

# Linear approximation of planar spatial databases using transitive-closure logic

Floris Geerts and Bart Kuijpers  
Limburgs Universitair Centrum  
B-3590 Diepenbeek, Belgium  
{floris.geerts,bart.kuijpers}@luc.ac.be

## ABSTRACT

We consider spatial databases in the plane that can be defined by polynomial constraint formulas. Motivated by applications in geographic information systems, we investigate linear approximations of spatial databases and study in which language they can be expressed effectively. Specifically, we show that they cannot be expressed in the standard first-order query language for polynomial constraint databases but that an extension of this first-order language with transitive closure suffices to express the approximation query in an effective manner. Furthermore, we introduce an extension of transitive-closure logic and show that this logic is complete for the computable queries on linear spatial databases. This result together with our first result implies that this extension of transitive-closure logic can express all computable topological queries on arbitrary spatial databases in the plane.

## Categories and Subject Descriptors

H.2.3 [Database Management]: Languages—*Query Languages*

## General Terms

Languages, Theory

## 1. INTRODUCTION

Spatial database systems [1, 6, 9, 17, 18, 27] are concerned with the representation and manipulation of data that have a geometrical or topological interpretation. Conceptually, spatial databases are possibly infinite sets of points in a real space  $\mathbf{R}^n$ . The framework of constraint databases, introduced in 1990 by Kanellakis, Kuper and Revesz [19] and [24], provides an elegant and powerful model for spatial databases. In the setting of the constraint model, a spatial database is finitely represented as a Boolean combination of polynomial equalities and inequalities. The head of the left arrow depicted in Figure 1 is an example of a spatial database in the plane that can be described by the formula

$$(10x)^2 + (y - 9)^2 \geq 100 \wedge (10x)^2 + (y + 9)^2 \geq 100 \wedge -1 \leq x \leq 0 \wedge -9 \leq 10y \leq 9.$$

Several authors have argued that the restriction to linear polynomial constraints provides a sufficiently general framework for spatial database applications [15, 29, 30]. Indeed, in geographic information systems (GIS), which is one of the main application areas for spatial databases, *linear approximations* are used to model spatial objects (for an overview of this field since the early 90's we refer to [1, 6, 9, 17, 18, 27]). There are also more pragmatic reasons for the restriction to linear approximations. Existing implementations of the constraint model (see the work on the system DEDALE [12, 13, 15]), are also restricted to linear polynomial constraints, not only because there is a conceptually easier way to evaluate queries expressed in the standard first-order query language for these constraint databases and because there is a slight gain in data complexity ( $\text{NC}^1$  while it is  $\text{NC}$  for arbitrary polynomial constraints [16]), but mainly because of the existence of numerous efficient algorithms for geometrical operations on linear figures [26].

Linear approximation of general spatial databases is therefore a natural spatial database query that is relevant to spatial database practice (an example of an approximation is given in Figure 1). We show that the approximation query cannot be expressed in  $\text{FO} + \text{POLY}$ , the standard first-order query language of polynomial constraint databases, however. We therefore turn to the more expressive language  $\text{FO} + \text{POLY} + \text{TC}$ , namely the extension of  $\text{FO} + \text{POLY}$  with transitive closure. We show that in this language, for a given bounded spatial database in the plane, a linear spatial database can be defined that, from a topological point of view, is homeomorphic to the given database and, from a metric point of view, approximates the given spatial database within a specified  $\varepsilon$ -error margin. For an arbitrary given spatial database, an homeomorphic linear database can be defined in  $\text{FO} + \text{POLY} + \text{TC}$ . In general, unbounded spatial databases cannot be approximated linearly.

The language  $\text{FO} + \text{POLY} + \text{TC}$  shares with most programming languages the disadvantage that the evaluations of its formulas are not guaranteed to terminate. However, if the evaluation of an  $\text{FO} + \text{POLY} + \text{TC}$  formula on a constraint database input terminates, it evaluates to an output within the constraint model. More specifically, we show that the evaluation of our approximation query terminates and that it works correctly. In the course of proving these results, we

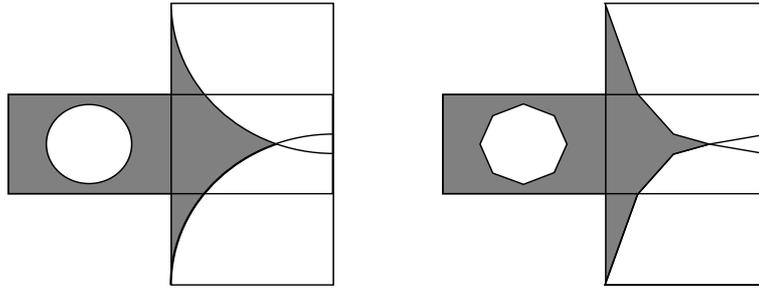


Figure 1: A spatial database and an  $\varepsilon$ -approximation

also show that the tangent line in a point to a curve in a spatial database can be expressed in  $\text{FO} + \text{POLY}$  and that as a consequence, for instance, cusp points and corner points of curves can be identified in  $\text{FO} + \text{POLY}$ .

It should be noted that it is *not* obvious (in fact, this problem is open) that the approximation query, or any other computable query for that matter, is expressible in  $\text{FO} + \text{POLY} + \text{TC}$ . However, if the transitive closure operation is extended with a stop condition, we denote this language with  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ , then it is not difficult to see that  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  is computationally complete on finite spatial databases (with rational coordinates). Using recent results by Vandeurzen et al. on the encoding of linear spatial databases by means of finite spatial databases [3, 8, 31] (this encoding and the corresponding decoding are expressible in  $\text{FO} + \text{POLY}$ ), we show that  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  is also computationally complete on linear spatial databases (that can be described by linear polynomials with *integer* or *rational* coefficients).

We also show that  $\text{FO} + \text{POLY} + \text{TC}$  allows for the expression of aggregation operators, such as the area of regions and the length of curves, on linear spatial databases (that may even be described by linear polynomials with *algebraic* coefficients). Benedikt and Libkin [4] have already shown that these aggregate operators are not first-order definable, but that they can be expressed in suitable extensions of  $\text{FO} + \text{POLY}$ .

We will also show a refinement of our approximation result, to the effect that the resulting approximation can be guaranteed to be definable by means of linear polynomials with integer coefficients. This refinement implies that  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  is computationally complete on arbitrary planar spatial databases for what concerns the Boolean *topological queries*. For a number of spatial database applications exact metric information is not important and only topological properties of the spatial data are relevant. In such applications, only topological queries are of interest. This class of queries has received ample attention in the constraint database literature (for an overview we refer to Chapter 10 in [24] and references therein [23]).

As an example, we examine the topological connectivity query and show that the connectivity test for arbitrary spatial databases is expressible in  $\text{FO} + \text{POLY} + \text{TC}$ .

The question remains open whether or not  $\text{FO} + \text{POLY} + \text{TC}$  is complete for all computable queries on arbitrary planar spatial databases in the polynomial constraint model.

This paper is organized as follows. In Section 2, we define spatial databases and the languages  $\text{FO} + \text{POLY}$  and  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ . We also review some topological properties of spatial databases. In Section 3, the expression of the linear approximation of a spatial database is given and proven to work correctly. Also results concerning the expression of tangent lines in  $\text{FO} + \text{POLY}$  and of aggregation operators for linear databases in  $\text{FO} + \text{POLY} + \text{TC}$  are discussed in this section. Finally, Section 4 contains all the completeness results.

## 2. PRELIMINARIES AND DEFINITIONS

In this section, we define spatial databases and the standard first-order query language for spatial databases. We also define two extensions of this query language with transitive closure. Finally, we review some basic properties concerning the local topological behavior of spatial databases.

### 2.1 The constraint model for spatial databases

Let  $\mathbf{R}$  denote the set of the real numbers, and  $\mathbf{R}^2$  the real plane.

*Definition 1.* A *spatial relation* is a geometrical figure in  $\mathbf{R}^2$  that can be defined as a Boolean combination (union, intersection and complement) of sets of the form  $\{(x, y) \mid p(x, y) > 0\}$ , where  $p(x, y)$  is a polynomial with integer coefficients in the real variables  $x$  and  $y$ . A *spatial database* is a finite sequence of spatial relations.

A **Z-linear** (resp. **A-linear**) *spatial relation* is a spatial relation that can be defined in terms of linear polynomials with integer (resp. algebraic) coefficients. A **Z-linear** (resp. **A-linear**) *spatial database* is a finite sequence of **Z-linear** (resp. **A-linear**) spatial relations.

