \mid A_k)(\vec{x}), \vec{v} \rangle = \mathrm{d}_{\vec{x}}(f \mid A_k)(\vec{v}).$$

Next, for each $\boldsymbol{\sigma} \in \{<, =, >\}^n$, we define the following continuous vector field on $A_k^{\boldsymbol{\sigma}}$:

$$\xi_k^{\boldsymbol{\sigma}} = -\operatorname{grad}(f_{n+1} \mid A_k^{\boldsymbol{\sigma}}) \prod_{j \in J_k^{\boldsymbol{\sigma}} \setminus \{n+1\}} |\operatorname{grad}(f_j | A_k^{\boldsymbol{\sigma}})|.$$

Clearly, $\xi_k^{\boldsymbol{\sigma}}(\vec{x}) = 0$ if and only if $\vec{x} \in A_k^{\boldsymbol{\sigma}}$ is a critical point of at least one of the $f_j \mid A_k^{\boldsymbol{\sigma}}$'s with $j \in J_k^{\boldsymbol{\sigma}}$. In Fig. 6 we have depicted the vector field $\xi$ in case where $A = \mathbb{R}^2$ and $\vec{p} = 0$.

By conditions (7) and (8), $\varepsilon > 0$ is such that within $\operatorname{int}(B^n(\vec{p}, \varepsilon))$ no critical points occur. Hence, $\xi_k^{\boldsymbol{\sigma}}(\vec{x})(f_j) < 0$ for $\vec{x} \in A_k^{\boldsymbol{\sigma}} \cap \operatorname{int}(B^n(\vec{p}, \varepsilon))$ and each $j \in J_k^{\boldsymbol{\sigma}}$. (Here, we take the alternative view of tangent vectors where the vector $\xi_k^{\boldsymbol{\sigma}}(\vec{x})$ is seen as a function which maps real-valued functions on $A_k^{\boldsymbol{\sigma}}$ to a real number [5, Chapter 8].)

Since parts $A_1, \ldots, A_n$ are $C^1$-manifolds, we can clearly obtain a continuous flow on each part $A_i^{\boldsymbol{\sigma}}$ by integrating the continuous vector field $\xi_k^{\boldsymbol{\sigma}}$. In general, however, we cannot expect to obtain a continuous flow on the set $\operatorname{cl}(A)$ by just putting the flows

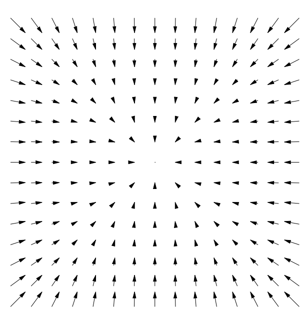

**Fig. 6.** The vector field $\xi$ for $A = \mathbb{R}^2$ and $\vec{p} = 0$.

together. Therefore the vector fields $\xi_k^{\boldsymbol{\sigma}}$ will be transformed into vector fields $\eta_k^{\boldsymbol{\sigma}}$ which are controlled along the boundary of each part. Integrating these new vector fields and putting the flows together then results in a continuous flow on $\mathrm{cl}(A)$ [10].

Technically, the controlled vectors field $\eta_k^{\boldsymbol{\sigma}}$ are constructed inductively on the dimension of the parts by means of a $C^1$-controlled tube system for the $C^1$-Whitney decomposition $A_1, \ldots, A_n$ [23], [22, Lemmas I.1.3 and I.1.5]. Moreover, these vector fields can be chosen such that they also satisfy $\eta_k^{\boldsymbol{\sigma}}(\vec{x})(f_j) < 0$ for $\vec{x} \in A_k^{\boldsymbol{\sigma}} \cap \mathrm{int}(B^n(\vec{p}, \varepsilon))$ and $j \in J_k^{\boldsymbol{\sigma}}$. Since $\mathrm{cl}(A) \cap \mathrm{int}(B^n(\vec{p}, \varepsilon))$ is locally closed, the vector fields $\eta_k^{\boldsymbol{\sigma}}$ admit a continuous flow $\theta_i^{\boldsymbol{\sigma}}$ which nicely fit together into a continuous flow $\Theta$ on $\mathrm{cl}(A) \cap \mathrm{int}\,(B^n(\vec{p}, \varepsilon))$ [22, Lemma I.1.6].

