The expressive power of first-order topological languages

Ernest Davis[*]
Dept. of Computer Science
New York University
davise@cs.nyu.edu

November 29, 2012

Abstract

We consider the expressive power of the first-order structure $\langle \Omega, \mathtt{C} \rangle$ where Ω is either of two of different domains of extended regions in Euclidean space, and $\mathtt{C(x,y)}$ is the topological relation "Region $\mathtt{x}$ is in contact with region $\mathtt{y}$." We prove two main theorems:

- Let $\mathcal{P}[\mathbb{Q}]$ be the domain of bounded, non-empty, rational polyhedra in two- or three-dimensional Euclidean space. A relation Γ over $\mathcal{P}[\mathbb{Q}]$ is definable in the structure $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$ if and only if Γ is arithmetic and invariant under rational PL-homeomorphisms of the space to itself. We also extend this result to a number of other domains, including the domain of all polyhedra and the domain of semi-algebraic regions.

- Let $\mathcal{R}$ be the space of bounded, non-empty, closed regular regions. in n-dimensional Euclidean space. Any analytical relation over lower-dimensional (i.e. empty interior) compact point sets that is invariant under homeomorphism is implicitly definable in the structure $\langle \mathcal{R}, \mathtt{C} \rangle$.

Keywords: First-order language, topological language, qualitative spatial reasoning, expressivity.

One much-studied approach to qualitative spatial representation is to use a logical structure where the domain of entities is some natural collection of extended spatial regions, and where the primitive relations are qualitative relations over regions. Such structures vary along three primary dimensions:

- The underlying logic. For example, constraint logic, existential logic, or first-order logic.

- The choice of domain of regions. For example, rational polyhedra, real polyhedra, semi-algebraic regions, subanalytic regions, topologically regular regions.

- The choice of relations. Here there is a wide range of choices, which can be categorized in terms of the underlying geometry. For instance, there are *set theoretic* or *mereological* properties, such as $\mathtt{PartOf(x,y)}$. There are *topological* properties such as $\mathtt{ExternallyConnnected(x,y)}$ or $\mathtt{Connected(x)}$. There are *affine* properties, such as $\mathtt{Convex(x)}$. There are *metric* properties, such as $\mathtt{Sphere(x)}$ or $\mathtt{Closer(x,y,z)}$. And, of course, combinations of these can be used.

For each of these structures, there are a number of metalogical questions that can be asked. What is the computational complexity of the decision problem for sentences in the structure? What is the expressivity of the structure; i.e. what features or relations can be expressed as open formulas?

Can the structure be axiomatized? Are two structures equivalent? Can one structure be defined in terms of another?

In this paper, we address two closely related questions of this kind. In both, the metalogical issue we address is expressivity; the logic is first-order logic; the sole relation is $C(x,y)$, meaning "Region x is in contact with y". We consider two different domains of regions: the domain of rational polyhedra, denoted $\mathcal{P}[\mathbb{Q}]$ and the domain of all bounded, topologically closed regular regions, denoted $\mathcal{R}$. (These are formally defined in definitions 1 and 4 below.)

We characterize the expressivity of these two structures with reference to the first- and second-order theories of the natural numbers. Properties of natural numbers definable in the first-order theory of the natural numbers, with addition and multiplication, are known as *arithmetic* properties. Properties of natural numbers and of sets of natural numbers definable in the second-order theory of the natural numbers are known as *analytical* properties. The vast majority of all properties of Euclidean geometry or real analysis that are discussed in the mathematical literature can be expressed quite straightforwardly as analytical properties. These two categories have natural extensions to the domains of polyhedra and of regular regions, as we will discuss below (definitions 17 and 25).

We will prove the following two main theorems:

> **Theorem 1**: Let $\mathcal{P}[\mathbb{Q}]$ be the domain of rational polyhedra in two- or three-dimensional Euclidean space. A relation Γ over $\mathcal{P}[\mathbb{Q}]$ is definable in the structure $\langle \mathcal{P}[\mathbb{Q}], C \rangle$ if and only if Γ is arithmetic and invariant under rational PL-homeomorphisms of the space to itself. We also extend this result to a number of other domains, including the domain of all polyhedra and the domain of semi-algebraic regions.

> **Theorem 3**: Any analytical relation over lower-dimensional compact point sets (i.e. point sets with empty interior) that is invariant under homeomorphism is implicitly definable in the structure $\langle \mathcal{R}, C \rangle$.

The two proofs have a similar flavor and share some lemmas, which is the reason for including them together in the same paper. Both proofs rely on *labelling* the geometric entities with numerical values. In the structure $\langle \mathcal{P}[\mathbb{Q}], C \rangle$, it is possible to construct natural numbers and finite sequences of natural numbers as first-order constructions; these can be used to label the vertices of a simplicial complex and then to characterize the combinatorial structure of the complex as a tuple of numerical labels. In the structure $\langle \mathcal{R}, C \rangle$, it is possible define real numbers as first-order constructions; these can be used to label each of a countable, dense subset of points in a lower-dimensional entity with its coordinates in a standard embedding.

The paper is structured as follows. Section 1 discusses related work. Section 2 sets up some groundwork, which applies more or less equally in the two structures: how basic topological relations can be defined (section 2.1); how lower-dimensional entities and basic relations over lower-dimensional entities can be defined (section 2.2); and how natural numbers, finite tuples of natural numbers, arithmetic relations, real numbers, countable sequences of reals, and analytical relations can be defined (section 2.3). Section 3 gives the proof of theorem 1. Section 4 gives a "warm-up" to theorem 3 by proving a very special case (theorem 2): The property "Point set S is homeomorphic to the k-cube," where k is less than the dimensionality of the embedding space, is definable in $\langle \mathcal{R}, C \rangle$. The essential construction is the same here as in theorem 3, but the more restricted setting makes the proof easier to follow. Section 5 gives the proof of theorem 3. Section 6 shows that the results here and in [9] partially answer a question raised by Hahmann and Gruninger in [15]. Section 7 concludes.

1 Related Work

In this section we will survey only computability and expressivity results for first-order theories. The literature analyzing constraint and existential languages is much larger; see [7] and [21] for recent surveys.

The study of geometric theories based around on extended regions began with de Laguna [10] and Whitehead [30] and was greatly advanced by Clarke [6]. Gerla [12] presents a history of such theories. Borgo and Masolo [5] carry out a systematic comparison of the semantic models.

Tarski [28] studied a language over a universe of regions with the two primitives C(x,y) and S(x) (x is a sphere); he was not specific as to the logic he has in mind, but his exposition includes second-order constructions (e.g. the definition of a point in terms of a set of concentric spheres.) This theory was made more formal by Bennett [1], who developed first- and second-order axiomatizations of this theory and proves that the axioms are categorical.

Grzegorczyk [14] proved that the first-order language over regions in the plane with the predicate C(x,y) is undecidable. A similar result was rediscovered independently, with a very different proof, by Dornheim [11]. Grzegorczyk encodes an integer k as a region with k connected components; we will use this encoding throughout this paper. The theorem holds even if the universe of regions is restricted to polygons (indeed, even if it is restricted to polygons with horizontal and vertical edges.)

Randell, Cui, and Cohn [25] introduced the region connection calculus (RCC), and demonstrated that the RCC primitives and other important topological properties can be defined in the first-order logic over regular regions with the predicate C(x,y). The expressivity of this language was further studied by Gotts [13].

Pratt-Hartmann and Schoop [23] studied the expressivity of the language over the space of polygonal regions in the plane with the primitive C(x,y) and the language over the same universe with the primitives Conn(x) and P(x,y). They prove that both languages are strong enough to specify completely the topological relations among any n-tuple of regions. (More precisely, the first language can achieve this over the open plane; the second can achieve it only for regions within the closed plane.) Pratt-Hartmann [24] achieved the analogous result for polyhedral regions in three-space; we rely on this theorem and on some other lemmas from this paper in section 3 below (lemmas 14 and 15). In large measure, indeed, section 3 can be viewed as essentially an extension to this part of [24], achieving a biconditional by incorporating Grzegorczyk's construction of the natural numbers. [24] also proves numerous other metalogical theorems for this theory and related theories.

Randell, Cui, and Cohn [25] also studied the first order language with predicates C(x,y) and Convex(x), and presented some useful qualitative properties that could be defined in that language. Pratt [22] proved that this language is powerful enough to express real addition and multiplication, and demonstrated a number of meta-level properties of this and related languages. Davis [8] proved that, over a broad range of domains of regions, this language can express any analytical property that is invariant under affine transformations, and that the language with the single primitive Closer(x,y,z) (Region x is closer to y than to z) could express any analytical property that is invariant under similarity transformations.

Bennett et al. [2] studied the first-order language over regular regions with predicates P(x,y) and Sphere(x). (It is easily shown that this is mutually definable with the structure $\langle \mathcal{U},$Closer$\rangle$ as long as the universe $\mathcal{U}$ contains spheres.) In a further paper [3] they studied the application of this language to representing rigid body motion. Borgo et al. [4] considered the first order language over regions with the single predicate P(x,y); the two predicates P(x,y) and SC(x) (x is strongly connected); and the three predicates P(x,y), SC(x) and Sphere(x).

A very thorough survey of the literature on first-order languages over regions, together with original

strong results on metalogical properties of first-order languages over regions in one-dimensional space and of existential languages in dimensions 2 and higher can be found in [21].

2   Preliminaries

Throughout we will be dealing with regions in n-dimensional Euclidean space $\mathbb{E}^n$ for fixed $n \geq 2$. ($\mathbb{E}^1$ is different; see [21] for a discussion.) The symbol n will *always* refer to the dimension of the space.

We begin with formal definitions of the two domains of regions that are the primary focus of this paper.

Definition 1 *A subset $\mathbf{R}$ of $\mathbb{E}^n$ is* topologically closed regular *if it is equal to the closure of its interior. For the remainder of this paper we will simply say "regular" to mean "topologically closed regular". The symbol $\mathcal{R}$ will denote the class of non-empty, bounded, regular subsets of $\mathbb{E}^n$.*

Since we will extensively be using the theory of simplices in section 3, we will use that as the basis of our definition of polyhedra.