We remark that in mathematical terminology, spatial relations are called *semi-algebraic sets*. Linear spatial relations are also referred to as *semilinear sets*. Every **Z-linear** relation is also **A-linear**, but the converse isn't true, e.g., the **A-linear** relation  $\{(x, y) \in \mathbf{R}^2 \mid y = \sqrt{2}\}$  is not **Z-linear**. Note that  $p(x, y) = 0$  is used to abbreviate  $\neg(p(x, y) > 0) \wedge \neg(\neg p(x, y) > 0)$ .

In this paper, we will use FO + POLY, the relational calculus augmented with polynomial inequalities as standard spatial query language.

*Definition 2.* A formula in FO + POLY is a first-order logic formula  $\varphi(x_1, \dots, x_n, S_1, \dots, S_m)$  built, using the logical connectives and quantifiers, from two kinds of atomic formulas:  $S_\ell(x, y)$  and  $p(x_1, \dots, x_k) > 0$ , where each  $S_\ell$ , for  $1 \leq \ell \leq m$ , is a binary relation name representing a spatial relation from the input database and  $p(x_1, \dots, x_k)$  is a polynomial in the variables  $x_1, \dots, x_k$  with integer coefficients.

Variables in such formulas are assumed to range over  $\mathbf{R}$ . Tarski's quantifier-elimination property for first-order logic over the reals guarantees that FO + POLY formulas can be evaluated effectively on spatial database inputs and their result is a spatial relation (in  $\mathbf{R}^n$ ) that also can be described by means of polynomial constraints over the reals [7, 28].

If  $\varphi(x_1, \dots, x_n, S_1, \dots, S_m)$  is an FO + POLY formula,  $a_1, \dots, a_n$  are reals, and  $(A_1, \dots, A_m)$  is a spatial database, we denote by  $(a_1, \dots, a_n, A_1, \dots, A_m) \models \varphi(x_1, \dots, x_n, S_1, \dots, S_m)$  that  $(a_1, \dots, a_n, A_1, \dots, A_m)$  satisfies  $\varphi$ .

## 2.2 Transitive closure logic

A second language that we will use is FO + POLY + TC, an extension of FO + POLY with recursion.

*Definition 3.* A formula in FO + POLY + TC is a formula built in the same way as an FO + POLY formula, with the following extra formation rule: if  $\psi(x_1, \dots, x_k, y_1, \dots, y_k, S_1, \dots, S_m)$  is a formula,  $\vec{s}$  and  $\vec{t}$  are  $k$  dimensional vectors of variables, then

$$[\text{TC}_{\vec{x}, \vec{y}} \psi(\vec{x}, \vec{y}, S_1, \dots, S_m)](\vec{s}, \vec{t}),$$

is also a formula whose free variables are  $\vec{s}$  and  $\vec{t}$ . It is assumed that  $\vec{x}$  and  $\vec{y}$  are pairwise distinct.

We must evaluate FO + POLY + TC formulas on spatial databases in an effective way. Moreover, in analogy with FO + POLY formulas, we want as output a polynomial constraint description of the result of the evaluation.

For the first-order subformulas of FO + POLY + TC, this is done as for FO + POLY: the description of the database relations are plugged in to the formula, and quantifier elimination is employed.

For a formula of the form  $[\text{TC}\psi](\vec{s}, \vec{t})$ , we produce constraint descriptions of the successive stages  $\text{TC}_i R_\psi$  in the computation of the transitive closure, as follows. The binary (vector) relation  $R_\psi$  is defined, given a database  $\mathcal{D} = (A_1, \dots, A_m)$ , as  $\{(\vec{a}, \vec{b}) \mid (\vec{a}, \vec{b}, \mathcal{D}) \models \psi(\vec{x}, \vec{y}, S_1, \dots, S_m)\}$ . We put the first stage  $\text{TC}_1 R_\psi := R_\psi$ , and for  $i > 1$ , the stage  $\text{TC}_i R_\psi$  equals  $\text{TC}_{i-1} R_\psi \cup \{(\vec{x}, \vec{y}) \mid (\exists \vec{z})((\vec{x}, \vec{z}) \in \text{TC}_{i-1} R_\psi \wedge (\vec{z}, \vec{y}) \in R_\psi)\}$ . For each  $i$ , we can effectively check, whether

$$\text{TC}_i R_\psi = \text{TC}_{i-1} R_\psi.$$

If we find such an  $i$ , we can output a constraint description of  $\text{TC}_i R_\psi$  (or  $\text{TC}_{i-1} R_\psi$  for that matter) as the result. If

we never find such an  $i$ , the computation does not end and the result is undefined. Finally, we check whether the pair of vectors  $(\vec{s}, \vec{t})$  is in the result.

An example when no  $i$  exists is the following. Consider the relation  $R = \{(x, y) \in \mathbf{R}^2 \mid y = 2x\}$ . At stage  $i$ , we have the relation  $\text{TC}_i R = \{(x, y) \in \mathbf{R}^2 \mid (\exists j)(0 \leq j \leq i \text{ and } y = 2^j x)\}$ .

We now extend the language FO + POLY + TC with a stop condition.

*Definition 4.* The language FO+POLY+TC $^\Sigma$  is defined as FO+POLY+TC, except that instead of the transitive closure formation rule, we now have a the following formation rule. If  $\psi(x_1, \dots, x_k, y_1, \dots, y_k, S_1, \dots, S_m)$  is a formula,  $\vec{s}$  and  $\vec{t}$  are  $k$  dimensional vectors of variables, then

$$[\text{TC}_{\vec{x}, \vec{y}} \psi(\vec{x}, \vec{y}, S_1, \dots, S_m) \mid_T \sigma(S_1, \dots, S_m, T)](\vec{s}, \vec{t}),$$

is also a formula with free variables  $\vec{s}$  and  $\vec{t}$ , and where  $T$  denotes a new binary relation name representing a stage in the computation of the transitive closure of  $R_\psi$ . The *stop condition*  $\sigma$  is a formula with may use  $T(\vec{x}, \vec{y})$  as extra atomic formula. It is assumed that  $\vec{x}$  and  $\vec{y}$  are pairwise distinct.

The evaluation of an FO+POLY+TC $^\Sigma$  formula is as follows. For the first-order subformulas of FO + POLY + TC $^\Sigma$ , this is done as for FO + POLY.

For a formula of the form  $[\text{TC}\psi \mid_T \sigma](\vec{s}, \vec{t})$ , we again produce constraint descriptions of the successive stages  $\text{TC}_i^\Sigma R_\psi$  in the computation of the transitive closure, as follows. The binary (vector) relation  $R_\psi$  is defined as before, and we put the first stage  $\text{TC}_1^\Sigma := R_\psi$ . Then for  $i > 1$ , stage  $\text{TC}_i^\Sigma R_\psi$  equals  $\text{TC}_{i-1}^\Sigma R_\psi \cup \{(\vec{x}, \vec{y}) \mid (\exists \vec{z})((\vec{x}, \vec{z}) \in \text{TC}_{i-1}^\Sigma R_\psi \wedge (\vec{z}, \vec{y}) \in R_\psi)\}$ .

The difference with the evaluation of an FO + POLY + TC formula is that for each  $i$  we check (again, this can be done in an effective way) whether

$$\text{TC}_i^\Sigma R_\psi = \text{TC}_{i-1}^\Sigma R_\psi,$$

or, whether

$$(\mathcal{D}, \text{TC}_i^\Sigma) \models \sigma(S_1, \dots, S_m, T).$$

If we find such an  $i$ , we can output a constraint description of  $\text{TC}_i R_\psi$  as the result. If we never find such an  $i$ , the computation does not end and the result is undefined. Finally, we check whether the pair of vectors  $(\vec{s}, \vec{t})$  is in the result.

Remark that FO + POLY + TC is included in FO + POLY + TC $^\Sigma$ . Indeed, when  $\sigma := \text{false}$  we obtain the definition of FO + POLY + TC. It is an open problem whether the inclusion is strict.