Let $\mathrm{Box} = [p_1 - b_1, p_1 + a_1] \times \cdots \times [p_n - b_n, p_n + a_n] \subseteq \mathrm{int}(B^n(\vec{p}, \varepsilon))$. We will show that the flow $\Theta$ induces the required homeomorphism between $A \cap \mathrm{Box}$ and $\mathrm{Cone}(A \cap \mathrm{bd}(\mathrm{Box}), \vec{p})$, proving that $\varepsilon$ is a box cone radius in $\vec{p}$ for $A$.

The flow $\Theta$ is the union of flows $\theta_k^{\boldsymbol{\sigma}}$ for $k = 1, \ldots, n$ and $\boldsymbol{\sigma} \in \{<, =, >\}^n$. Each individual flow $\theta_k^{\boldsymbol{\sigma}}$ can be chosen to be the continuous map

$$\mathbb{R} \times (A_k^{\boldsymbol{\sigma}} \cap \mathrm{int}(B^n(\vec{p}, \varepsilon))) \to A_k^{\boldsymbol{\sigma}} \cap \mathrm{int}(B^n(\vec{p}, \varepsilon)),$$

which is the unique solution of

$$\frac{d}{dt}\theta_k^{\boldsymbol{\sigma}}(t, \vec{x}) = \eta_k^{\boldsymbol{\sigma}}(\vec{x})$$

for $\vec{x} \in A_k^{\boldsymbol{\sigma}} \cap \mathrm{int}(B^n(\vec{p}, \varepsilon))$ and initial condition $\theta_k^{\boldsymbol{\sigma}}(0, \vec{x}) = \vec{x}$ for $\vec{x} \in A_k^{\boldsymbol{\sigma}} \cap \mathrm{bd}(\mathrm{Box})$.

Since $\eta_k^{\boldsymbol{\sigma}}(\vec{x})(f_j) < 0$ for $j \in J_k^{\boldsymbol{\sigma}}$, we have that

$$|p_j - (\theta_k^{\boldsymbol{\sigma}}(t_1, \vec{x}))_j| > |p_j - (\theta_k^{\boldsymbol{\sigma}}(t_2, \vec{x}))_j| \qquad \text{for any} \quad t_1 < t_2, \tag{9}$$

for all $j \in J_k^{\boldsymbol{\sigma}}$. This means that the coordinates of the *integral curves* $\gamma_{\vec{x}}\colon\, t \to \theta_k^{\boldsymbol{\sigma}}(t, \vec{x})$ are either monotone decreasing for increasing values of $t$, or equal to $p_j$ for some $j$.

For each $t \in (0, 1]$, define the box $\mathrm{Box}_t = [p_1 - tb_1, p_1 + ta_1] \times \cdots \times [p_n - tb_n, p_n + ta_n]$ and the mapping (illustrated in Fig. 7)

$$\begin{aligned} (h_k^{\boldsymbol{\sigma}})_t\colon\, A_k^{\boldsymbol{\sigma}} \cap \mathrm{bd}(\mathrm{Box}) &\to (A_k^{\boldsymbol{\sigma}} \cap \mathrm{bd}(\mathrm{Box}_t)) \backslash \{\vec{p}\}, \\ \vec{x} &\mapsto \gamma_{\vec{x}} \cap \mathrm{bd}(\mathrm{Box}_t). \end{aligned}$$

We now define $h_t\colon\, \mathrm{cl}(A) \cap \mathrm{bd}(\mathrm{Box}) \to \mathrm{cl}(A) \cap \mathrm{bd}(\mathrm{Box}_t)$ by $h_t(\vec{x}) = (h_k^{\boldsymbol{\sigma}})_t(\vec{x})$, where $k$ and $\boldsymbol{\sigma}$ are the unique indices such that $\vec{x} \in A_k^{\boldsymbol{\sigma}}$, and show that it is a homeomorphism

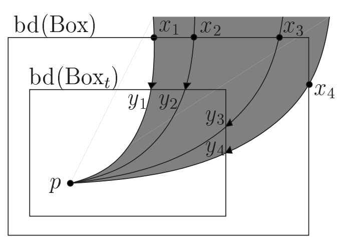


**Fig. 7.** The mapping $h_t$ maps $\vec{x}_i$, $i = 1, 2, 3, 4$, to the intersections of the integral curves with $\mathrm{bd}(\mathrm{Box}_t)$ which are the points $\vec{y}_i$, $i = 1, 2, 3, 4$.

for each $t \in (0, 1]$. It suffices to prove that each $(h_k^{\boldsymbol\sigma})_t$ is a function, since bijectivity and continuity follow directly from the properties of flows of vector fields [10].