Definition 2 *Let $0 \leq k \leq n$. Let $V = \{\mathbf{v}^1 \ldots \mathbf{v}^{k+1}\}$ be a set of $k+1$ points in $\mathbb{E}^n$ that do not lie in any affine space of dimension $k-1$. Then the* (closed) k-dimensional Euclidean simplex *with vertex set V is the convex hull of V.*

If V is a set of $k+1$ points as above and Q is a non-empty subset of V then the convex hull of Q is a face *of V.*

Definition 3 *A* Euclidean complex *is a finite set $\mathcal{C}$ of Euclidean simplices satisfying the following:*

- *if $U \in \mathcal{C}$ and V is a face of U then $V \in \mathcal{C}$.*

- *if $U \in \mathcal{C}$ and $V \in \mathcal{C}$ then either $U \cap V$ is empty or $U \cap V$ is both a face of U and a face of V.*

The extent *of $\mathcal{C}$, denoted $|\mathcal{C}|$, is the union of all the elements of $\mathcal{C}$: $|\mathcal{C}| = \cup_{U \in \mathcal{C}} U$. A* piece *of $\mathcal{C}$ is any maximal element of $\mathcal{C}$; that is, any element that is not a face of some other element. The set of vertices of $\mathcal{C}$ is denoted* Vertices($\mathcal{C}$).

Definition 4 *A point set S is a* polytope *if $S = |\mathcal{C}|$ for some Euclidean complex $\mathcal{C}$. A polytope is* fully dimensional *if S is the union of n-dimensional simplices; equivalently, if S is topologically closed regular. A fully dimensional polytope is called a* polyhedron. *A polytope is* rational *if the coordinates of all its vertices are rational. The set of all rational polyhedra is denoted $\mathcal{P}[\mathbb{Q}]$.*

There are two arbitrary choices in the above domains. The first is the decision to use closed regular regions rather than open regular regions (i.e. regions equal to the interior of their closure); this is entirely arbitrary since the domain of closed regular regions and the domain of open regular regions are isomorphic. The second is to restrict the domain to bounded regions and exclude unbounded regions. This is largely a matter of taste; some researchers, most notably Pratt-Hartmann, prefer to use domains of unbounded regions [24]. The theorems in this paper remain valid if the domains are defined to include unbounded regions and the proofs require only very slight modification. In an analysis that relies on simplicial decomposition of polyhedra, it is slightly simpler to use bounded polyhedra.

In this paper, we deal primarily with the two domains $\mathcal{P}[\mathbb{Q}]$, the class of bounded, non-empty rational polyhedra in $\mathbb{E}^n$, and $\mathcal{R}$, the class of all bounded, non-empty, regular regions in $\mathbb{E}^n$. Some further domains of regions will be considered in section 3.5. Almost all of the discussion in this section will apply to both domains; we will use the symbol Ω to mean either domain. Throughout this paper a *region* will be an element of whichever domain is under discussion; other subsets of $\mathbb{E}^n$ will be called "point sets".

The fundamental topological relation we will consider in this paper is $\texttt{C(x,y)}$ meaning "x and y are in contact"; that is, they have at least one point in common.

Notation: We will use boldface lower case symbols like $\mathbf{p}$ for points; boldface upper case symbols like $\mathbf{R}$ for regions; and italicized upper-case symbols like S for other point sets and other entities. We will use calligraphic font like $\mathcal{T}$ for sets of point sets, both simplicial complexes and domains of regions.

The symbol $\mathbb{R}^k$ means the space of all k-tuples of real numbers; in section 5 especially, it will be important to distinguish this numeric space from the (isomorphic, but conceptually different) geometric space $\mathbb{E}^n$. We will use vector notation $\vec{v}$ for vectors (i.e. k-tuples) in $\mathbb{R}^k$. Other sequences, finite or infinite, will be designated using hat notation $\hat{c}$. We will coerce sequences to the corresponding sets; that is, we will use expressions like $\vec{x} \in \hat{c}$ to mean "there is a i such that $\vec{x} = \hat{c}_i$." We assume throughout that there is a standard coordinate system $e : \mathbb{E}^n \mapsto \mathbb{R}^n$ mapping points to vectors; note that such an assumption is inherent in the very concept of "rational polyhedra".

The symbol $P \setminus Q$ is the set difference P minus Q (*not* the regularized difference). S° denotes the interior of S. Both S^- and $\mathrm{Closure}(S)$ denote the closure of S relative to the space $\mathbb{E}^n$. The boundary of S relative to $\mathbb{E}^n$ is denoted $\mathrm{Bd}(S) = S^- \setminus S^\circ$ (*not* the boundary of S in the sense of a manifold.) The function $d(\mathbf{p}, \mathbf{q})$ is the Euclidean distance. The open ball centered at $\mathbf{p}$ of radius r is denoted $B(\mathbf{p}, r)$ and the closed ball is denoted $\bar{B}(\mathbf{p}, r)$.

The symbol $\mathbb{C}^k$ will denote the standard k-dimensional cube, $[0, 1]^k$. The symbol $\mathbb{F}^k$ will denote the standard embedding of the k-cube in $\mathbb{E}^n$, $\mathbb{C}^k \times \{0\}^{n-k}$.

We will use typewriter font like $\texttt{C(x,y)}$ for expressions in the formal languages, with lower-case letters for variables and upper-case letters for constant, function, and predicate symbols. We will sometimes be imprecise about the distinction between the formal symbol and the entity it denotes; in view of the logical simplicity of what we are doing, that should cause no distress.

Slightly at variance with standard terminology, we will say that regular region $\mathbf{R}$ is a "neighborhood" of point $\mathbf{p}$ if $\mathbf{p} \in \mathbf{R}^\circ$.

2.1 Basic Topological Relations

Let Ω be either the domain $\mathcal{P}[\mathbb{Q}]$ or the domain $\mathcal{R}$.

We begin by defining some basic topological predicates, including the standard RCC relations in terms of the predicate, $\texttt{C(x,y)}$, x is in contact with y. These definitions have been known at least since [25]. In all of the following, the variables range over regions in Ω.

```
P(x,y)  ≡ ∀_z C(z,x) ⇒ C(z,y).
PP(x,y) ≡ P(x,y) ∧ ¬P(y,x).
DC(x,y) ≡ ¬C(x,y).
O(x,y)  ≡ ∃_z P(z,x) ∧ P(z,y).
DS(x,y) ≡ ¬O(x,y).
```

```
EC(x,y) ≡ C(x,y) ∧ DS(x,y).

OV (x,y) ≡ O(x,y) ∧ ¬P(x,y) ∧ ¬P(y,x).

TPP(x,y) ≡ PP(x,y) ∧ ∃_z EC(z,x) ∧ EC(z,y).

NTPP(x,y) ≡ PP(x,y) ∧ ¬TPP(x,y).

EQ(x,y) ≡ P(x,y) ∧ P(y,x).

Union(x,y,z) ≡ ∀_a C(a,z) ⇔ C(a,x) ∨ C(a,y).

Intersect(x,y,z) ≡ ∀_a P(a,z) ⇔ P(a,x) ∧ P(a,y).

Dconn(x) ≡ ∃_{a,b} Union(a,b,x) ∧ DC(a,b).

Conn(x) ≡ ¬Dconn(x).

InteriorConnected(x) ≡
      ∀_w NTPP(w,x) ⇒ ∃_z PP(w,z) ∧ NTPP(z,x) ∧ Conn(z).

IntConnComp(x,y) ≡
      P(x,y) ∧ InteriorConnected(x) ∧
      [∀_z P(x,z) ∧ InteriorConnected(z) ⇒ z= x].
```

It is well known, and easily proved, that over either domain $\mathcal{R}$ or $\mathcal{P}[\mathbb{Q}]$, the above definitions do indeed define the intended relations.

Lemma 1 *Consider the structure $\langle \Omega, \mathtt{C} \rangle$ where Ω may be either $\mathcal{R}$ or $\mathcal{P}[\mathbb{Q}]$, and $\mathtt{C(x,y)}$ holds if $\mathtt{x} \cap \mathtt{y}$ is non-empty. Then the relations defined in the above formal definitions satisfy the following:*

P(x,y) *iff* x *is a subset of* y.
PP(x,y) *iff* x *is a proper subset of* y.
DC(x,y) *iff* x *is disconnected from* y; *that is,* $\mathtt{x} \cap \mathtt{y} = \emptyset$.
O(x,y) *iff* x *and* y *overlap; that is* $\mathtt{x} \cap \mathtt{y}$ *has non-empty interior.*
DS(x,y) *iff* x *is disjoint from* y; *that is, the interior of* $\mathtt{x} \cap \mathtt{y}$ *is empty.*
EC(x,y) *iff* x *is externally connected to* y.
OV(x,y) *iff* x *and* y *properly overlap; that is, they overlap but neither is a subset of the other.*
TPP(x,y) *iff* x *is a tangential proper part of* y. *That is,* x *is a proper subset of* y
      and the boundary of x *meets the boundary of* y.
NTPP(x,y) *iff* x *is a non-tangential partial part of* y. *That is,* x *is a proper subset of* y
      and the boundary of x *does not meet the boundary of* y.
EQ(x,y) *iff* x *and* y *are equal.*
Union(x,y,z) *iff* z *is the set union of* x *and* y. Intersect(x,y,z) *iff* z *is the regularized intersection of* x *and* y.
Dconn(x) *iff* x *is a disconnected region.*
Conn(x) *iff* x *is a connected region.*
InteriorConnected(x) *iff the interior of* x *is connected.*
IntConnComp(x,y) *iff* x *is an interior connected component of* y;
      i.e. the closure of a maximal connected open subset of y.

Proof: These are well-known and easily proved [24].

2.2   Lower Dimensional Entities

Definition 5 *A point set S is* lower-dimensional *if the interior of S is the empty set.*

We characterize a lower-dimensional entity S in terms of an externally-connected pair $\mathbf{A}, \mathbf{B} \in \Omega$ such that $S = \mathbf{A} \cap \mathbf{B}$. Such a point-set will be called an "iregion" (for intersection region); the collection of iregions will be denoted Ω'.