## 2.3 Spatial relations are locally conical

In this section, we discuss a fundamental topological property of spatial databases. First, we introduce the metric and topological terminology needed in the remainder of the paper. The Euclidean distance between two points  $p, q \in \mathbf{R}^2$

will be denoted by  $d(p, q)$ . If  $p$  is a point of  $\mathbf{R}^2$  and  $\varepsilon > 0$  a real number, we denote by  $B^2(p, \varepsilon)$  the closed disk with center  $p$  and radius  $\varepsilon$  and by  $S^1(p, \varepsilon)$  its boundary. A *homeomorphism*  $h : \mathbf{R}^2 \rightarrow \mathbf{R}^2$  is a continuous bijective function whose inverse is also continuous. An *isotopy* of the plane is a homeomorphism of the plane that is isotopic to the identity. Two homeomorphisms  $f$  and  $g$  are *isotopic* if there is a continuous function  $F : \mathbf{R}^2 \times [0, 1] \rightarrow \mathbf{R}^2$  such that for each  $t \in [0, 1]$ , the function  $F_t : \mathbf{R}^2 \rightarrow \mathbf{R}^2 : p \mapsto F(p, t)$  is a homeomorphism and  $F_0$  is  $f$  and  $F_1$  is  $g$ . Two sets are said to be *isotopic* if there is an isotopy that maps one to the other. Finally, if  $A$  is a spatial relation, we denote its topological border by  $\partial A$ , its topological closure by  $\bar{A}$  and its interior by  $A^\circ$ .

By a *cone* with top  $p = (x_p, y_p)$  and base  $B \subseteq \mathbf{R}^2$ , we mean the set  $\{\lambda \cdot (x_p, y_p) + (1 - \lambda) \cdot (x', y') \mid 0 \leq \lambda \leq 1 \text{ and } (x', y') \in B\}$ . Spatial relations have the following fundamental topological property.

*Property 1.* [5, Theorem 9.3.5] For a spatial relation  $A$  and a point  $p$  in the plane there exists a radius  $\varepsilon_{A,p}$  such that for each  $0 < \varepsilon < \varepsilon_{A,p}$ ,  $B^2(p, \varepsilon) \cap A$  is isotopic to the cone with top  $p$  and base  $S^1(p, \varepsilon) \cap A$ .

We call such a radius  $\varepsilon_{A,p}$ , a *cone radius of  $p$  in  $A$* . If  $A$  is clear from the context, we just call it a cone radius of  $p$  and denote it by  $\varepsilon_p$ . We remark that a spatial relation is also conical towards infinity (or around the point  $\infty$ ), i.e., there exists a radius  $\varepsilon_\infty$  such that the spatial relation is conical outside the disk  $B^2((0, 0), \varepsilon_\infty)$ . The authors have shown the following expressiveness property.

*Property 2.* [11] For a given spatial relation  $A$  and point  $p$ , there exists an FO + POLY formula which defines a cone radius  $\varepsilon_p$ . For a given spatial relation  $A$ , there exists an FO + POLY formula which defines a radius  $\varepsilon_\infty$ .

More concretely, the value defined by this FO + POLY formula is one third of the maximal value such a cone radius can take (and three times the minimal value for the point  $\infty$ ).

The bases of the cones are finite unions of points and open arc segments on the circles  $S^1(p, \varepsilon_p)$ . Moreover, it can be shown that only a finite number of points  $p$  in the plane have a cone with base *different* from the following six types: two points on the circle, a closed arc segment on the circle (i.e. the union of an open arc segment and two points) and the circle itself for  $p$  belonging to the spatial relation  $A$ , and the complements (on the circle) of these three types, for  $p$  belonging to the complement of  $A$ . The points which have a cone of one of these six types, are called *regular* points of  $A$ . Non-regular points are called *singular*. It is known that a spatial relation has only a finite number of singular points. We also remark that the regularity/singularity of a point is expressible in FO + POLY [21].

Since Property 2 is only proven for planar spatial databases and the construction of the approximation in the next section heavily depends on this result, our results are restricted

to planar spatial databases. It is not clear if this Property also holds for spatial databases in higher dimensions.

### 3. LINEAR APPROXIMATION OF PLANAR SPATIAL DATABASES

First, we define the concept of an  $\varepsilon$ -approximation of a spatial relation, i.e., a linear spatial database that, from a topological point of view, is isotopic to a given spatial relation and, from a metric point of view, approximates it within an  $\varepsilon$ -error margin.

*Definition 5.* Let  $A$  be a spatial relation,  $(A_1, \dots, A_m)$  a spatial database and  $\varepsilon > 0$  a real number.

We call an  $\mathbf{A}$ -linear spatial database  $(B_1, \dots, B_m)$  an (*algebraic*)  $\varepsilon$ -approximation of  $(A_1, \dots, A_m)$  if there exists an isotopy  $h$  of the plane such that  $h(A_i) = B_i$  and if for every point  $p \in \partial A_i$  there exists a point  $q \in \partial B_i$  such that  $d(p, q) \leq \min\{\varepsilon, \varepsilon_{A_i,p}\}$  (where  $\varepsilon_{A_i,p}$  is the cone radius of  $p$  in  $A_i$ ), for every  $1 \leq i \leq m$ . We call an  $\mathbf{A}$ -linear spatial relation  $B$  an (*algebraic*)  $\varepsilon$ -approximation of  $A$  if the database  $(B)$  is an  $\varepsilon$ -approximation of  $(A)$ . If an  $\varepsilon$ -approximation of a spatial database or a spatial relation is furthermore  $\mathbf{Z}$ -linear, we call it a *rational  $\varepsilon$ -approximation*.

As we will show,  $\varepsilon$ -approximations always exists for *bounded* spatial relations and databases. This is no longer true for unbounded spatial relations. Consider e.g., the unbounded spatial relation  $\{(x, y) \in \mathbf{R}^2 \mid y = x^2\}$ . It is easy to see that this parabola cannot be approximated by a finite number of line segments, and hence has no  $\varepsilon$ -approximation for any  $\varepsilon > 0$ .

We first show that approximation cannot be expressed in FO + POLY.

**PROPOSITION 3.1.** *Let  $\varepsilon > 0$  be a real number. A query that returns an  $\varepsilon$ -approximation of a spatial relation is not expressible in FO + POLY.*

**PROOF.** Let  $R$  be a binary relation over the reals. Let  $\varphi$  be the query on  $R$  that returns the corner points of an  $\varepsilon$ -approximation of the circle through the points of  $R$  if  $R$  consists of three non-collinear points and that returns  $R$  itself in any other case. Clearly, the construction of a circle through three points is expressible in FO + POLY. The same holds for the selection of the corner points of a linear spatial relation (see also Lemma 3.3 of the next section). Hence, if we assume that an  $\varepsilon$ -approximation can be defined in FO + POLY, then  $\varphi$  is also expressible in FO + POLY. However, the number of corner points,  $|\varphi(R)|$ , can be made arbitrarily large by choosing  $R$  to consist of three points far enough apart points. This contradicts the Dichotomy Theorem [3], which guarantees the existence of a polynomial  $p_\varphi$  such that  $|\varphi(R)| < p_\varphi(|R|)$  on any finite relation  $R$ .  $\square$

#### 3.1 Effective algebraic approximations

We now show that an  $\varepsilon$ -approximation can be expressed in classical FO + POLY + TC (i.e., FO + POLY + TC without

the use of stop conditions). In the course of doing so we will also show that the tangent line in a point to a semi-algebraic curve can be expressed in FO + POLY (if it exists). More specifically, let  $A$  be a spatial relation and  $p$  be a point on a curve of  $\partial A$ . It is a well-known fact of algebraic geometry that the left and right tangent half-line in  $p$  always exist [2]. The following property shows how these half-lines can be defined in FO + POLY. For a proof sketch we refer to the Appendix.

**LEMMA 3.2.** *Let  $A$  be a spatial relation and  $p = (x_p, y_p)$  be a point of  $\partial A$  whose cone is based on two points. There exists a formula  $\varphi_{\text{tangent}}(S, x, y, u, v)$  in FO+POLY such that  $(A, x_p, y_p, a, b) \models \varphi_{\text{tangent}}(S, x, y, u, v)$  if and only if the line  $a(x - x_p) + b(y - y_p) = 0$  is the left or right tangent line to  $\partial A$  in  $p$ .*

**PROOF.** Let  $A$  be a spatial relation and  $p = (x_p, y_p)$  be a point of  $\partial A$  as in the statement of the lemma. The desired formula  $\varphi_{\text{tangent}}$  can be written as a disjunction  $\bigvee_{i=1}^4 \varphi_{\text{tangent}}^i$ , where,  $(A, x_p, y_p, a, b) \models \varphi_{\text{tangent}}^i(S, x, y, u, v)$  if and only if the line  $a(x - x_p) + b(y - y_p) = 0$  is the tangent line in  $p$  to  $A$ , in the  $i$ th quadrant of the plane in  $p$ .

Since all formulas  $\varphi_{\text{tangent}}^i$  are similar, we only describe the case of the first quadrant of  $p$ . The cone of  $p$  is based on two points, hence, locally in the first quadrant, each of the two branches of  $\partial A$  in  $p$  is defined by a single polynomial equation  $P_1(x, y) = 0$ , resp.  $P_2(x, y) = 0$ .