Firstly, for each $\vec{x} \in A_k^{\boldsymbol\sigma} \cap \mathrm{bd}(\mathrm{Box})$ there exists a point $\vec{y} \in A_k^{\boldsymbol\sigma} \cap \mathrm{bd}(\mathrm{Box}_t)$ such that $(h_k^{\boldsymbol\sigma})_t(\vec{x}) = \vec{y}$. Indeed, it follows from property (9) that $\lim_{t' \to +\infty} \theta_k^{\boldsymbol\sigma}(t', \vec{x}) = \vec{p}$, and, hence, the curve $\gamma_{\vec{x}}$ must have an intersection with $\mathrm{bd}(\mathrm{Box}_t)$.

Secondly, suppose that for an $\vec{x} \in A_k^{\boldsymbol\sigma} \cap \mathrm{bd}(\mathrm{Box})$, $(h_k^{\boldsymbol\sigma})_t(\vec{x}) = \{\vec{y}_1, \vec{y}_2\}$. By definition of $(h_k^{\boldsymbol\sigma})_t$, this happens when $\theta_k^{\boldsymbol\sigma}(t_1, \vec{x}) = \vec{y}_1$ and $\theta_k^{\boldsymbol\sigma}(t_2, \vec{x}) = \vec{y}_2$, or, in other words, when the integral curve $\gamma_{\vec{x}}$ intersects $\mathrm{bd}(\mathrm{Box}_t)$ twice. Clearly, property (9) implies that $\vec{y}_1 \neq \vec{y}_2$. However, we show that this also is impossible. We give the argument for $n = 2$, the $n$-dimensional case being analogous.

In this case, $\mathrm{Box}_t = [p_1 - tb_1, p_1 + ta_1] \times [p_2 - tb_2, p_2 + ta_2]$. By property (9), there exists a unique $\boldsymbol\sigma \in \{<, =, >\}^2$ such that $\theta_k^{\boldsymbol\sigma}(t, \vec{x}) \subseteq \mathrm{Box}_t^{\boldsymbol\sigma}$.[1] Suppose that $\theta_k^{\boldsymbol\sigma}(t, \vec{x}) \subseteq B_t^{(>,>)}$, the other cases being analogous. Hence, either

1. $\vec{y}_1 = (p_1 + ta_1, p_2 + ta_2)$; or
2. $\vec{y}_1 = (p_1 + ta_1, u)$ with $p_2 < u \leq p_2 + ta_2$; or
3. $\vec{y}_1 = (v, p_2 + ta_2)$ with $p_1 < v \leq p_1 + ta_1$.

Suppose that we are in the second case. Then $\vec{y}_2 = (u', v')$ with $p_1 < u' < p_1 + ta_1$ and $p_2 < v' < u$. Since $\vec{y}_2 \in \mathrm{Box}_t$, $v'$ must be equal to $p_2 + ta_2$, but this is impossible since we have that $p_2 < v' \leq u \leq p_2 + ta_2$. Hence, $(h_k^{\boldsymbol\sigma})_t$ is a function.

Next, we define the homeomorphism

$$h_1 := (0, 1] \times (\mathrm{cl}(A) \cap \mathrm{bd}(\mathrm{Box})) \to \mathrm{cl}(A) \cap \mathrm{Box} \backslash \{\vec{p}\},$$

by $h_1(t, \vec{x}) := h_t(\vec{x})$.

Since $A_0, \ldots, A_n$ is compatible with $A$ and the flows of which $h_1$ is constructed preserve the connected components of the $A_k$'s, the restriction $h_1^{-1} \mid ((A \cap \mathrm{Box}) \backslash \{\vec{p}\})$ will also be a homeomorphism between $(A \cap \mathrm{Box}) \backslash \{\vec{p}\}$ and $(0, 1] \times (A \cap \mathrm{bd}(\mathrm{Box}))$. Since the cylinder $(0, 1] \times (A \cap \mathrm{bd}(\mathrm{Box}))$ is homeomorphic to $\mathrm{Cone}(A \cap \mathrm{bd}(\mathrm{Box}), \vec{p}) \backslash \{\vec{p}\}$, e.g., by the homeomorphism

$$h_2(t, \vec{x}) := (1 - t)\vec{p} + t\vec{x},$$

we obtain a homeomorphism

$$h_3 := h_2 \circ h_1^{-1} \colon\ (A \cap \mathrm{Box}) \backslash \{\vec{p}\} \to \mathrm{Cone}(A \cap \mathrm{bd}(\mathrm{Box}), \vec{p}) \backslash \{\vec{p}\}.$$