In these definitions, we will use semi-colons to group together the pairs of region arguments that are defining an iregion or some other entity such as a number. This is purely for readability; logically, the semi-colon is equivalent to a comma.

We now show that we can define the following basic relations on iregions.

> `Point(a,b)`: Iregion $\mathtt{a} \cap \mathtt{b}$ consists of a single point.
> `PtInReg(a,b;r)`: $\mathtt{a} \cap \mathtt{b}$ is a point and is in region $\mathtt{r}$.
> `PtIn(a,b;c,d)`: $\mathtt{a} \cap \mathtt{b}$ is a point and is in iregion $\mathtt{c} \cap \mathtt{d}$.
> `SamePoint(a,b;c,d)`: Iregions $\mathtt{a} \cap \mathtt{b}$ and $\mathtt{c} \cap \mathtt{d}$ consist of the same single point.
> `Subset(a,b;c,d)`: $(\mathtt{a} \cap \mathtt{b}) \subset (\mathtt{c} \cap \mathtt{d})$.
> `SameIregion(a,b;c,d)`: $\mathtt{a} \cap \mathtt{b} = \mathtt{c} \cap \mathtt{d}$.
> `Region(a,b)`: Iregion $\mathtt{a} \cap \mathtt{b}$ is in fact a region.

> `Point(a,b)` $\equiv$
> `EC(a,b)` $\wedge$ `[`$\forall_{\mathtt{x,y}}$ `PP(x,a)` $\wedge$ `PP(y,b)` $\wedge$ `C(x,b)` $\wedge$ `C(y,a)` $\Rightarrow$ `C(x,y)]`.
>
> `PtInReg(a,b;r)` $\equiv$
> `Point(a,b)` $\wedge$ $\forall_{\mathtt{x}}$ `PP(x,a)` $\wedge$ `EC(x,b)` $\Rightarrow$ `C(x,r)`.
>
> `PtIn(a,b;c,d)` $\equiv$ `PtInReg(a,b;c)` $\wedge$ `PtInReg(a,b;d)`.
>
> `SamePoint(a,b;c,d)` $\equiv$ `PtIn(a,b;c,d)` $\wedge$ `PtIn(c,d;a,b)`.
>
> `Subset(a,b;c,d)` $\equiv$ $\forall_{\mathtt{x,y}}$ `PtIn(x,y;a,b)` $\Rightarrow$ `PtIn(x,y;c,d)`.
>
> `SameIregion(a,b;c,d)` $\equiv$ `Subset(a,b;c,d)` $\wedge$ `Subset(c,d;a,b)`.
>
> `Region(a,b)` $\equiv$ $\exists_{\mathtt{r}}$ `SameIregion(a,b;r,r)`.

Note that any region $\mathtt{r} = \mathtt{r} \cap \mathtt{r}$, so a region is also an iregion.

In $\mathcal{P}[\mathbb{Q}]$, a particular point $\mathtt{p}$ can be indicated using the predicate `Point(a,b)` only if $\mathtt{p}$ has rational coordinates.

The correctness of the above formal definition of `Point` as a formula over $\mathcal{R}$ can be proved as follows:

Lemma 2 *Over the domain $\mathcal{R}$, two regions $\mathbf{A}, \mathbf{B}$ meet at a single point if and only if they satisfy the definition of* `Point` *given above.*

Proof: Left to right. Suppose that $\mathbf{A}$ and $\mathbf{B}$ meet only at point $\mathbf{p}$. Then they are certainly `EC`. Let $\mathbf{X}, \mathbf{Y}$ be regions such that `PP(X,A)`, `C(X,B)`, `PP(Y,B)`, `C(Y,A)`. Since `C(X,B)`, $\mathbf{X}$ must have a point in common with $\mathbf{B}$ and since `PP(X,A)` that point can only be $\mathbf{p}$. Likewise $\mathbf{p} \in \mathbf{Y}$, so `C(X,Y)`.

Right to left. Suppose that $\mathbf{A}$ and $\mathbf{B}$ are `EC` and meet in two points $\mathbf{p}$ and $\mathbf{q}$. Let $\mathbf{U}$ and $\mathbf{V}$ be disconnected regions such that $\mathbf{p} \in \mathbf{U}^\circ$ and $\mathbf{q} \in \mathbf{V}^\circ$. Let $\mathbf{X}$ and $\mathbf{Y}$ be the regularized intersections of $\mathbf{U}$ with $\mathbf{A}$ and of $\mathbf{V}$ with $\mathbf{B}$ respectively. It is easily shown that $\mathbf{X}$ and $\mathbf{Y}$ are non-empty and that they satisfy the conditions of the definition. ∎.

The proof of the correctness of the other definitions is entirely straightforward.

We can therefore, without affecting the expressivity, extend our domain to the domain of iregions with an equality relation, all the predicates over regions defined in section 2.1 plus the predicates

`Point(x)`, `Region(x)`, `Subset(x,y)`, and `PtIn(p,x)`. with the obvious interpretations. That is, any sentence ϕ in the new structure with variables ranging over iregions can be translated into a sentence θ in the structure $\langle \Omega, \mathtt{C} \rangle$, replacing every iregion variable in ϕ by a pair of region variables, and replacing each predicate in ϕ by its definition as shown above.

2.3 Arithmetic and analytical relations

In this section we review the definitions of the classes of arithmetic relations and analytical relations and we discuss how other structures can be defined in terms of them. (These definitions are adapted from [16]; see especially section 3.5, pp. 114-124. The proofs in this section are presented in greater detail in [21], section 4.3.3).

The symbol $\mathbb{N}$ represents the set of natural numbers. $\mathbb{N}^*$ (Kleene star) is the set of finite sequences of natural numbers. $\mathbb{N}^{\mathbb{N}}$ is the set of all countable sequences of natural numbers.

Definition 6 *A relation over $\mathbb{N}$ is* arithmetic *if it can be defined as an open formula in the structure $\langle \mathbb{N}, +, * \rangle$ (the natural numbers with addition and multiplication).*

Any computable function can be expressed in the language over the natural numbers with $+$ and $*$ [16]; hence the class of arithmetic relations includes all computable functions.

Using Gödel numbering, it is possible to define a function `TupleToNat` from $\mathbb{N}^*$ to $\mathbb{N}$ that maps each finite sequence $\hat{x} \in \mathbb{N}^*$ to a unique value. Basic operations on $\hat{x}$ such as extracting the ith element are then definable as arithmetic relations. We can therefore extend the notion of an arithmetic relation to be any relation definable in the structure $\langle \mathbb{N} \cup \mathbb{N}^*, +, *, [\,], \mathtt{TupleToNat} \rangle$ where $\hat{x}[i]$ is the indexing function returning the ith element of tuple $\hat{x}$. Iterating the operation, we can encode finite tuples of tuples, finite sets of tuples, and so on as natural numbers; we will use the symbol $\mathbb{N}^*$ loosely to include any such entity.

Definition 7 *Let $\mathcal{U}$ be the domain $\mathbb{N} \cup \mathbb{N}^{\mathbb{N}}$ and let $[\,]$ be the indexing function, as above. A relation over $\mathcal{U}$ is* analytical *if it is definable in the structure $\langle \mathcal{U}, +, *, [\,] \rangle$.*

A real number can be represented as a sequence of integers (the bits in its binary expansion); a countable sequence of real numbers can likewise be represented as a sequence of integers, by dovetailing. It is easily shown that real addition, real multiplication, and indexing into the sequence of real numbers correspond to analytical relations over the corresponding sequence of integers. See [8] or [21] for details, though the result is certainly not original there. We will therefore use all these entities and relations freely in our constructions below.

A positive natural number n will be encoded in our geometric language as a region with n interior-connected components. This encoding and the definitions of two regions representing the same number and of addition were given by Grzegorczyk [14]. Our definition of multiplication is different from Grzegorczyk's, but either will do.

We must begin by defining the property of having finitely many interior-connected components, and thus being a representative of a proper natural number. In the domain $\mathcal{P}[\mathbb{Q}]$, this goes without saying, because all polyhedra have finitely many interior-connected components, but in $\mathcal{R}$ it must be dealt with. We will impose the further condition that a region $\mathbf{R}$ used to represent a natural number must have "separated interior-connected components"; that is, if $\mathbf{A}$ and $\mathbf{B}$ are interior-connected components of $\mathbf{R}$, then $\mathbf{A}$ and $\mathbf{B}$ have no boundary points in common. Suppose that $\mathbf{R}$ is a region in $\mathcal{R}$ with infinitely many, separated, interior-connected components. Since $\mathbf{R}$ is bounded, any infinite sequence of interior-connected components $\mathbf{C_1}, \mathbf{C_2} \ldots$ has an accumulation point $\mathbf{p}$. Since

the interior-connected components are separated, the point **p** can be an element of at most one of these; exclude this, if there is one, take the union of the rest, and let **Q** be the closure of that union. Then **Q** has the following properties:

- **Q** ∈ 𝓡.

- **Q** ⊂ **R**.

- Every interior-connected component of **Q** is an interior-connected component of **R**.

- **Q** contains at least one point **p** that is not part of any region that is an interior-connected component of **Q**.

Conversely, if **R** has finitely many, separated, interior-connected components, then no region **Q** can have this property.

We define the property `SeparatedICC(r)` (separated interior-connected components) as follows:

```
SeparatedICC(r) ≡
Region(r) ∧
[∀_{a,b} IntCommComp(a,r) ∧ IntCommComp(b,r) ⇒ [a=b ∨ DC(a,b)]].
```

We can therefore define the property `NaturalNum(r)` as follows:

```
NaturalNum(r) ≡
SeparatedICC(r) ∧
¬∃_q Region(q) ∧ PP(q,r) ∧
      [∀_w IntCommComp(w,q) ⇒ IntCommComp(w,r)] ∧
      [∃_p PtIn(p,q) ∧ ¬∃_w IntConnComp(w,q) ∧ PtIn(p,w)].
```

In the domain $\mathcal{P}[\mathbb{Q}]$ we take `NaturalNum` to be simply the universal relation, since all regions in that domain have finitely many interior-connected components.