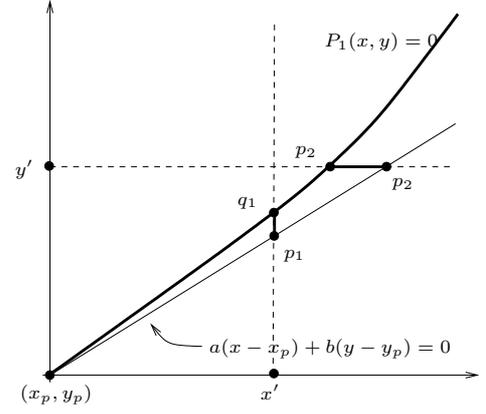
We further assume that locally only the branch of  $\partial A$ , defined by  $P_1(x, y) = 0$  lies in the first quadrant of  $p$  (if the two branches are located in the first quadrant, a further case analysis is required). In the first quadrant, the right partial derivative  $\lim_{x \downarrow x_p} \frac{\partial P_1}{\partial x}(x_p, y_p) = a$ , and the left partial derivative  $\lim_{y \downarrow y_p} \frac{\partial P_1}{\partial y}(x_p, y_p) = b$  exist, and  $(a, b) \neq (0, 0)$ . Assume that  $a \neq 0$  and  $b \neq 0$  (the wanted formula will be a disjunction of three disjuncts, where each disjunct corresponds to either  $a \neq 0 \wedge b \neq 0$ ,  $a = 0 \wedge b \neq 0$ , or  $a \neq 0 \wedge b = 0$ ). The Implicit Function Theorem guarantees the existence of two differentiable functions  $g(x)$  and  $f(y)$  such that locally, in the first quadrant,  $P_1(x, g(x)) = 0$ , and  $P_1(f(y), y) = 0$ . This implies that

$$\frac{dg}{dx}(x_p) = -\frac{a}{b} \quad \text{and} \quad \frac{df}{dy}(y_p) = -\frac{b}{a}.$$

We express this by an FO + POLY formula  $\varphi_{\text{tangent}}^1$ , which expresses the existence of a number  $\delta > 0$  such that for each  $x', y'$  with  $x_p < x' < x_p + \delta$  and  $y_p < y' < y_p + \delta$ ,

$$\begin{aligned} |p_1 - q_1| &< \varepsilon(x' - x_p), \\ |p_2 - q_2| &< \varepsilon(y' - y_p), \end{aligned}$$

where  $p_1$  and  $p_2$  are the intersection points of the line  $a(x - x_p) + b(y - y_p) = 0$  with the lines  $x = x'$  and  $y = y'$  respectively, and  $q_1 = |g(x') - y_p|$  and  $q_2 = |f(y') - x_p|$  are the intersection points of  $\partial A$  with the lines  $x = x'$  and  $y = y'$  respectively. Hence,  $\varphi_{\text{tangent}}^1$  defines the coefficients of the tangent half-line in  $p$  to  $\partial A$  in the first quadrant. Note that, since all multiples of  $\lambda \cdot (a, b)$  satisfy the formula  $\varphi_{\text{tangent}}^1$ , we select the one such that  $\lambda \cdot (a + b) = 1$ .  $\square$



**Figure 2: Illustration of the definition of the tangent line**

The formula  $\varphi_{\text{tangent}}(S, x, y, u, v)$  defines the tangent lines to a curve in a point. Basically, there are three possible configurations of the left and right tangent line.

- First, they can be collinear and oriented in opposite directions; in this case there exists a *unique tangent line* to  $\partial A$  in  $p$ .
- Second, they can be collinear and oriented in the same direction; in this case  $p$  is a *cusp point* of  $\partial A$ .
- Third, the left and right tangent line can form a constant angle in  $p$ . In this case two different algebraic curves of  $\partial A$  meet in  $p$ , and we call  $p$  a *corner point* of  $\partial A$  or  $A$ .

**LEMMA 3.3.** *Let  $A$  be a spatial relation. The cusp points, the corner points and the points with a certain fixed tangent line to  $\partial A$  can be defined in FO + POLY.*

Before turning to the main result of this section, we remark that cusp and corner points of a spatial relation are finite in number [2].

**THEOREM 3.4.** *Given  $\varepsilon > 0$  and a spatial database  $(A_1, \dots, A_m)$  in which each  $A_i$  is bounded, there exists an FO + POLY + TC formula which defines an algebraic  $\varepsilon$ -approximation of  $A_1, \dots, A_m$ .*

**PROOF.** We first discuss the case  $m = 1$  in considerable detail and next sketch the case  $m > 1$ .

- $m = 1$ : Let  $A$  be a bounded spatial relation, and  $\varepsilon > 0$  a real number. The wanted formula performs two main operations: first, locally within  $\varepsilon$ -environments it rectifies  $A$  around its singular points, its cusp points, its corner points and points on a non-linear curve with a vertical tangent line to  $\partial A$ ; secondly, outside these  $\varepsilon$ -environments (where  $A$  behaves smoothly) it computes an  $\varepsilon$ -approximation. As we will show, the first operation can be performed in FO + POLY, whereas the second needs the additional power of the transitive closure.

For the first operation, we determine in FO + POLY the singular points of  $A$  (using Property 2), the cusp points, the

corner points and the points on a non-linear curve with a vertical tangent line to  $\partial A$  (using Lemma 3.3). For such points  $p$ , we compute the minimum of  $\varepsilon$  and the cone radius  $\varepsilon_p$  of  $p$  (given by Property 2). Next, we replace  $B^2(p, \min\{\varepsilon, \varepsilon_p\}) \cap A$  by the cone with top  $p$  and base  $S^1(p, \min\{\varepsilon, \varepsilon_p\}) \cap A$ . This locally rectifies  $A$  around these points within an  $\varepsilon$ -error margin. Let the union of the interiors of all the above balls  $B^2(p, \min\{\varepsilon, \varepsilon_p\})$  be called  $B$ , and denote  $A \setminus B$  by  $\hat{A}$ .

The part of  $\partial A$  in  $\hat{A}$  is a finite union of simple algebraic curve segments, which we denote by  $\gamma$ . The set  $\gamma$  no longer has singular points, cusp points or corner points (except for the endpoints). It can also no longer contain closed curves (i.e., curves homeomorphic to the unit circle), since such closed curves contain at least two points with a vertical tangent line. We further remark that, apart from the endpoints, each point of  $\gamma$  has a unique tangent line to  $\gamma$ . Consider the function  $\mathcal{E} : \gamma \rightarrow \mathbf{R}^+ : p \mapsto \varepsilon_{A,p}$ , where  $\varepsilon_{A,p}$  is the cone radius of  $p$  in  $A$  given by Property 2. The mapping  $\mathcal{E}$  is a continuous semi-algebraic function and therefore maps the compact set  $\gamma$  onto a compact subset  $\mathcal{E}(\gamma)$  of  $\mathbf{R}^+$ . The set  $\mathcal{E}(\gamma)$  is therefore closed and bounded in  $\mathbf{R}^+$  and as a consequence has a minimum element  $\varepsilon_0$ . Let  $\varepsilon_0^*$  be the minimum of the input value  $\varepsilon$  and  $\varepsilon_0$ . Remark that  $\varepsilon_0^*$  can be defined in FO + POLY.

We now construct in FO+POLY+TC linear  $\varepsilon$ -approximations of the curves in  $\gamma$ , by walking with steps of length  $\varepsilon_0^*$  along these curves. For this we define in FO + POLY the binary point relation *Step* as  $\{(p, q) \in \gamma \times \gamma \mid d(p, q) = \varepsilon_0^*\}$ . Let  $\partial\gamma$  denote the set of endpoints of the curves in  $\gamma$ . Next, we define the set

$$\{p \in \gamma \mid (\exists q) \in \partial\gamma \text{ and } (p, q) \in \text{TC}(\text{Step})\}.$$

This set is finite and will be the set of corner points of the  $\varepsilon$ -approximation of  $\hat{A}$ . Note that no stop condition is needed, because the transitive closure of the relation *Step* terminates. We define the neighbors of a corner point  $p$ , as those corner points  $q$ , such that no other corner point  $r$  satisfies  $d(p, r) < d(p, q)$  and  $d(q, r) < d(p, q)$ . It is clear that a corner point  $p$  has always two neighbors, except when  $p \in \partial\gamma$ , in which case it has only one.