The homeomorphism $h_3$ can be trivially extended to the point $\vec{p}$, resulting in the desired homeomorphism $h$. □

*Proof of Claim* 2. We now observe that for each $\vec{p}$ there exists a positive real number $\varepsilon > 0$ such that both (7) and (8) are satisfied.

Indeed, it is clear that the distance between $\vec{p}$ and $A_0 \backslash \{\vec{p}\}$ is strictly positive. Let $\varepsilon_0$ be this distance.

[1] Let $X \subseteq \mathbb{R}^n$ and $\sigma \in \{<, =, >\}^n$. Then $X^{\sigma} = \{\vec{x} \in X \mid x\ \sigma_1\ 0, \ldots, x\ \sigma_n\ 0\}$.

Next, a *critical value* of $f_j \mid A_k^{\sigma}$ is the image by $f_j \mid A_k^{\sigma}$ of a critical point. The set of critical points of $f_j \mid A_k^{\sigma}$ is semi-algebraic and admits a $C^1$-cell decomposition $\mathcal{C} = \{C_1, \dots, C_m\}$ such that $f_j \mid C_\ell$ is $C^1$ [25]. Sard's theorem for $C^1$-mapping [28] implies that each $f_j \mid C_\ell$ attains only a finite number of values. Hence the image by $f_j \mid A_k^{\sigma}$ of the set of critical points is finite. This is true for every $k = 1, \dots, n$ and every $\boldsymbol{\sigma} \in \{<, =, >\}^n$ and $j \in J_k^{\sigma}$.

Denote by $\varepsilon_{jk}^{\sigma}$ the minimal critical value of $f_j \mid A_k^{\sigma}$. (If there are no critical values, set $\varepsilon_{jk}^{\sigma} = 1$.) Note that $\varepsilon_{jk}^{\sigma}$ can only be zero for points of $A_k^{\sigma} \cap f_j^{-1}(0)$. However, this implies that $j \notin J_k^{\sigma}$, which is impossible. Hence, any $0 < \varepsilon < \min\{\varepsilon_0, \sqrt{\varepsilon_{jk}^{\sigma}/n} \mid k = 1, \dots, n, \boldsymbol{\sigma} \in \{<, =, >\}^n, j \in J_k^{\sigma}\}$ is a good one. □

This concludes the proof of Theorem 1. □

**Theorem 2.** *There exists an FO-expressible box cone radius query.*

*Proof.* Let $\mathcal{S}$ be a schema containing a relation name $S$ of arity $n$, and let $D$ be a database over $\mathcal{S}$. By Claim 1, we can define the following cone radius query:

$$Q_{\text{radius}}(D) := \{(\vec{p}, r) \in \mathbb{R}^n \times \mathbb{R} \mid \vec{p} \in S^D \text{ and } r \in (0, \varepsilon)\},$$

where $\varepsilon$ is such that conditions (7) and (8) are satisfied for the semi-algebraic set $A = S^D$. Let us express this query in FO.

We define the critical point query as

$$\begin{aligned} Q_{\text{crit}}(D) \;:=\; \{(\vec{p}, \vec{x}) \in \mathbb{R}^n \times \mathbb{R}^n \mid\ & \vec{x} \in Q_{0\text{-Whitney}}(D) \text{ or } \vec{p} \in S^D \\ & \text{and } \vec{x} \in Q_{k\text{-Whitney}}(D) \text{ for a some } k,\ f_j(\vec{x}) \neq 0 \text{ and} \\ & \vec{x} \text{ is a critical point of } f_j \mid Q_{k\text{-Whitney}}(D) \\ & \text{for a certain } j = 1, \dots, n+1\}. \end{aligned}$$