To define the relation of two regions having the same number of interior-connected components, we begin by defining a condition that is sufficient but not necessary: Region **A** can be shrunk to **P**, preserving the number of interior-connected components, if every interior-connected component of **A** contains exactly one interior-connected component of **P** and vice versa.

```
Shrink(a,p): a can be shrunk to p, preserving the number of interior-connected com-
ponents.
Shrink(a,p) ≡
      NaturalNum(a) ∧ NaturalNum(p) ∧
      [∀_x IntConnComp(x,a) ⇒ ∃¹_y IntConnComp(y,p) ∧ P(y,x)] ∧
      [∀_y IntConnComp(y,p) ⇒ ∃¹_x IntConnComp(x,a) ∧ P(y,x)] .
```

One can then define the general case of two regions **A** and **B** that have the same number of interior-connected components: Shrink **A** down to **P**, where the interior-connected components of **P** are small dots, and shrink **B** down to **Q**. Match up the components of **P** to the component of **Q** by using a region **C** each of whose components stretches from a component of **P** to one of **Q**. Then **P** and **Q** are both shrinkings of **C** (Figure 1).

```
SameNum(a,b): a and b represent the same natural number.
SameNum(a,b) ≡
∃_{p,q,c} Shrink(a,p) ∧ Shrink(b,q) ∧ Shrink(c,p) ∧ Shrink(c,q).
```
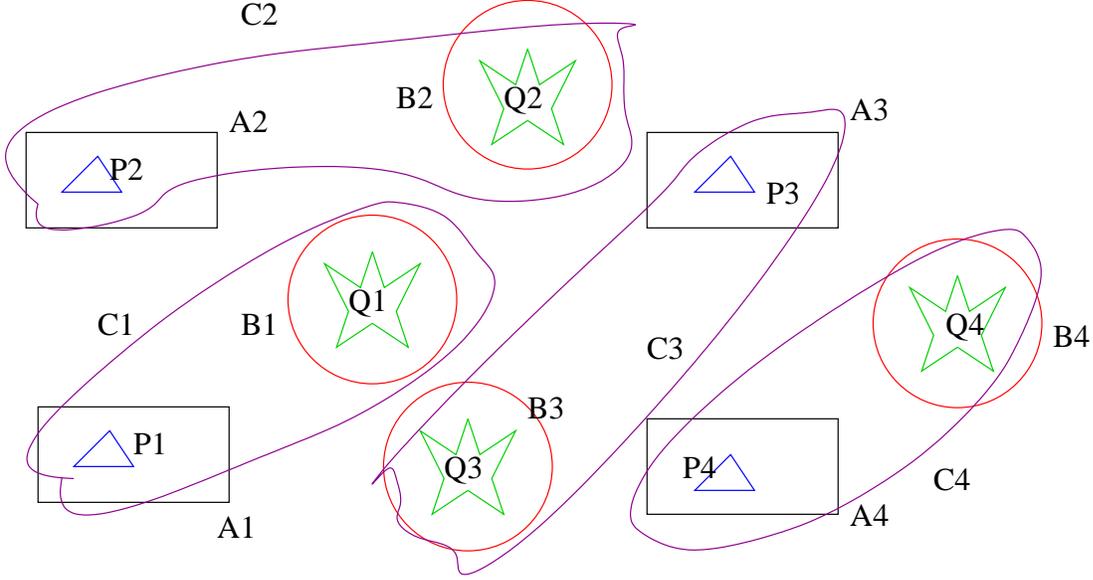
A is the union of the rectangles. *B* is the union of the circles.
C is the union of the irregular shapes. *P* is the union of the triangles.
Q is the union of the stars.

Figure 1: SameNum(A, B)

The sum of two numbers is just their disjoint union.

```
Plus(a,b,c):  a+b=c.
Plus(a,b,c) ≡
NaturalNum(a) ∧ NaturalNum(b) ∧ NaturalNum(c) ∧
        ∃_{d,e} SameNum(b,d) ∧ DC(a,d) ∧ Union(a,d,e) ∧ SameNum(e,c).
```

The number r is the product $p \cdot q$ if $r \subset p$ and every connected component of p contains exactly q components of r.

```
Times(p,q,r):  p·q = r.
Times(p,q,r) ≡
NaturalNum(p) ∧ NaturalNum(q) ∧ NaturalNum(r) ∧
        ∃_w SameNum(w,r) ∧ P(w,p) ∧
            ∀_{c,d} IntConnComp(c,p) ∧ Intersect(w,c,d) ⇒ SameNum(d,q).
```

We have thus established the following lemma:

Lemma 3 *The structure $\langle \mathbb{N}, +, * \rangle$ can be defined in terms of the structure $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$. Any arithmetic relation over $\mathbb{N}$ or over $\mathbb{N}^*$ can be defined in the structure $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$.*

In the structure $\langle \mathcal{R}, \mathtt{C} \rangle$ we can go further, taking advantage of the fact that $\mathcal{R}$ includes pathological regions with infinitely many interior-connected components. A sequence $\hat{s} = \langle \hat{s}[1], \hat{s}[2] \ldots \rangle$ is represented in terms of three regions $\mathbf{A}, \mathbf{B}, \mathbf{C}$. Region $\mathbf{A}$ has infinitely many interior-connected components, $\mathbf{A}_1, \mathbf{A}_2 \ldots$. Region $\mathbf{B}$ is constructed so that $\mathbf{A}_i \cap \mathbf{B}$ has i connected components (this
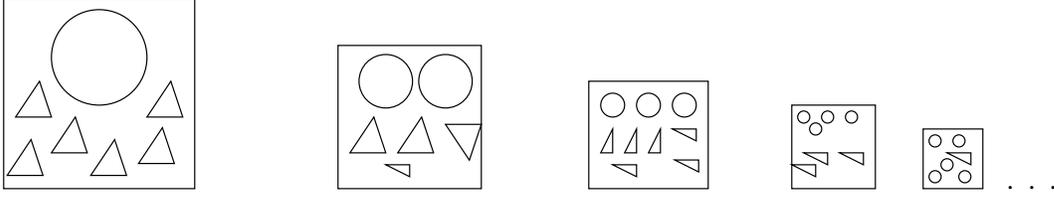
Figure 2: Representation of the sequence 6,4,6,3,1 ...

enables us to associate the index i with the particular component of $\mathbf{A}_i$). Region $\mathbf{C}$ is constructed so that $\mathbf{A}[i] \cap \mathbf{C}$ has $\mathtt{s}[i]$ interior-connected components (Figure 2).

> Index(a,b,c;i,p): s[i]=p, where s is encoded as $\langle \mathtt{a}, \mathtt{b}, \mathtt{c} \rangle$.
> Index(a,b,c,i,p) $\equiv$
> $\exists_{\mathtt{x,y,z}}$ IntConnComp(x,a) $\wedge$ Intersect(x,b,y) $\wedge$ Intersect(x,c,z) $\wedge$
>   SameNum(y,i) $\wedge$ SameNum(z,p).
>
> IsSequence(a,b,c): $\langle \mathtt{a}, \mathtt{b}, \mathtt{c} \rangle$ represents a sequence s if each natural number i appears at least once as an index, and s[i] is always uniquely defined and a natural number.
>
> IsSequence(a,b,c) $\equiv$
> $\forall_{\mathtt{i}}$ NaturalNum(i) $\Rightarrow$ $\exists_{\mathtt{p}}^{1}$ NaturalNum(p) $\wedge$ Index(a,b,c,i,p).

We have thus proven the following:

Lemma 4 *The structure $\langle \mathbb{N} \cup \mathbb{N}^{\mathbb{N}}, +, *, [] \rangle$ is definable in the structure $\langle \mathcal{R}, \mathtt{C} \rangle$. Any analytical relation over $\mathbb{N} \cup \mathbb{N}^{\mathbb{N}}$ is definable in $\langle \mathcal{R}, \mathtt{C} \rangle$.*

3   Arithmetic topological relations over polyhedra

In this section, we characterize the properties that are definable in the structure $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$. As we have shown, starting with the structure $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$ we can define an extended structure that includes all iregions (polytopes), $\mathbb{N}$ and $\mathbb{N}^{*}$ together with the relations on iregions defined in section 2.2 and the relations over $\mathbb{N}$ and $\mathbb{N}^{*}$ defined in section 2.3. This extended structure will be denoted $\mathcal{P}'$. Any relation that is definable in $\mathcal{P}'$ is implicitly definable in $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$; any relation over elements of $\mathcal{P}[\mathbb{Q}]$ that is definable in $\mathcal{P}'$ is explicitly definable in $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$.

Our final theorem is only proven for two- and three-dimensional Euclidean space $\mathbb{E}^2$ and $\mathbb{E}^3$, but we have stated as many lemmas as possible in terms of $\mathbb{E}^n$ in general. We have not attempted to generalize to topological spaces other than $\mathbb{E}^n$.

In section 3.1 we show how the topology of a polyhedron can be represented as a tuple of numbers, and we define the class of arithmetic relations over polytopes. In section 3.2 we discuss how the vertices of a simplicial complex can be labelled using regions. In section 3.3 we show how the structure of a triangulation can be defined in terms of a fixed number of regions. we complete the proof in section 3.4. In section 3.5 we extend this to some other domains, such as the domain of all polyhedra and the domain of semi-algebraic regions.

3.1   The structural description of a polyhedron

We begin with some standard definition in the theory of simplices.

11
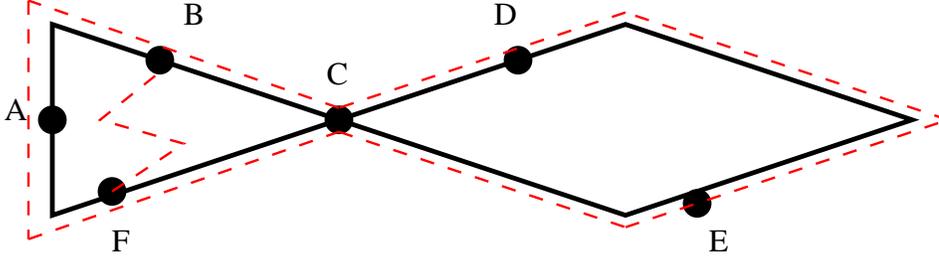
Figure 3: A triangulation

Definition 8 *A k-ball in $\mathbb{E}^n$ is a point set that is homeomorphic to a k-dimensional simplex.*

If U is a k-ball, let $@U$ be the boundary of U in the manifold sense. That is, if h is a homeomorphism from a k-dimensional simplex S to U, and F is the union of the faces of S, then $@U = h(F)$.