Now let  $(a, b)$  be the coefficients of the tangent line to  $A$  in  $p$ , as defined by  $\varphi_{\text{tangent}}$  of Lemma 3.2. The coefficients of the line perpendicular to the tangent line in  $p$  to  $A$  (call this line  $\ell$ ) are then given by the pair  $(-b, a)$ . Let  $q$  be a neighbor on  $\gamma$  of  $p$ , let  $H$  be the half plane containing  $q$  and bounded by  $\ell$ . We now replace the intersection  $B^2(p, d(p, q)) \cap A \cap H$  by the cone with top  $p$  and base the intersection  $S^1(p, d(p, q)) \cap A \cap H$ . Let  $\hat{A}_\varepsilon$  be the union of the semi-algebraic sets constructed from  $\hat{A}$  by performing the above procedure for each corner point, and of the semi-algebraic set  $B$ , which is previously defined. We will now show that the semi-algebraic set  $\hat{A}_\varepsilon$  is indeed an  $\varepsilon$ -approximation of the given spatial relation  $A$ .

It is clear that  $\hat{A}_\varepsilon$  is an  $\mathbf{A}$ -linear spatial relation. Indeed, since  $\gamma$  consists of a finite number of compact curve segments, there exists a bound on the lengths of these curve segments [2, Proposition 4.6.10]. As a consequence, the number of steps needed to step over a curve in  $\gamma$ , starting from

$p \in \partial\gamma$  and arriving closer than  $\varepsilon$  to the other endpoint  $q \in \partial\gamma$ , is also bounded. This means that only a finite number of stages in the computation of the transitive closure are needed to find the corner points of  $\gamma$ . These corner points, together with the singular points, cusp points, points on a non-linear curve with vertical tangent line are finite. Moreover, locally around these finitely many points linear constructions are done. This gives rise to a linear spatial relation.

It is also clear from the construction of the spatial relation  $\hat{A}_\varepsilon$  that the set  $\hat{A}_\varepsilon$  satisfies the distance condition in Definition 5.

Finally, the set  $\hat{A}_\varepsilon$  is isotopic to  $A$ . Lemma 1 guarantees locally the existence of isotopies, mapping the database  $A$  locally to the cones, which are the building blocks in the construction of  $\hat{A}_\varepsilon$ . Using standard techniques from topology in  $\mathbf{R}^2$ , it can be shown that these isotopies can be merged into a single isotopy  $h$  of the plane, such that  $h(A) = \hat{A}_\varepsilon$  [25].

Hence,  $\hat{A}_\varepsilon$  is an  $\varepsilon$ -approximation of  $A$ .

•  $m > 1$ : If the spatial database contains more than one relation, we can follow a similar strategy, except that the interactions between the different input relations have to be taken into account. First, for an input database  $(A_1, \dots, A_m)$ , we compute  $2^m$  spatial relations (in FO + POLY), namely, every set

$$A_1^{\alpha_1} \cap A_2^{\alpha_2} \cap \dots \cap A_m^{\alpha_m},$$

where  $A_i^{\alpha_i}$  can be  $A_i$  or  $A_i^c$ , the complement of  $A_i$ . Then we start, as in the case  $m = 1$ , by determining in each of these  $2^m$  sets the “significant points” (i.e., the singular points, cusp points, corner points and points on a non-linear curve with a vertical tangent line). For each such point  $p$ , we determine its cone radius with respect to the sets where  $p$  is significant and let  $\varepsilon_p$  be equal to the minimum of the thus obtained values. In each of the  $2^m$  sets we uniformly do a local rectification in  $B^2(p, \min\{\varepsilon, \varepsilon_p\})$ . Then, as in the case  $m = 1$ , what remains to be approximated is located in  $D$ , the complement in the plane of the union of the above mentioned balls. In each of the  $2^m$  sets, the part lying in  $D$  consists of a finite union of simple algebraic curve segments. For each of the  $2^m$  sets a minimal step value  $\varepsilon_{0,i}^*$  is determined (see the case  $m = 1$ ). We take  $\varepsilon_{0,i}^*$  to be 1 for sets with no curves. Then we let  $\varepsilon_0^*$  be equal to the minimum of  $\{\varepsilon_{0,1}^*, \dots, \varepsilon_{0,2^m}^*\}$ . Then the curves are approximated uniformly (in all the sets in which they appear) using this value  $\varepsilon_0^*$  (again, as in the case of  $m = 1$  by an FO + POLY + TC formula). Since all approximations are done uniformly, all topological relations between the different sets are respected and the approximations of the individual sets  $A_i$  are obtained as the union of the appropriate sets among the  $2^m$  computed approximations.  $\square$

The following property is a direct consequence of the construction in the proof of the previous Theorem 3.4.

**COROLLARY 3.5.** *Given an  $\mathbf{A}$ -linear spatial database, the  $\varepsilon$ -approximation, defined by the FO + POLY + TC construction of Theorem 3.4, is the spatial database itself.*

**THEOREM 3.6.** *There is an FO + POLY + TC formula that defines, on a given  $\varepsilon > 0$  and a given spatial database  $(A_1, \dots, A_m)$ , an  $\mathbf{A}$ -linear spatial database  $(B_1, \dots, B_m)$  for which there is an isotopy  $h$  of the plane such that  $h(A_i) = B_i$ , for  $1 \leq i \leq m$ .*

**PROOF.** (Sketch) We only sketch the proof for  $m = 1$ . By Property 2, there exists an FO + POLY formula that defines a value  $\varepsilon_{A,\infty}$  such that the spatial relation  $A$  is conical towards  $\infty$  outside the disk  $B^2((0,0), \varepsilon_{A,\infty})$ . Outside this disk an isotopic rectification of  $A$  can be defined in FO + POLY. Inside this disk, a linear database isotopic to  $B^2((0,0), \varepsilon_{A,\infty}) \cap A$  can be defined in FO + POLY + TC, using Theorem 3.4. Since the two isotopies above coincide on  $S^1((0,0), \varepsilon_{A,\infty})$ , they form the desired isotopy of the plane.  $\square$

### 3.2 Aggregation

If  $A$  is a bounded  $\mathbf{A}$ -linear spatial relation, we denote by  $\text{Area}(A)$  the area of  $\overline{A}$ , and by  $\text{Length}(A)$  the total length of the curves in  $\partial A$ .

**PROPOSITION 3.7.** *There exist FO + POLY + TC formulas that compute for a given bounded  $\mathbf{A}$ -linear spatial relation  $A$ , the values  $\text{Area}(A)$  and  $\text{Length}(A)$ .*

**PROOF.** (Sketch) Given a linear spatial relation  $A$  that is fully two-dimensional, the set  $B$  of its corner points and singular points can be determined in FO + POLY. From  $B$  a triangular covering of the plane can be computed, namely by first considering the set of all lines  $\mathcal{L}$  between two points of  $B$ , and second by constructing the set of triangles  $\mathcal{T} = \{(p_1, p_2, p_3) \mid p_i = L \cap M, \text{ for } i = 1, 2, 3 \text{ and with } L, M \in \mathcal{L}\}$ . Next, consider the set of all triangles of  $\mathcal{T}$  belonging to  $A$  and select only those triangles which do not contain any other triangle. This construction is in FO + POLY, and these triangles clearly define a triangulation of  $A$ . Moreover, the area of each triangle in this triangulation of  $A$  can also be determined in FO + POLY. This results in a finite bag  $C$  of real numbers. The sum of the values in  $C$  can be computed in FO + POLY + TC. We show this for the case where  $C = \{c_1 < \dots < c_n\}$  is a set (for bags the argument is similar but more elaborate). Define the binary relation  $D = \{((x, c_i), (x + c_i, c_{i+1})) \mid x \in \mathbf{R} \text{ and } 1 \leq i < n\}$ . Then  $y + c_n$  is the sum of the elements in  $C$ , where  $y$  is such that  $((0, c_1), (y, c_n))$  belongs to the transitive closure of  $D$ . A similar argument can be used for  $\text{Length}(A)$ .  $\square$

### 3.3 Rational approximations

We now refine the previous two theorems to *rational  $\varepsilon$ -approximations*. We need the following technical lemma concerning the rational approximation of real numbers.