**Claim.** *Let $Z$ be a $C^1$ semi-algebraic set in $\mathbb{R}^n$ of dimension $k$. Then $\vec{x} \in \mathbb{R}^n$ is a critical point of $f_j \mid Z$ if and only if the tangent space of $Z$ is parallel to $\{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_j = 0\}$. Similarly, $\vec{x} \in \mathbb{R}^n$ is a critical point of $f_{n+1} \mid Z$ if and only if the tangent space of $Z$ in $\vec{x}$ is orthogonal to $\vec{p} - \vec{x}$.*

*Proof of Claim.* We prove the claim for $f = f_{n+1} = (p_1 - x_1)^2 + \dots + (p_n - x_n)^2$, the other case being subcases. We compute the differential $d_{\vec{x}}(f \mid Z)$ as follows: Locally around $\vec{x}$, we may assume that the projection on the first $k$ coordinates $\Pi\colon Z \to U \subseteq \mathbb{R}^k$ is a homeomorphism. By definition of the differential, $d_{\vec{x}}(f \mid Z) = (d_{(x_1,\dots,x_k)} g)(d_{(x_1,\dots,x_k)} \Pi^{-1})^{-1}$, where $g = (f \mid Z) \circ \Pi^{-1}$. By the $C^1$ Inverse Function Theorem, we may assume that $\Pi^{-1}\colon U \to Z\colon (x_1, \dots, x_k) \mapsto (x, \dots, x_k, \varphi_{k+1}, \dots, \varphi_n)$, where $\varphi_i(x_1, \dots, x_k)$ are $C^1$-mappings, and hence $g\colon U \mapsto \mathbb{R}\colon (x_1, \dots, x_k) = \sum_{i=1}^{k} (p_i - x_i)^2 + \sum_{j=k+1}^{n} (p_j - \varphi_j(x_1, \dots, x_k))^2$. An elementary calculation shows that the differential of $f \mid Z$ in $\vec{x}$ is the vector

$$d_{\vec{x}}(f \mid Z) = 2\left(\left((p_i - x_i) + \sum_{j=k+1}^{n}(p_j - x_j)\frac{\partial\varphi_j}{\partial x_i}(x_1, \ldots, x_k)\right)_{i=1,\ldots,k}, \underbrace{0, \ldots, 0}_{n-k \text{ times}}\right).$$

Since $d_{(x_1,\ldots,x_k)}\Pi^{-1}$ is an isomorphism between the tangent space $\mathrm{T}_{(x_1,\ldots,x_k)}U$ of $U$ in the projection $\Pi(\vec{x})$, and the tangent space $\mathrm{T}_{\vec{x}}Z$ of $Z$ in $\vec{x}$, any tangent vector $(v_1, \ldots, v_n) \in \mathrm{T}_{\vec{x}}Z$ is of the form $(d_{(x_1,\ldots,x_k)}\Pi^{-1})(v_1, \ldots, v_k)$. More specifically, any tangent vector $\vec{v} \in \mathrm{T}_{\vec{x}}Z$ can be written as

$$(v_1, \ldots, v_n) = \left(v_1, \ldots, v_k, \sum_{i=1}^{k}\frac{\partial\varphi_{k+1}}{\partial x_i}(x_1, \ldots, x_k)v_i, \ldots, \sum_{i=1}^{k}\frac{\partial\varphi_n}{\partial x_i}(x_1, \ldots, x_k)v_i\right).$$

Hence, the product

$$d_{\vec{x}}(f \mid Z)\vec{v} = 2\sum_{i=1}^{k}(p_i - x_i)v_i + 2\sum_{j=k+1}^{n}(p_j - x_j)\left(\sum_{i=1}^{k}\frac{\partial\varphi_j}{\partial x_i}(x_1, \ldots, x_k)v_i\right)$$

is equal to $2\sum_{i=1}^{n}(x_i - p_i)v_i$. This implies that the differential mapping $d_{\vec{x}}(f \mid Z)$ is not surjective if and only if $2\sum_{i=1}^{n}(p_i - x_i)v_i = 0$ for all tangent vectors $\vec{v} \in \mathrm{T}_{\vec{x}}Z$. It is clear that this implies the claim. □