Definition 9 *A triangulation $\mathcal{T}$ is a finite set of 0-balls (points), 1-balls, 2-balls, . . . n-balls satisfying the following:*

- *If U is a k-ball in $\mathcal{T}$, then there are exactly $k + 1$ 0-balls $\mathbf{v} \in \mathcal{T}$ such that $\mathbf{v} \in @U$. This set is denoted* Vertices(U).

- *If $U \in \mathcal{T}$, then for any non-empty proper subset $S \subset$ Vertices(U) there exists a unique element $V \in \mathcal{T}$ such that Vertices$(V) = S$. V is called a* face *of U relative to $\mathcal{T}$.*

- *If $U, V \in \mathcal{T}$ then either $U \cap V = \emptyset$ or $U \cap V$ is both a face of U and a face of V.*

- *$@U$ is equal to the union of the faces of U.*

Definition 10 *A* piece *of a triangulation $\mathcal{T}$ is an element of $\mathcal{T}$ that is not a face of any other element. A triangulation $\mathcal{T}$ is* fully dimensional *if every piece of $\mathcal{T}$ is a n-ball; equivalently, if $|\mathcal{T}|$ is topologically closed regular.*

A Euclidean complex is a special case of a triangulation, in which the elements are all simplices: The elements of a Euclidean complex are Euclidean simplices: points, lines, triangles, tetrahedra. The elements of a triangulation are polytopes (because the space of regions is defined to be $\mathcal{P}[\mathbb{Q}]$) whose topological relations satisfy the same constraints as the simplices of a Euclidean complex. Being a Euclidean complex is a geometric property of a set of polytopes, but certainly not a topological property, and hence not definable in $\langle \mathcal{P}[\mathbb{Q}], \mathsf{C} \rangle$. Being a triangulation *is* a topological property of a set of polytopes, and, as we shall show below, it *is* definable in $\langle \mathcal{P}[\mathbb{Q}], \mathsf{C} \rangle$.

Figure 3 shows a triangulation. The vertices are A . . . F. The 1-balls are shown with dashed lines; their vertex sets are { A,B }, { A,F }, { B,C }, { B,F } { C,D }, { C,E }, { C,F }, and { D,E }. The 2-balls are the two parts of the triangle on the left and the rhombus on the right; their vertex sets are { A,B,F }, {C,B,F} and {C,D,E }.

If $\mathcal{T}$ is a triangulation, $\mathcal{D}$ is a subtriangulation of $\mathcal{T}$ and $\mathbf{P} = |\mathcal{D}|$, then we will write $\mathcal{D} = \mathcal{T}|_{\mathbf{P}}$, read "$\mathcal{T}$ restricted to $\mathbf{P}$".

Definition 11 *Let $\mathbf{P}^1 \ldots \mathbf{P}^k$ be polyhedra. A* triangulated package *for $\mathbf{P}^1 \ldots \mathbf{P}^k$ is a triangulation $\mathcal{T}$ such that*

- $|\mathcal{T}|$ *is an n-ball;*

- *For $i = 1$ to k there exists a sub-triangulation $\mathcal{D}^i \subset \mathcal{T}$ such that $\mathbf{P}^i = |\mathcal{D}^i|$.*

Lemma 5 *If $\mathbf{P}^1 \dots \mathbf{P}^k$ are rational polyhedra, then there exists a rational Euclidean complex that is a triangulated package for $\mathbf{P}^1 \dots \mathbf{P}^k$.*

Proof: Immediate from the triangulation theorem ([26] p. 16).

Definition 12 *Let $\mathbf{P}^1 \dots \mathbf{P}^k$ be polyhedra, and let $\mathcal{T}$ be a triangulated package for $\mathbf{P}^1 \dots \mathbf{P}^k$. Let $\mathcal{D}^i = \mathcal{T}|_{\mathbf{P}^i}$. Let Φ be an injection from the vertices of $\mathcal{T}$ to $\mathbb{N}$; that is, an assignment of a unique identification code to each vertex. The* structural description *of $\mathbf{P}^1 \dots \mathbf{P}^k$ based on $\mathcal{T}, \Phi$ is the $k + 1$-tuple of sets of $n + 1$-tuples*

$$\begin{aligned}
\langle \; & \{\Phi(\text{Vertices}(\mathbf{U})) \mid \mathbf{U} \text{ is an n-ball in } \mathcal{T}\}, \\
& \{\Phi(\text{Vertices}(\mathbf{U})) \mid \mathbf{U} \text{ is an n-ball in } \mathcal{D}^1\}, \\
& \quad \dots \\
& \{\Phi(\text{Vertices}(\mathbf{U})) \mid \mathbf{U} \text{ is an n-ball in } \mathcal{D}^k\} \\
\rangle &
\end{aligned}$$

Definition 13 *Let $\hat{\mathbf{P}} = \langle \mathbf{P}^1 \dots \mathbf{P}^k \rangle$ be a k-tuple of polyhedra. A $k + 1$-tuple of sets of n-tuples E is a* structural description *of $\hat{\mathbf{P}}$ if there exist $\mathcal{T}, \Phi$ such that E is the structural description of $\hat{\mathbf{P}}$ based on $\mathcal{T}, \Phi$.*

An example is shown in figure 4. $\mathbf{P}^1$ is the rectangle, and $\mathbf{P}^2$ is the triangle. The edges of the triangulation are shown in dashed lines. The structural description of $\mathbf{P}^1 \dots \mathbf{P}^k$ based on $\mathcal{T}, \Phi$ is

⟨ { { 1,2,3 }, { 1,2,5 }, { 1,3,4 }, { 2,3,6 }, { 2,4,5 }, { 2,4,6 }, { 3,4,6 } },
  { { 2,3,6 }, { 2,4,5 }, { 2,4,6 }, { 3,4,6 } },
  { { 1,2,3 }, { 2,3,6 } }
⟩

Lemma 6 *Any homeomorphism h from the unit $n - 1$-sphere to itself can be extended to a homeomorphism from the unit n-ball to itself. (Weak Alexander trick).*

Proof: Define $h(\vec{x}) = |\vec{x}| \cdot h(\vec{x}/|\vec{x}|)$. ∎

Lemma 7 *If triangulations $\mathcal{T}$ and $\mathcal{D}$ have the same structural description then they are homeomorphic.*

Proof: Since $\mathcal{T}$ and $\mathcal{D}$ have the same structural description, there is an isomorphism $\iota : \mathcal{T} \mapsto \mathcal{D}$ that takes each k-cell in $\mathcal{T}$ to a k-cell in $\mathcal{D}$. We define a homeomorphism h from $\mathcal{T}$ to $\mathcal{D}$ by induction on the dimension of the cells.

For each vertex v in $\mathcal{T}$, define $h(v) = \iota(v)$.

Assume that h has been defined over the cells of dimension $k - 1$. Let U be a k-cell in $\mathcal{T}$ and let $V = \iota(W)$. Since U and V are k-cells, there are homeomorphisms g_U and g_V from the unit k-ball to U and V respectively.
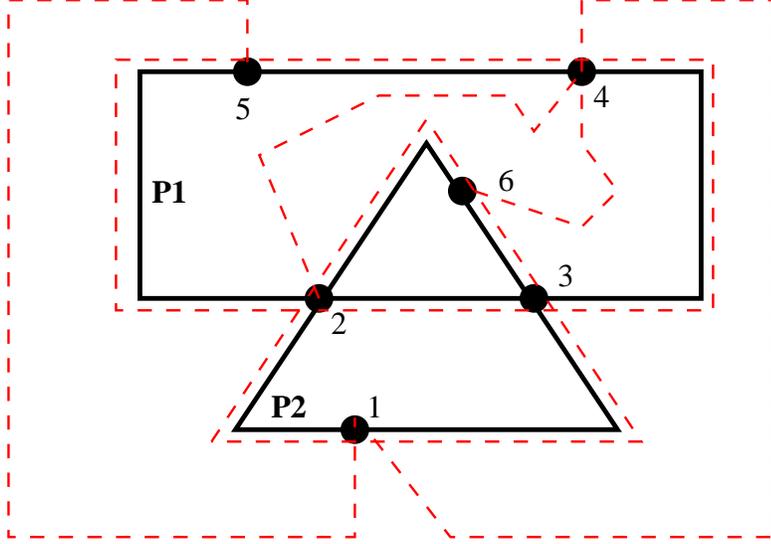
Figure 4: Triangulated Package

Let @U and @V be the boundaries of U and V; then g_U and g_V map the unit $k-1$ sphere into @U and @V. Moreover @U and @V are unions of $k+1$ cells of dimension $k-1$ in $\mathcal{T}$ and $\mathcal{D}$, so by induction h is a homeomorphism from @U to @V. Therefore $q = g_V^{-1} \circ h \circ g_U$ is a homeomorphism from the unit k-sphere to itself. By lemma 6, we can extend q to be a homeomorphism from the unit k-ball to itself. Then $h_{U,V} = g_V \circ q \circ g_U^{-1}$ is a homeomorphism from U to V that agrees with h on @U. Thus we can extend h to include $h_{U,V}$. Since cells in $\mathcal{T}$ and $\mathcal{D}$ intersect only along lower-dimensional faces, this process never creates inconsistencies or discontinuities. ∎

Definition 14 *A triangulation $\mathcal{T}$ is* realizable in $\mathbb{E}^n$ *if there exists a Euclidean complex in $\mathbb{E}^n$ with the same structural description.*

Note that, though being Euclidean is not a topological property of a triangulation, being realizable in $\mathbb{E}^n$ is a topological property, since it is invariant under homeomorphism.