**LEMMA 3.8.** *There exists an FO + POLY + TC $^\Sigma$  formula that computes, for a given  $\varepsilon > 0$  and a given finite set  $A$  of real numbers, a rational  $\varepsilon$ -approximation of  $A$ .*

**PROOF.** First we compute the minimal natural number  $N$ , such that  $\varepsilon N > 1$ . This can be done in FO + POLY + TC $^\Sigma$  as follows. Let  $\psi_1(x, y) = 0 \leq x \wedge y = x + 1$  and

$\sigma_1(\varepsilon, T_1) = (\exists x)((0, x) \in T_1 \wedge (\varepsilon(x - 1) \leq 1 < \varepsilon x))$ . Hence, the natural number  $N$  is defined by the formula  $\Psi_1(x, \varepsilon) = \max_x[\text{TC}\psi_1 \upharpoonright_{T_1} \sigma_1](0, x)$ . Let  $A = A^- \cup A^+$ , with  $A^\tau = \{x \mid \text{sign}(x) = \tau \wedge x \in A\}$ , with  $\tau \in \{-, +\}$ . From the definition of  $N$  it follows that every element of  $A$  lies within distance  $\varepsilon$  to a multiple of  $1/N$ . Suppose that  $m_1$  and  $m_2$  are the smallest natural numbers such that  $A \subseteq [p, q]$ , with  $p = -m_1/N$  and  $q = m_2/N$ . We then construct the sets  $I^+ = [0, m_2] \cap \mathbf{N}$ , and  $I^- = [0, m_1] \cap \mathbf{N}$ . We focus on the set  $I^+$ , the set  $I^-$  being analogous. Let  $x_{\max} = \max\{x \mid x \in A^+\}$ . Let  $\psi_2 = \psi_1$ , and  $\sigma_2(\varepsilon, T_2) = (\exists x)(\exists N)((0, x) \in T_2 \wedge |Nx_{\max} - x| \leq N\varepsilon \wedge \Psi_1(N, \varepsilon))$ . Then the set  $I^+$  is defined by the formula  $\Psi_2^+(n, A, \varepsilon) = [\text{TC}\psi_2 \upharpoonright_{T_2} \sigma_2](0, n)$ . (The set  $I^-$  is defined by  $\Psi_2^-$ .) Denote the union  $I^- \cup I^+$  with  $I$ , and let  $x \in A$ . The set  $I$  is then defined by  $\Psi_2 = \Psi_2^- \vee \Psi_2^+$ . We define the  $\varepsilon$ -approximation of  $x$  as either  $x$  itself, in case that  $x$  is a multiple of  $1/N$ , or as the unique multiple of  $1/N$  which is  $\varepsilon$  close to  $x$ , otherwise. We express the first case as  $\varphi_1(x, A, \varepsilon) = (x \in A) \wedge (\Psi_2(x, A, \varepsilon))$ , and the second case as  $\varphi_2(x, A, \varepsilon) = (x \in A) \wedge \neg(\Psi_2(x, A, \varepsilon))$ . A rational  $\varepsilon$ -approximation  $\hat{A}_\varepsilon$  is then defined by  $\{\hat{a} \mid (\hat{a}, A, \varepsilon) \models \Psi_3(x, A, \varepsilon)\}$ , with  $\Psi_3(\hat{x}, A, \varepsilon) = (\exists x)((\varphi_1(x, A, \varepsilon) \wedge \hat{x} = x) \vee (\varphi_2(x, A, \varepsilon) \wedge N\hat{x} \in I \wedge |\hat{x} - xN| \leq \varepsilon N))$ .  $\square$

**THEOREM 3.9.** *Rational  $\varepsilon$ -approximations of spatial databases are expressible in FO + POLY + TC $^\Sigma$ .*

**PROOF.** (Sketch)

- $m = 1$ : By Theorem 3.4, an  $\mathbf{A}$ -linear  $\varepsilon/2$ -approximation  $\hat{A}_{\varepsilon/2}$  of  $A$  can be defined in FO + POLY + TC. The singular points and corner points of  $\hat{A}_{\varepsilon/2}$  can be determined by an FO + POLY formula. The  $x$  and  $y$  coordinates of the singular and corner points of  $\hat{A}_{\varepsilon/2}$  form two finite sets of real algebraic numbers, which can be approximated by Lemma 3.8 by a rational  $\varepsilon/4$ -approximation. With these approximated singular points and corner points a  $\mathbf{Z}$ -linear relation can be built that approximates  $\hat{A}_{\varepsilon/2}$  with an  $\varepsilon/2$ -error margin and therefore satisfies the restrictions of the statement of this theorem.

- $m > 1$ : Here, we proceed as in the case  $m = 1$  and determine in the  $2^m$   $\mathbf{A}$ -linear approximations, provided by Theorem 3.4, the singular points and corner points and approximate these by rational numbers uniformly in all of the  $2^m$  sets. We follow the same construction as in the case  $m = 1$ .  $\square$

Analogously, we get the following result.

**THEOREM 3.10.** *There is an FO + POLY + TC $^\Sigma$  formula that defines, on a given spatial database  $(A_1, \dots, A_m)$ , a  $\mathbf{Z}$ -linear spatial database  $(B_1, \dots, B_m)$  for which there is an isotopy  $h$  of the plane such that  $h(A_i) = B_i$ , for  $1 \leq i \leq m$ .*

## 4. EXPRESSIVENESS RESULTS

In the previous section, we have proven that a specific query, namely the  $\varepsilon$ -approximation, is expressible in FO + POLY + TC. In this section, we show a general result on the expressiveness of FO + POLY + TC $^\Sigma$ . Specifically, we prove that FO + POLY + TC $^\Sigma$  is computationally complete on  $\mathbf{Z}$ -linear spatial databases, and computationally complete for

the Boolean topological queries on arbitrary databases. We prove this in several steps. First, we show that any Turing computable function on  $\mathbf{N}$  is expressible by an  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  formula. Secondly, we prove that there exists an  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  definable encoding of a  $\mathbf{Z}$ -linear spatial database by a natural number. The corresponding decoding is also in  $\text{FO} + \text{POLY}$ . This implies the completeness of  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  on  $\mathbf{Z}$ -linear spatial databases. Finally, the latter completeness result, together with the existence of an  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ -definable  $\mathbf{Z}$ -linear set (namely the one provided by Theorem 3.10), which is isotopic to a given arbitrary spatial database, shows that  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  can express all the computable Boolean topological queries on arbitrary spatial databases.

## 4.1 Recursive functions on the naturals

We now show that  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  is computationally complete on the set of natural numbers  $\mathbf{N}$ .

LEMMA 4.1. *For every  $\mu$ -recursive function  $f : \mathbf{N}^k \rightarrow \mathbf{N}$  there exists an  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  formula  $\varphi_f(y)$  over a single  $k$ -ary relation  $S$ , such that on any singleton input relation  $S = \{(n_1, \dots, n_k)\}$ , we have  $(S, y) \models \varphi_f$  iff  $y = f(n_1, \dots, n_k)$ .*

PROOF. Since the functions  $\text{zero}(x) := 0$ ,  $\text{proj}_i(x_1, \dots, x_k, y) := x_i$ , the successor relation  $\text{succ}(x, y)$ , and the composition of  $\mu$ -recursive functions, are straightforwardly expressible in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ , we focus on the formula of primitive recursion and the  $\mu$ -operator on  $\mu$ -recursive functions on  $\mathbf{N}$ .

Let  $f(x_1, \dots, x_k, 0) = g(x_1, \dots, x_k)$  and  $f(x_1, \dots, x_k, i + 1) = h(x_1, \dots, x_k, f(x_1, \dots, x_k, i), i)$ , with  $g$  and  $h$   $\mu$ -recursive functions. We denote the vector  $(x_1, \dots, x_k)$  with  $\vec{x}$ . We consider as singleton input relation  $\{(\vec{x}, i)\}$ , with all  $x_i$  and  $i$  natural numbers.

Let  $\psi_1(u, v) := (0 \leq u) \wedge (v = u + 1)$ , and  $\sigma_1(i, T) := (\exists j)(0, j) \in T \wedge j = i$ . The interval  $[0, i] \cap \mathbf{N}$  is then defined by  $\{x \mid [\text{TC}\psi_1 \mid_T \sigma_1](0, x)\}$ . Next, define  $\psi_2(j, y', k, y) := ((j \in [0, i] \cap \mathbf{N}) \wedge (j < i) \wedge (k = j + 1) \wedge \varphi_h(\vec{x}, y', j, y))$  and  $\sigma_2 := \text{false}$ . The  $\mu$ -recursive function  $f$  can now be expressed in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  as follows.

$$\varphi_f(\vec{x}, i, y) := (i = 0 \wedge \varphi_g(\vec{x}, y)) \vee (\exists y') (\varphi_g(\vec{x}, y') \wedge [\text{TC}^\Sigma \psi_2 \mid_T \sigma_2]((0, y'), (i, y))),$$

where  $\varphi_g$  and  $\varphi_h$  are  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  formulas for the functions  $g$  and  $h$ .