The proof of the theorem now continues as follows. The tangent space query

$$Q_{\text{tangent}}(D) := \{(\vec{x}, \vec{v}) \in \mathbb{R}^n \times \mathbb{R}^n \mid S^D \text{ is } C^1,\ \vec{x} \in S^D \text{ and } \vec{v} \in \mathrm{T}_{\vec{x}}S^D\}$$

is expressible in FO [20, Lemma 2]. Because the orthogonality and parallelism of two vectors can be easily expressed in FO, the formula

$$\begin{aligned}
\varphi_{\text{crit}}(\vec{p}, \vec{x}) \;=\; & (\varphi_{0\text{-part}}(S)(\vec{x}) \wedge (\vec{x} \neq \vec{p})) \vee S(\vec{p}) \\
& \wedge \bigvee_{j=1}^{n}\Bigg(\bigvee_{k=1}^{n}(\forall\vec{v}\varphi_{\text{tangent}}(\varphi_{j\text{-part}}(S))(\vec{x}, \vec{v}) \wedge (x_k \neq p_k) \rightarrow (p_k - x_k)v_k = 0) \\
& \qquad \vee \forall\vec{v}\varphi_{\text{tangent}}(\varphi_{j\text{-part}}(S))(\vec{x}, \vec{v}) \wedge (\vec{x} \neq \vec{p}) \\
& \qquad \rightarrow (p_1 - x_1)v_1 + \cdots + (p_n - x_n)v_n = 0\Bigg)
\end{aligned}$$

expresses $Q_{\text{crit}}$ correctly by the above claim. Here, $\varphi_{j\text{-part}}$ denotes an FO formula expressing $Q_{j\text{-part}}$ for $j = 0, \ldots, n$, and $\varphi_{\text{tangent}}$ is an FO formula expressing $Q_{\text{tangent}}$.

Let $\varphi_{\text{val}}(\vec{p}, r)$ be the FO formula which expresses the query which returns the critical values of critical points given by $Q_{\text{crit}}$.

By the above there exists a minimal critical value and any value smaller than this minimal value is a cone radius. We therefore conclude that the query expressed in FO as

$$\varphi_{\text{radius}}(\vec{p}, r) := (\forall r')(\varphi_{\text{val}}(\vec{p}, r') \rightarrow r < r')$$

is a cone radius query, as desired. □

**Theorem 3.** *There exists an FO-expressible cone radius query.*

*Proof.* The proof is the same as for Theorems 1 and 2, except that the constructed $C^1$-Whitney decomposition of $\mathrm{cl}(A)$ only needs to be compatible with $A$ and only the critical points with respect to $f_{n+1}$ must be considered. □

**Corollary 1.** *Every box cone radius is a cone radius.*

*Proof.* This follows from the fact that the set of critical points identified in the proof of Theorem 1 is a superset of those identified in the proof of Theorem 3. □

## Acknowledgments

The author thanks Jan Van den Bussche for many interesting discussions and critical remarks, which led to improvements in the presentation of the paper. The author also thanks Masahiro Shiota for his feedback on the construction of the vector fields in the proof of Theorem 1.

## References

1. R. Benedetti and J.J. Risler. *Real Algebraic and Semi-Algebraic Sets*. Hermann, Paris, 1990.
2. M. Benedikt, G. Dong, L. Libkin, and L. Wong, Relational expressive power of constraint query languages. *Journal of the ACM*, 45(1):1–34, 1998.
3. M. Benedikt, M. Grohe, L. Libkin, and L. Segoufin. Reachability and connectivity queries in constraint databases. In *Proceedings of the* 19*th ACM Symposium on Principles of Database Systems*, pages 104–115. ACM Press, New York, 2000.
4. J. Bochnak, M. Coste, and M.-F. Roy. *Real Algebraic Geometry*, volume 36 of Ergebenisse der Mathematik und ihrer Grenzgebiete. Springer-Verlag, Berlin, 1998.
5. D. Gauld. *Differential Topology*: *An Introduction*, volume 72 of A Series of Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker, New York, 1982.
6. F. Geerts. Linear approximation of semi-algebraic spatial databases using transitive closure logic, in arbitrary dimension. In G. Ghelli and G. Grahne, editors, *Proceedings of the* 8*th International Workshop on Databases and Programming Languages*, volume 2397 of Lecture Notes in Computer Science, pages 182–197. Springer-Verlag, Berlin, 2002.
7. F. Geerts and B. Kuijpers. Linear approximation of planar spatial databases using transitive-closure logic. In *Proceedings of the* 19*th ACM Symposium on Principles of Database Systems*, pages 126–135. ACM Press, New York, 2000.
8. F. Geerts and B. Kuijpers. Expressing topological connectivity of spatial databases. In A. Mendelzon and R. Connor, editors, *Research Issues in Structured and Semistructured Database Programming*, volume 1949 of Lecture Notes in Computer Science, pages 224–238. Springer-Verlag, Berlin, 2001.
9. F. Geerts, B. Kuijpers, and J. Van den Bussche. Linearization and completeness results for terminating transitive closure queries on spatial databases. Submitted, 2002.
10. C.G. Gibson, K. Wirthmüller, A.A. du Plessis, and E.J.N. Looijenga. *Topological Stability of Smooth Mappings*, volume 552 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1976.
11. M. Grohe and L. Segoufin. On first-order topological queries. *Transaction on Computational Logic*, 3(3):336–358, 2002.
12. S. Grumbach and G. Kuper. Tractable recursion over geometric data. In G. Smolka, editor, *Proceedings of the* 3*rd Conference on Principles and Practice of Constraint Programming*, volume 1330 of Lecture Notes in Computer Science, pages 450–462. Springer-Verlag, Berlin, 1997.