Definition 15 *Two tuples of regions $\hat{\mathbf{P}}$ and $\hat{\mathbf{Q}}$ are* similarly situated *if there is a homeomorphism h from $\mathbb{E}^n$ to itself such that $h(\hat{\mathbf{P}}) = \hat{\mathbf{Q}}$. They are* rational PL-similarly situated *if h is rational piecewise linear.*

Lemma 8 *In $\mathbb{E}^2$ and $\mathbb{E}^3$, if two rational polytopes or two tuples of rational polytopes are homeomorphic, then they are rational PL-homeomorphic; and if they are similarly situated, then they are rational PL-similarly situated.*

Proof: In $\mathbb{E}^2$ and $\mathbb{E}^3$ if two tuples of polytopes are homeomorphic, then they are PL-homeomorphic[1] [20] p. 253. By [9] lemma 15, if two tuples of rational polytopes are PL-homeomorphic then they

---

[1]This theorem is one formulation of the problem known as the *Hauptvermutung*. The statement does not hold in all dimensions; the first counter-example was found by Milnor [19]. In the anonymous article, "Polyhedron, abstract" in the online *Encyclopedia of Mathematics*, it is stated that the result is known not to hold in dimensions 5 and greater, and it seems to be the case that the result is open in dimension 4. However, the quality of this article is challenged in a comment on the website, as of November 19, 2012, and I have been unable to find a statement of the that is current, clear, and authoritative. In particular, results of this kind are generally stated in terms of the dimension of the polyhedra rather than the dimension of the embedding space, which is the For example, the fact that the statement is false for 5-dimensional polyhedra does not necessarily imply that it is false for polyhedra in $\mathbb{E}^5$.

are rational PL-homeomorphic. ∎

Lemma 9 *In $\mathbb{E}^2$ and $\mathbb{E}^3$, let $\hat{\mathbf{P}}$ and $\hat{\mathbf{Q}}$ be k-tuples of polyhedra. If $\mathbf{P}$ and $\mathbf{Q}$ are PL-similarly situated, then every structural description of $\hat{\mathbf{P}}$ is a structural description of $\hat{\mathbf{Q}}$. Conversely, if any structural description of $\hat{\mathbf{P}}$ is also a structural description of $\hat{\mathbf{Q}}$, then they are PL-similarly situated.*

Proof: The first statement is trivial: If h is a PL-homeomorphism of $\mathbb{E}^n$ that maps P to Q, then h maps any triangulated package for P into a structurally identical triangulated package for Q.

For the second statement: Let $\mathcal{T}$ and $\mathcal{D}$ be triangulated packages for $\hat{\mathbf{P}}$ and $\hat{\mathbf{Q}}$ respectively that have identical structural descriptions. By lemmas 7 and 8, $\mathcal{T}$ and $\mathcal{D}$ are PL-homeomorphic. But since $|\mathcal{T}|$ and $|\mathcal{D}|$ are n-balls and contain all of $\hat{\mathbf{P}}$ and $\hat{\mathbf{Q}}$, it is easily shown that the PL-homeomorphism from $|\mathcal{T}|$ to $|\mathcal{D}|$ can be extended to a PL-homeomorphism from $\mathbb{E}^n$ to itself. ∎

Definition 16 *Let $\mathbf{P}^1 \dots \mathbf{P}^k$ be k polyhedra. Let $\mathcal{T}$ be a Euclidean complex that is a triangulated package for $\mathbf{P}^1 \dots \mathbf{P}^k$. Let Φ be a function from $\mathrm{Vertices}(\mathcal{T})$ to $\mathbb{N}$. The $k+2$ tuple $\langle \hat{c}, \hat{v}^0 \dots \hat{v}^k \rangle$ is an exact description of $\mathbf{P}^1 \dots \mathbf{P}^k$ based on $\mathcal{T}$ and Φ if*

- *$\hat{c} = \{ \langle \Phi(v), Coordinates(v) \rangle \mid v \text{ is a vertex of } \mathcal{T} \}$*

- *$\langle \hat{v}^0 \dots \hat{v}_k \rangle$ is the structural description of $\langle \mathbf{P}^1 \dots \mathbf{P}^k \rangle$ based on $\mathcal{T}, \Phi$.*

Clearly, a tuple of this kind $\langle \hat{c}, \hat{v}^0 \dots \hat{v}^k \rangle$ determines the polyhedra $\mathbf{P}^1 \dots \mathbf{P}^k$ uniquely; there is at most one tuple of polyhedra for which $\langle \hat{c}, \hat{v}^0 \dots \hat{v}^k \rangle$ is an exact description. (There may be none, if the description is ill-formed or geometrically inconsistent.) We define the partial function $\eta(\hat{c}, \hat{v}^0 \dots \hat{v}^k)$ as mapping a putative exact description onto a tuple of polyhedra. It is easily seen that the relation "$\eta(\hat{c}, \hat{v}^0 \dots \hat{v}^k)$ defines a tuple of polyhedra" is a computable and hence an arithmetic relation.

Definition 17 *Let $\Gamma(\mathbf{P}^1 \dots \mathbf{P}^k)$ be a k-place relation over $\mathcal{P}[\mathbb{Q}]$. Let $\nabla[\Gamma]$ be the $k+2$-place relation over $\mathbb{N}^*$, "$\eta(\hat{c}, \hat{x}^0 \dots \hat{x}^k)$ is defined and $\Gamma(\eta(\hat{c}, \hat{x}^0 \dots \hat{x}^k))$". The relation Γ is said to be arithmetic if $\nabla[\Gamma]$ is arithmetic.*

Lemma 10 *Let $\Gamma(\mathbf{P}^1 \dots \mathbf{P}^k)$ be an arithmetic k-place relation over $\mathcal{P}[\mathbb{Q}]$ that is invariant under rational PL-homeomorphisms of $\mathbb{E}^n$ to itself. Let $\nabla[\Gamma]$ be defined as in definition 17. Let $\Psi(\hat{v}^0 \dots \hat{v}^k)$ be the $k+1$-place relation over $\mathbb{N}^*$.*

$$\Psi(\hat{v}^0 \dots \hat{v}^k) \equiv \exists_{\hat{c}} \, \Gamma(\hat{c}, \hat{v}^0 \dots \hat{v}^k)$$

Then:

- a. *For any $\mathbf{P}^1 \dots \mathbf{P}^k \in \mathcal{P}[\mathbb{Q}]$ and $\hat{v}^0 \dots \hat{v}^k \in \mathbb{N}^*$, if $\langle \hat{v}^0 \dots \hat{v}^k \rangle$ is a structural description of $\langle \mathbf{P}^1 \dots \mathbf{P}^k \rangle$, then $\Psi(\hat{v}^0 \dots \hat{v}^k)$ if and only if $\Gamma(\mathbf{P}^1 \dots \mathbf{P}^k)$.*

- b. *Ψ is an arithmetic relation over $\mathbb{N}^*$.*

Proof of a: Let $\hat{v}^0 \dots \hat{v}^k$ be a structural description of $\mathbf{P}^1 \dots \mathbf{P}^k$. By the assumption of the lemma, $\Psi(\hat{v}^0 \dots \hat{v}^k)$ if and only if there exist $\mathbf{Q}^1 \dots \mathbf{Q}^k$ such that $\langle \hat{v}^0 \dots \hat{v}^k \rangle$ is a structural description of $\langle \mathbf{Q}^1 \dots \mathbf{Q}^k \rangle$ and $\Gamma(\mathbf{Q}^1 \dots \mathbf{Q}^k)$. By lemma 9 $\mathbf{Q}^1 \dots \mathbf{Q}^k$ are PL-similarly situated to $\mathbf{P}^1 \dots \mathbf{P}^k$. Since Γ is invariant under PL-homeomorphism, we have $\Gamma(\mathbf{Q}^1 \dots \mathbf{Q}^k)$ if and only if $\Gamma(\mathbf{P}^1 \dots \mathbf{P}^k)$.

Proof of b: By definition 17, since Γ is arithmetic, $\nabla(\Gamma)$ is arithmetic, so Ψ is also arithmetic.

∎

In our construction here, we introduce the coordinate tuple $\hat{c}$ in definition 16, only to immediately remove it by existential quantification in lemma 10. (By contrast, the construction of sections 4 and 5 involve actually labelling points with their coordinates in a coordinate system.) Since we are only representing topological relations, which are determined by the structural description, the formula Ψ of lemma 10 depends only on the structural description. However, we nonetheless include the coordinates in definition 17 because we wish to define the concept of "an arithmetic relation over regions" in general, independently of the particular context of topological relations. For example, the relation "The distance from $\mathbf{P}^1$ to $\mathbf{P}^2$ is less than 2" is certainly an arithmetic relation over the polyhedra, since it can be computed from the exact descriptions, though it is not a topological relation.

It is not immediately obvious what kinds of topological relations over polyhedra are in fact arithmetic. Suppose, for example, that we are interested in the relation `PLHomeomorphicStructures(û, v̂)` that holds if there exists complexes $\mathcal{A}$ and $\mathcal{B}$ such that $\mathbf{u}$ is a structural description of $\mathcal{A}$, $\mathbf{v}$ is a structural description of $\mathcal{B}$, and $|\mathcal{A}|$ is PL-homeomorphic to $|\mathcal{B}|$. It is difficult to write down an algorithm that computes this relation in $\mathbb{E}^2$ and $\mathbb{E}^3$ and it is undecidable for manifolds of dimension 4 and higher [18].[2] However, combining the powers of arbitrary quantification and Gödel numbering, it is easy to show that this and similar relations are indeed arithmetic.[3] We proceed as follows: If rational polyhedra $\mathbf{A}$ and $\mathbf{B}$ are PL-homomorphic, then there exists a rational PL-homeomorphism Γ from $\mathbf{A}$ to $\mathbf{B}$ with finitely many cells. Such a homeomorphism can be encoded as a finite data structure with integer fields; that is, an element of $\mathbb{N}^*$. The structural description is another element of $\mathbb{N}^*$ and the function that applies a rational homeomorphism thus encoded to a Euclidean complex is certainly computable. The relation `PLHomeomorphicStructure(û, v̂)` can therefore be expressed as follows: "There exist rational coordinate assignments $\hat{c}, \hat{d}$ and an encoding $\hat{g}$ of PL-homeomorphism Γ such that the image under Γ of the instantiation of $\hat{u}$ with coordinates $\hat{c}$ is equal to the instantiation of $\hat{v}$ with coordinates $\hat{d}$."