The  $\mu$ -operator on a  $\mu$ -recursive function  $g$ , denoted by  $f(\vec{x}) := \mu y g$ , is either defined as the smallest  $y$  such that  $g(\vec{x}, y) = 0$  and for each  $z < y$ ,  $g(\vec{x}, z)$  is defined and non-zero, or is undefined when such a  $y$  does not exist. We consider as singleton input relation  $\{\vec{x}\}$ , with all  $x_i$  natural numbers. Let  $\psi(y', y) := (0 \leq y') \wedge (y = y' + 1) \wedge (\exists z > 0) \varphi_g(\vec{x}, y, z)$  and  $\sigma(\vec{x}, T) := (\exists y)((0, y) \in T \wedge \varphi_g(\vec{x}, y, 0))$ . We can express the  $\mu$ -operator in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  as follows.

$$\varphi_f(\vec{x}, y) := (y = 0 \wedge \varphi_g(\vec{x}, y, 0)) \vee \max_y [\text{TC}^\Sigma \psi \mid_T \sigma](0, y),$$

with  $\varphi_g$  the  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  formula for the function  $g$ .  $\square$

## 4.2 Encoding/decoding of $\mathbf{z}$ -linear spatial databases

It has been shown by Vandeurzen et al. [31] that any  $\mathbf{Z}$ -linear spatial database has a finite geometric representation by means of a finite database over  $\mathbf{Q}$  consisting of  $(n + 1)^2$ -ary tuples, with  $n$  the dimension of the  $\mathbf{Z}$ -linear spatial database. Basically, this geometric representation contains the projective coordinates<sup>1</sup> of a complete triangulation of the linear spatial relations. Moreover, this representation can be computed by an  $\text{FO} + \text{POLY}$  formula  $\varphi_{\text{encode}}$ . Also, the corresponding decoding, which computes the  $\mathbf{Z}$ -linear spatial relation out of its finite geometric representation, can be performed using a  $\text{FO} + \text{POLY}$  formula  $\varphi_{\text{decode}}$  [31].

Using classical encoding techniques [10] (see also [14]), we now show how these finite rational databases can be encoded into a single natural number in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ . We also show that the corresponding decoding can be performed in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ .

Since the arity of the tuples in the finite representation of a  $\mathbf{Z}$ -linear spatial relation is 9, we first encode a projective coordinate triple  $q = (x, y, z)$  of rational coordinates into a 6-tuple of natural numbers  $p = ((d_x, n_x), (d_y, n_y), (d_z, n_z))$ , where  $d$  and  $n$  denote the denominator and nominator of the rational coordinates. Then we encode this  $6$ -tuple in the natural number  $n_p = 2^{d_x} \times 3^{n_x} \times 5^{d_y} \times 7^{n_y} \times 11^{d_z} \times 13^{n_z}$ , and then decode each tuple  $(p_1, \dots, p_9)$ , into the natural number  $2^{n_{p_1}} \times 3^{n_{p_2}} \times \dots \times 23^{n_{p_9}}$ . Hence, to finalize the encoding of a finite relation over the rationals is to encode a finite set  $R = \{a_1, \dots, a_\ell\}$  of integers as a single integer. We can assume that  $a_1 < a_2 < \dots < a_n$  (since the original point tuples were different). We will encode this set  $R$  as the natural number  $n = 2^{a_1} \times 3^{a_2} \times \dots \times p_\ell^{a_\ell}$ , with  $p_\ell$  the  $\ell$ th prime number.

This encoding can be expressed in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ . Indeed, first we generate successively pairs of natural numbers which enumerate the set of rational numbers  $\mathbf{Q}$ . We represent a rational number  $p/q$  by the pair  $(p, q)$ . This can be done by defining in  $\text{FO} + \text{POLY}$ , a successor relation on pairs of natural numbers. E.g., consider the interval  $[-1, 1] \cap \mathbf{Q}$ . The following successor relation is easily defined in  $\text{FO} + \text{POLY}$ :  $(0, 1) \mapsto (1, 1) \mapsto (-1, 1) \mapsto (0, 2) \mapsto (1, 2) \mapsto (-1, 2) \mapsto (0, 3) \mapsto (1, 3) \mapsto (-1, 3) \mapsto (2, 3) \mapsto \dots$ . Also outside  $[-1, 1]$ , an analogous successor relation can be defined in  $\text{FO} + \text{POLY}$ . We then stop this enumeration when for every rational number  $r$  in the projective coordinates of the representation of a  $\mathbf{Z}$ -linear spatial relation, a pair  $(p, q)$  is found such that  $q \times x = p$ . This clearly is in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ . Secondly, the cardinality  $\ell$  of  $R$  can be determined in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ . By completeness, the  $\ell$ -th prime number  $p_\ell$  can be computed in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ , and also the set  $A = \{2, 3, 4, \dots, p_\ell\}$ . The subset of  $A$  consisting of elements  $z$  such that there are no  $x, y \in A$ , such that  $z = x \times y$  can be defined. This gives the set of the first

<sup>1</sup>Projective coordinates are used to deal with unbounded databases and the unbounded triangles or lines in their triangulation.

$\ell$  prime numbers. In  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ , the set of pairs  $\{(p_1, a_1), \dots, (p_\ell, a_\ell)\}$  can be computed. Because of completeness, exponentiation  $x^y$  is expressible in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ , we only need to show how to multiply the elements in the set  $R' = \{p_1^{a_1}, \dots, p_\ell^{a_\ell}\}$ . First, define the binary relation  $M = \{((x, p_i^{a_i}), (x \times p_i^{a_i}, p_{i+1}^{a_{i+1}})) \mid x \in \mathbf{R} \text{ and } 1 \leq i < \ell\}$ . Then  $y \times p_\ell^{a_\ell}$  is the product of the elements in  $R'$ , where  $y$  is such that  $((1, p_1^{a_1}), (y, p_\ell^{a_\ell}))$  belongs to the transitive closure of  $M$ . To decode a natural number into a finite rational relation mainly requires the factoring of natural numbers into prime factors. This can also be done in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ . Indeed, given a natural number  $n$ , the set  $A = \{2, 3, 4, \dots, n\}$  can be computed in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ . From  $A$ , the divisors of  $n$ , the prime divisors of  $n$ , and the highest powers of these prime divisors that still divide  $n$  can be computed. Hence also the set of pairs  $(p_i, a_i)$  such that  $n = \prod p_i^{a_i}$  can be computed.

### 4.3 Completeness results

In this section, we conclude that  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  is computationally complete on  $\mathbf{Z}$ -linear spatial databases, and computationally complete for the Boolean topological queries on arbitrary databases.

**THEOREM 4.2.** *The language  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  is computationally complete on  $\mathbf{Z}$ -linear spatial databases.*

**PROOF.** This follows directly from the finite geometrical encoding/decoding of  $\mathbf{Z}$ -linear sets, presented in the previous section, and the fact that any Turing computable function on  $\mathbf{N}$  is expressible by an  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  formula (see Section 4.1).  $\square$

A Boolean query  $Q$  is said to be *topological*, if for any two spatial databases  $A = (A_1, \dots, A_m)$  and  $B = (B_1, \dots, B_m)$  for which there exists an isotopy  $h$  of the plane such that  $h(A_i) = B_i$ ,  $1 \leq i \leq m$ ,  $Q(A_1, \dots, A_m) = Q(B_1, \dots, B_m)$  holds.<sup>2</sup>

**THEOREM 4.3.** *All computable Boolean topological queries on spatial databases can be expressed in  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$ .*

**PROOF.** (Sketch) Let  $Q$  be a Boolean topological computable query. By Theorem 3.10, there exists an  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  formula  $Q_{\text{approx}}$  that defines, for any given spatial database  $(A_1, \dots, A_m)$ , a  $\mathbf{Z}$ -linear spatial database  $(B_1, \dots, B_m)$  for which there is an isotopy  $h$  of the plane such that  $h(A_i) = B_i$ , for  $1 \leq i \leq m$ . Since  $Q$  is computable, it is in particular computable on  $\mathbf{Z}$ -linear spatial databases, and therefore, by Theorem 4.2 expressible on these databases by an  $\text{FO} + \text{POLY} + \text{TC}^\Sigma$  formula  $Q_{\mathbf{Z}}$ . Now clearly,  $Q$  is equivalent to the composition of  $Q_{\mathbf{Z}}$  and  $Q_{\text{approx}}$ .  $\square$

There are also topological queries that can be expressed in  $\text{FO} + \text{POLY} + \text{TC}$  in a more straightforward manner, i.e.,

<sup>2</sup>In the literature, topological queries are also sometimes defined using homeomorphisms instead of isotopies. The difference is *not* essential.