13. P.C. Kanellakis, G.M. Kuper, and P.Z. Revesz. Constraint query languages. *Journal of Computer and System Sciences*, 51(1):26–52, 1995.
14. S. Kreutzer. Operational semantics for fixed-point logics on constraint databases. In R. Nieuwenhuis and A. Voronkov, editors, *Proceedings of the* 8*th International Conference on Logic for Programming and Automated Reasoning*, volume 2250 of Lecture Notes in Artificial Intelligence, pages 465–479. Springer-Verlag, Berlin, 2002.
15. B. Kuijpers, J. Paredaens, M. Smits, and J. Van den Bussche. Termination properties of spatial datalog. In D. Pedreschi and C. Zaniolo, editors, *Logic in Databases* (*LID* '96), volume 1154 of Lecture Notes in Computer Science, pages 101–116. Springer-Verlag, Berlin, 1996.
16. B. Kuijpers, J. Paredaens, and J. Van den Bussche. Topological elementary equivalence of closed semi-algebraic sets in the real plane. *Journal of Symbolic Logic*, 65(4):1530–1555, 2000.
17. B. Kuijpers and M. Smits. On expressing topological connectivity in spatial datalog. In V. Gaede, A. Brodsky, O. Gunter, D. Srivastava, V. Vianu, and M. Wallace, editors, *Proceedings of the* 2*nd Workshop on Constraint Databases and Applications*, volume 1191 of Lecture Notes in Computer Science, pages 116–133. Springer-Verlag, Berlin, 1997.
18. G.M. Kuper, J. Paredaens, and L. Libkin, editors. *Constraint Databases*. Springer-Verlag, Berlin, 2000.
19. J. Paredaens, J. Van den Bussche, and D. Van Gucht. Towards a theory of spatial database queries. In *Proceedings of the* 13*th ACM Symposium on Principles of Database Systems*, pages 279–288. ACM Press, New York, 1994.
20. E. Rannou. The complexity of stratification computation. *Discrete & Computational Geometry*, 19:47–79, 1998.
21. L. Segoufin and V. Vianu. Querying spatial databases via topological invariants. *Journal of Computer and System Sciences*, 61(2):276–301, 2000.
22. M. Shiota. *Geometry of Subanalytic and Semialgebraic Sets*. Birkhäuser, Basel, 1997.
23. M. Shiota. Personal communication.
24. L. van den Dries. Alfred Tarski's elimination theory for real closed fields. *Journal of Symbolic Logic*, 53(1):7–19, 1988.
25. L. van den Dries. *Tame Topology and o-minimal Structures*. Cambridge University Press, Cambridge, 1998.
26. L. van den Dries and C. Miller. Geometric categories and o-minimal structures. *Duke Mathematical Journal*, 82(2):497–540, 1996.
27. H. Whitney. Local properties of analytic varieties. *Differential and Combinatorial Topology.* "A symposium in honor of Marston Morse," pages 205–244. Princeton University Press, Princeton, NJ, 1965.
28. A.J. Wilkie. A theorem of the complement and some new o-minimal structures. *Selecta Mathematica, New Series*, 5:397–421, 1999.

*Received January* 14, 2002, *and in revised form July* 27, 2002. *Online publication July* 25, 2003.