3.2   Using regions to label vertices

We now show how the structural description of a triangulation can be instantiated as a pair of regions $\mathbf{L}, \mathbf{M}$ that label the vertices and n-balls of the triangulation.

Definition 18 *Let* $\mathbf{L}, \mathbf{M}$ *be a pair of regions; let* $\mathbf{p}$ *be a point; and let x be a natural number. The regions* $\mathbf{L}, \mathbf{M}$ *label* $\mathbf{p}$ *with x if there is an interior-connected component* $\mathbf{U}$ *of* $\mathbf{L}$ *such that* $\mathbf{p} \in \mathbf{U}^\circ$ *and* $\mathbf{U} \cap \mathbf{M}$ *has x interior-connected components.*

Figure 5 shows the labelling of the vertices of a complex with two triangles. $\mathbf{L}$ is the union of the small squares and $\mathbf{M}$ is the union of the very small pentagons.

Lemma 11 *Let* $\mathcal{T}$ *be a triangulation, and let* Φ *be a function from the vertices of* $\mathcal{T}$ *to* $\mathbb{N}$. *Then there exist* $\mathbf{L}, \mathbf{M} \in \mathcal{P}[\mathbb{Q}]$ *such that each vertex* $\mathbf{v}$ *of* $\mathcal{T}$ *is labelled* $\Phi(\mathbf{v})$.

Proof: Choose disjoint connected neighborhoods $\mathbf{U_v}$ of the vertices of $\mathcal{T}$. Let $\mathbf{L}$ be the union of the $\mathbf{U_v}$. Inside each $\mathbf{U_v}$ choose $\Phi(\mathbf{v})$ disjoint connected regions. Let $\mathbf{M}$ be the union of all these. ∎

---

[2]The theorem in [18] is stated in terms of homeomorphisms and polyhedra generally but the construction, which is purely combinatorial, is unchanged when the statement is restricted to PL-homeomorphisms and rational polyhedra.

[3]Whether homeomorphism over rational polyhedra, as distinguished from PL-homeomorphism, is an arithmetic relation in dimensions greater than 3, I do not know.
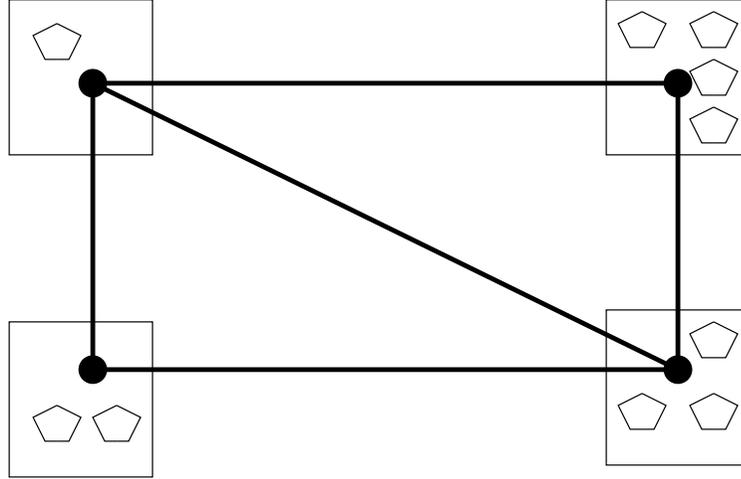
Figure 5: Labelling

3.3 Constructing triangulations from a fixed number of regions

In order to translate the definitions in section 3.1 into the language $\langle \mathcal{P}[\mathbb{Q}], \mathsf{C} \rangle$, it is necessary to quantify over triangulations, which means that a triangulation of arbitrarily many cells must be represented in terms of a *fixed* number of polyhedra. To accomplish this, we return to the discussion of iregions, begun in section 2.2. Recall that an iregion is a point set that is equal to the intersection of two regions. We will designate the space of iregions formed by regions in $\mathcal{P}[\mathbb{Q}]$ as $\mathcal{P}[\mathbb{Q}]'$. The first question is, what point sets are iregions? The answer is as general as one could reasonably hope for.

Lemma 12 *Any polytope is an iregion.*

Proof: Let $\mathcal{C}$ be a Euclidean complex. We wish to find polyhedra $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{A} \cap \mathbf{B} = |\mathcal{C}|$.

First let $\mathbf{R}$ be the union of all the n-balls in $\mathcal{C}$. We can simply include these in both $\mathbf{A}$ and $\mathbf{B}$. Let $\mathbf{A}_0 = \mathbf{B}_0 = \mathbf{R}$.

To handle the lower-dimensional pieces of $\mathcal{C}$ we proceed as follows. Let $\mathbf{U}$ be a open rational n-ball polytope that contains $\mathcal{C}$ in its interior. Consider the set difference $D_0 = U \setminus |\mathcal{C}|$. Let $P_1 \ldots P_m$ be the number of lower-dimensional pieces of $\mathcal{C}$. Now carry out the following procedure:

for $(i \leftarrow 1$ to $m)$ {
    using the triangulation theorem, construct a Euclidean complex $\mathcal{E}$ triangulating of D_{i-1}
        such that P_i is a face of a simplex in $\mathcal{E}$.
    let $\mathbf{T}_i$ and $\mathbf{W}_i$ be two n-dimensional simplices in $\mathcal{E}$ that have P_i as a face.
        /* if P_i is $n-1$ dimensional, these will be on opposite sides of P;
            otherwise, there will be more than two such simplices; choose any two */
    construct n-dimensional simplices $\mathbf{U}_i$ and $\mathbf{V}_i$ such that
        (a) P_i is a face of $\mathbf{U}_i$ and of $\mathbf{V}_i$;
        (b) $\mathbf{U}_i \setminus P_i \subset \mathbf{T}_i^\circ$ and $\mathbf{V}_i \setminus P_i \subset \mathbf{W}_i^\circ$;
    $\mathbf{A}_i \leftarrow \mathbf{A}_{i-1} \cup \mathbf{U}_i$;
    $\mathbf{B}_i \leftarrow \mathbf{B}_{i-1} \cup \mathbf{V}_i$;
    $D_i \leftarrow D_{i-1} \setminus (\mathbf{U}_i \cup \mathbf{V}_i)$.
}

Note that, by condition (b), $\mathbf{U}_i \cap |\mathcal{C}| = \mathbf{V}_i \cap |\mathcal{C}| = P_i$. In particular, $\mathbf{U}_i$ and $\mathbf{V}_i$ do not meet any of the other P_j.

It is easily shown that each step of this procedure can be executed; that $\mathbf{U}_i \cap \mathbf{B}_m = \mathbf{V}_i \cap \mathbf{A}_m = P_i$ and that therefore $\mathbf{A}_m \cap \mathbf{B}_m = \mathbf{R} \cup (\cup_{i=1}^m P_i) = |\mathcal{C}|$. ∎

We can now characterize a fully-dimensional triangulation $\mathcal{T}$ in $\mathbb{E}^n$ in terms of a region and n iregions (and thus $2n+1$ regions); namely, the region $|\mathcal{T}|$ and, for every $k < n$, the union of the k-dimensional elements of $\mathcal{T}$.

The following lemmas show that we can define the relations over triangulations that we need as formulas in $\langle \mathcal{P}[\mathbb{Q}], \mathbb{C} \rangle$.

For the remainder of this paper, we will describe the definitions of relations in English, but we will not give the formalism. It should in each case be clear how the formalism could be constructed, if desired.

Lemma 13 *In $\mathbb{E}^n$, the relation* `SimpleCurve(x,a,b)` *meaning "x is a simple curve with endpoints a and b" is definable in $\mathcal{P}'$.*

Proof: Iregion x is a simple curve from a to b iff a ≠ b and x is a setwise minimal connected point set containing a and b. Note that if x is a connected polytope containing a and b then there exists a piecewise-linear path y from a to b such that y ⊂ x. Therefore, x is setwise minimal only if it is equal to such a path y. ∎

Lemma 14 *In $\mathbb{E}^2$ the relation* `2Ball(x)` *meaning "x is a 2-ball" is definable in $\langle \mathcal{P}[\mathbb{Q}], \mathbb{C} \rangle$.*

Proof: See [24], theorem 2.89.

Lemma 15 *In $\mathbb{E}^3$ the relation* `3Ball(x)` *meaning "x is a 3-ball" is definable in $\langle \mathcal{P}[\mathbb{Q}], \mathbb{C} \rangle$.*

Proof: See [24], theorem 2.132.

Lemma 16 *In $\mathbb{E}^3$ the relation* `2Ball(x)` *meaning "x is a 2-ball" is definable in $\mathcal{P}'$.*

Proof: See [24], proposition 2.133.

Definition 19 *Let X be a 2-ball. A simple curve C separates X if $X \setminus C$ is not path connected.*

Lemma 17 *If X is a 2-ball, and $C \subset X$ is a simple curve that does not meet the boundary of X, then C does not separate X.*

Proof: By a well-known theorem, no simple curve separates $\mathbb{E}^2$ ([29] proposition 2.11). Since X° is homeomorphic to $\mathbb{E}^2$, no simple curve in X° separates X°. Since C is a positive distance from @X (they are disjoint compact sets), C does not separate @X from $X^\circ \setminus C$. ∎

A subtlety in the above argument should be noted. The result in [29] states that for any curve C, and any two points $\mathbf{u}, \mathbf{v} \in \mathbb{E}^2 \setminus C$, there exists a curve $P \subset \mathbb{E}^2 \setminus C$ with endpoints $\mathbf{u}$ and $\mathbf{v}$. In our discussion here, everything is placed in the space of rational polytopes; thus the desired statement is that, for any states that for any rational polyline C, and any two rational points $\mathbf{u}, \mathbf{v} \in \mathbb{E}^2 \setminus \mathbf{C}$, there exists a rational polyline $P \subset \mathbb{E}^2 \setminus C$ with endpoints $\mathbf{u}$ and $\mathbf{v}$. As regards the existence of P, this is in fact a stronger statement. However, the proof in [29] in fact constructs a polyline, which can easily be chosen to be a rational polyline.