avoiding the deviation via the encoding by natural numbers. As an illustration, we discuss the path-connectivity query. This expresses whether two given points  $p$  and  $q$  in a given spatial relation  $A$ , can be connected by a path lying entirely in  $A$ . This query can be expressed in  $\text{FO} + \text{POLY} + \text{TC}$  by first computing an  $\varepsilon$ -approximation of the spatial database  $(A, \{p, q\})$ . On this approximation,  $(\hat{A}, \{p', q'\})$ , a binary relation

$$\text{Path}(\hat{A}) = \{(r, s) \in \mathbf{R}^2 \times \mathbf{R}^2 \mid (\exists \lambda)(0 \leq \lambda \leq 1 \wedge (\forall t)(t = \lambda \cdot r + (1 - \lambda) \cdot s \rightarrow t \in \hat{A}))\}$$

can be defined in  $\text{FO} + \text{POLY}$  consisting of all pairs of points that can be connected by a straight line segment lying entirely in  $\hat{A}$ . Let  $\varphi_{\text{Path}}(\hat{A}, p', q')$  be the  $\text{FO} + \text{POLY} + \text{TC}$  formula that tests whether  $(p', q')$  belongs to the transitive closure of the relation  $\text{Path}(\hat{A})$ .

Using techniques similar to those used to prove correctness of a spatial Datalog program [20, 22] to test connectivity of spatial databases, the following can be shown.

**PROPOSITION 4.4.** *The  $\text{FO} + \text{POLY} + \text{TC}$  formula  $\varphi_{\text{Path}}(\hat{A}, p', q')$  correctly tests whether there exists a path between two points  $p'$  and  $q'$  of an  $\mathbf{A}$ -linear spatial relation  $\hat{A}$ , in particular, the evaluation of this formula always terminates on these inputs.*

## 5. ACKNOWLEDGEMENTS

The authors are grateful to Jan Van den Bussche and Leonid Libkin for helpful comments on issues of completeness.

## 6. REFERENCES

- [1] D. Abel and B. Ooi, editors. *Advances in Spatial Databases—3rd Symposium SSD'93*, volume 692 of *Lecture Notes in Computer Science*. Springer-Verlag, 1993.
- [2] R. Benedetti and J.-J. Risler. *Real Algebraic and Semi-Algebraic Sets*. Actualités Mathématiques. Hermann, 1990.
- [3] M. Benedikt and L. Libkin. Safe constraint queries. In *Proceedings of the 17th ACM Symposium on Principles of Database Systems (PODS'98)*, pages 99–108. ACM Press, 1998.
- [4] M. Benedikt and L. Libkin. Exact and approximate aggregation in constraint query languages. In *Proceedings of the 18th ACM Symposium on Principles of Database Systems (PODS'99)*, pages 102–113. ACM Press, 1999.
- [5] J. Bochnak, M. Coste, and M.-F. Roy. *Real Algebraic Geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. Folge 3*. Springer-Verlag, 1998.
- [6] A. Buchmann, editor. *Design and implementation of large spatial databases—First Symposium SSD'89*, volume 409 of *Lecture Notes in Computer Science*. Springer-Verlag, 1989.
- [7] G. Collins. Quantifier elimination for real closed fields by cylindrical algebraic decomposition. In *Lecture*

- Notes in Computer Science*, volume 33, pages 134–183, 1975.
- [8] F. Dumortier, M. Gyssens, L. Vandeurzen, and D. Van Gucht. On the decidability of semi-linearity for semi-algebraic sets and its implications for spatial databases. In *Proceedings of the 16th ACM Symposium on Principles of Database Systems (PODS'97)*, pages 68–77. ACM Press, 1997.
- [9] M. Egenhofer and J. Herring, editors. *Advances in Spatial Databases*, volume 951 of *Lecture Notes in Computer Science*. Springer, 1995.
- [10] H. Enderton. *A Mathematical Introduction to Logic*. Academic Press, 1972.
- [11] F. Geerts and B. Kuijpers. Expressing topological connectivity of spatial databases. In A. Mendelzon and R. Connor, editors, *Proceedings of Database Programming Languages (DBPL'99)*, Lecture Notes in Computer Science. Springer-Verlag, to appear, 2000.
- [12] S. Grumbach, P. Rigaux, M. Scholl, and L. Segoufin. DEDALE, a spatial constraint database. In S. Cluet and R. Hull, editors, *Proceedings of Database Programming Languages (DBPL'97)*, volume 1369 of *Lecture Notes in Computer Science*, pages 38–59. Springer-Verlag, 1998.
- [13] S. Grumbach, P. Rigaux, and L. Segoufin. The DEDALE system for complex spatial queries. In L. Haas and A. Tiwary, editors, *Proceedings of the ACM International Conference on Management of Data (SIGMOD'98)*, pages 213–224. ACM Press, 1998.
- [14] S. Grumbach and J. Su. First-order definability over constraint databases. In U. Montanari and F. Rossi, editors, *Principles and Practice of Constraint Programming. First International Conference, CP'95*, volume 976 of *Lecture Notes in Computer Science*, pages 121–136. Springer-Verlag, 1995.
- [15] S. Grumbach and J. Su. Towards practical constraint databases. In *Proceedings of the 15th ACM Symposium on Principles of Database Systems (PODS'96)*, pages 28–39. ACM Press, 1996.
- [16] S. Grumbach and J. Su. Queries with arithmetical constraints. *Theoretical Computer Science*, 173(1):151–181, 1997.
- [17] O. Gunther and H.-J. Schek, editors. *Advances in Spatial Databases—2nd Symposium SSD'91*, volume 525 of *Lecture Notes in Computer Science*. Springer-Verlag, 1991.
- [18] R. Güting, editor. *Advances in Spatial Databases—6th Symposium SSD'99*, volume 1651 of *Lecture Notes in Computer Science*. Springer-Verlag, 1999.
- [19] P. Kanellakis, G. Kuper, and P. Revesz. Constraint query languages. *Journal of Computer and System Science*, 51(1):26–52, 1995. A preliminary report appeared in the *Proceedings of the 9th ACM Symposium on Principles of Database Systems (PODS'90)*.
- [20] 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, International Workshop (LID'96)*, volume 1154 of *Lecture Notes in Computer Science*, pages 101–116. Springer, 1996.
- [21] B. Kuijpers, J. Paredaens, and J. Van den Bussche. Topological elementary equivalence of closed semi-algebraic sets in the real plane. In F.N. Afrati and P. Kolaitis, editors, *Database Theory (ICDT'97)*, volume 1186 of *Lecture Notes in Computer Science*, pages 432–446. Springer, 1996. to appear in *Journal of Symbolic Logic*.
- [22] B. Kuijpers and M. Smits. On expressing topological connectivity in spatial datalog. In V. Gaede, A. Brodsky, O. Günther, D. Srivastava, V. Vianu, and M. Wallace, editors, *Constraint Databases and Their Applications, 2nd International Workshop on Constraint Database Systems (CDB '97)*, volume 1191 of *Lecture Notes in Computer Science*, pages 116–133. Springer, 1997.
- [23] B. Kuijpers and V. Vianu. Topological queries. In G. Kuper, J. Paredaens, and L. Libkin, editors, *Constraint Databases*, pages 229–271. Springer-Verlag, 1999.
- [24] G. Kuper, J. Paredaens, and L. Libkin, editors. *Constraint Databases*. Springer-Verlag, 1999.
- [25] E. Moise. *Geometrical Topology in Dimensions 2 and 3*. Springer-Verlag, 1977.
- [26] M. Preparata and M. Shamos. *Computational Geometry*. Springer-Verlag, 1985.
- [27] M. Scholl and A. Voisard, editors. *Advances in Spatial Databases—5th symposium SSD'97*, volume 1262 of *Lecture Notes in Computer Science*. Springer-Verlag, 1997.
- [28] A. Tarski. *A Decision Method for Elementary Algebra and Geometry*. University of California Press, 1951.
- [29] L. Vandeurzen. *Logic-Based Query Languages for the Linear Constraint Database Model*. PhD thesis, Limburgs Universitair Centrum (LUC), 1999.
- [30] L. Vandeurzen, M. Gyssens, and D. V. Gucht. On the desirability and limitations of linear spatial query languages. In Egenhofer and Herring [9], pages 14–28.
- [31] L. Vandeurzen, M. Gyssens, and D. Van Gucht. An expressive language for linear spatial database queries. In *Proceedings of the 17th ACM Symposium on Principles of Database Systems (PODS'98)*, pages 109–118. ACM Press, 1998.