18

Lemma 18 *If X is a 2-ball, and $C \subset X$ is a simple curve that meets the boundary of X at a single point $\mathbf{p}$, then C does not separate X.*

Proof by contradiction. Suppose that C separates X; thus $X \setminus C$ has at least two path-connected components. It is clear that any two points on $@X \setminus \{\mathbf{p}\}$ are connected by a path that goes around $@X$ in the direction that avoids $\mathbf{p}$, so all of $@X$ is in one path-connected component of $X \setminus C$. Let $\mathbf{u}$ be a point in a different path-connected component of $X \setminus C$ and let $\mathbf{v}$ be a point in $@X \setminus \{\mathbf{p}\}$. Thus $\mathbf{u} \in X^\circ$. By the theorem cited in the proof of lemma 17, there is a simple curve $Q \subset \mathbb{E}^2 \setminus C$ with endpoints $\mathbf{u}$ and $\mathbf{v}$. Parametrize Q by a continuous function $\phi : [0, 1] \mapsto Q$, where $\phi(0) = \mathbf{u}, \phi(1) = \mathbf{v}$. Let $\psi(t)$ be the distance from $\phi(t)$ to the complement of X; thus $\psi(0) > 0$ and $\psi(1) = 0$. Let t' be the minimal value of t for which $\psi(t') = 0$. Then $\phi([0, t'])$ is a curve in $X \setminus C$ from $\mathbf{u}$ to $@X$, contrary to assumption. ∎

Lemma 19 *Let X be a 2-ball and let C be a setwise-minimal separator of X. Then the endpoints of C lie in the boundary of X.*

Proof: By lemmas 17 and 18 C meets the boundary of X at at least two points. Conversely, a simple curve that does meet the boundary of X at two separate points does divide X, by the Jordan curve theorem.

Suppose then that C meets the boundary of X at two points $\mathbf{p}$ and $\mathbf{q}$; then the segment of C between $\mathbf{p}$ and $\mathbf{q}$ divides X. Since C is setwise-minimal, it must be equal to that segment; that is, $\mathbf{p}$ and $\mathbf{q}$ are the endpoints of C. ∎

Lemma 20 *In either $\mathbb{E}^2$ or $\mathbb{E}^3$ the relation `2BallBd(x,y)` meaning "x is a 2-ball with boundary y" is definable in $\mathcal{P}'$.*

Proof: By lemma 19, `2BallBd(x,y)` can be defined as the property, "A point $\mathbf{p}$ is in $\mathbf{Y}$ iff there exists a curve C such that C is a setwise-minimal divider of $\mathbf{X}$ and $\mathbf{p}$ is an endpoint of C." ∎

Lemma 21 *In $\mathbb{E}^3$ the relation `3BallBd(x,y)` meaning "x is a 3-ball with boundary y" is definable in $\mathcal{P}'$.*

Proof: A point $\mathbf{p}$ is on the boundary of ball $\mathbf{X}$ if $\mathbf{p} \in \mathbf{X}$ and $\mathbf{p} \in \mathbf{E}$ for some $\mathbf{E}$ that is externally connected to $\mathbf{X}$. ∎

We will describe the construction of a three-dimensional triangulation; the two-dimensional case is exactly analogous.

We will give the construction of a triangulation $\mathcal{T}$ as a quadruple of four iregions $\langle V, E, F, \mathbf{B} \rangle$ where V is the union of all the vertices in $\mathcal{T}$, E is the union of the edges, F is the union of the faces, and $\mathbf{B} = |\mathcal{T}|$.

Definition 20 *In $\mathbb{E}^3$ let $\langle V, E, F, \mathbf{B} \rangle$ be a quadruple of iregions.*

- *A simple curve C is a constituent of E with respect to V if $C \subset E$, the endpoints of C are in V, and no other points in V are C.*

- *A 2-ball Q is a constituent of F with respect to E, V if $Q \subset F$, $@Q = Q \cap E$ is the union of three constituents of E, and $Q \cap V$ is a set of three vertices.*
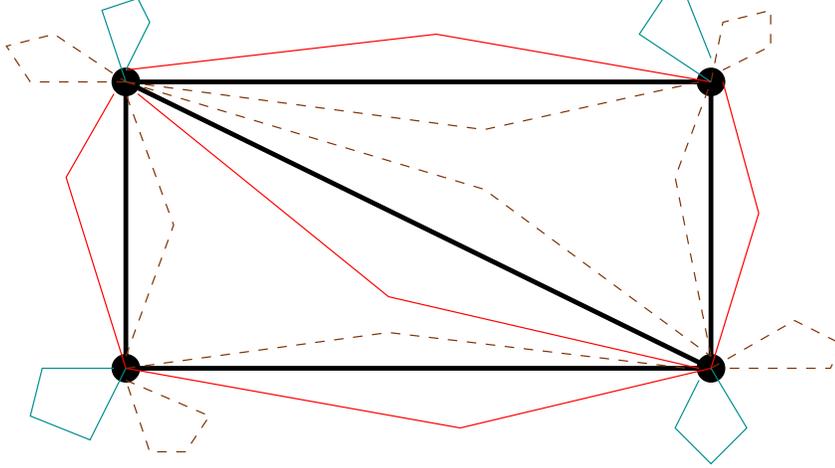
Figure 6: Construction of a complex

- *A 3-ball* $\mathbf{W}$ *is a* constituent of $\mathbf{B}$ with respect to E, V, F if $\mathbf{W} \subset \mathbf{B}$, $@\mathbf{W} = \mathbf{W} \cap F$ *is the union of four constituents of F, and* $\mathbf{W} \cap V$ *is a set of four vertices.*

Definition 21 *A quadruple $\langle V, E, F, \mathbf{B} \rangle$ represents a triangulation if it meets the following conditions.*

A. *V is a finite set of points.*

B. *E, F, and $\mathbf{B}$ each has finitely many constituents and are each equal to the union of their constituents.*

C. *The intersection of any two constituents is either a point in V, a constituent, or empty.*

Figure 6 shows this construction for a complex with two triangles. The complex is shown in thick lines. The set of vertices V is constructed as the intersection of the dashed quadrilaterals with the solid quadrilaterals. The set of edges E is constructed as the intersection of the dashed triangles with the solid, thin triangles.

Lemma 22 *A quadruple $\langle V, E, F, \mathbf{B} \rangle$ satisfies the conditions of definition 21 if and only if the set of vertices of V union the set of constituents of E union the set of constituents of F union the set of constituents of $\mathbf{B}$ is a triangulation.*

Proof: Immediate from the definitions. ∎

Lemma 23 *In $\mathbb{E}^3$ the relations* `Triangulation(V,E,F,B)` *meaning "$\langle V, E, F, \mathbf{B} \rangle$ represents a triangulation" and* `VertexConstituent(Q,V,E,F,B)`, `EdgeConstituent(Q,V,E,F,B)`, `FaceConstituent(Q,V,E,F,B)`, *and* `SolidConstituent(Q,V,E,F,B)`, *meaning that Q is a constituent of the specified kind in the triangulation represented by $\langle V, E, F, B \rangle$ are definable in $\mathcal{P}'$.*

Proof: The translation of definitions 20 and 21 to $\mathcal{P}'$ is straightforward, given the relations we have already defined. The finiteness conditions can be expressed using labelling, if necessary.

20

Lemma 24 *The relation* `StructuralDescription(v0 ...vk, P1 ...Pk)` *meaning "$\hat{v}^0 \ldots \hat{v}^k$ is a structural description of $\mathbf{P}^1 \ldots \mathbf{P}^k$ " is definable in $\mathcal{P}'$.*

Proof: Given the definitions of triangulations and their constituents in lemma 23, the translation of definition 11, is straightforward. The function Φ from the vertices to the natural numbers used in definitions 12 and 13 is instantiated as a pair of regions $\mathbf{L}, \mathbf{M}$ that label the vertices with Φ. Lemma 11 guarantees that such a pair of labelling regions can always be found. ∎

3.4 Completing the proof

Lemma 25 *In $\mathbb{E}^3$, let $\Gamma(\mathbf{P}^1 \ldots \mathbf{P}^k)$ be a relation over $\mathcal{P}[\mathbb{Q}]$ that is arithmetic and invariant under rational PL-homeomorphisms of the space to itself. Then Γ is definable in $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$.*

Proof: Define the relation $\Delta(\mathbf{P}^1 \ldots \mathbf{P}^k)$ as follows:

$\Delta(\mathbf{P}^1 \ldots \mathbf{P}^k)$ if and only if there exist:

- a triangulation $\mathcal{T}$ of $\mathbf{P}^1 \ldots \mathbf{P}^k$, represented as a quadruple $\langle V, E, F, \mathbf{B} \rangle$;
- a labelling $\mathbf{L}, \mathbf{M}$ of the vertices in V that assigns a different number $\Phi(\mathbf{u})$ to each vertex $\mathbf{u}$;

satisfying the following: Let $\langle \hat{v}^0 \ldots \hat{v}^k \rangle$ be the structural description of $\mathcal{T}$ based on Φ. Then $\langle \hat{v}^0 \ldots \hat{v}^k \rangle$ is realizable in $\mathbb{E}^n$ and satisfies $\nabla[\Gamma](\hat{v}^0 \ldots \hat{v}^k)$, where $\nabla[\Gamma]$ is the relation defined in definition 17.

Δ is clearly definable in $\mathcal{P}'$ using the relations constructed in section 3.3. By lemma 10, Δ is equivalent to Γ. ∎

Lemma 26 *Any relation definable in $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$ is invariant under rational PL-homeomorphisms of $\mathbb{E}^n$ to itself.*

Proof: Immediate from the fact that any rational PL-homeomorphisms maps $\mathcal{P}[\mathbb{Q}]$ to itself and leaves `C(x,y)` invariant. ∎

Lemma 27 *Any relation definable in $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$ is arithmetic.*

Proof: Immediate from the fact that `C(x,y)` is computable from exact descriptions of $\mathbf{X}$ and $\mathbf{Y}$ and hence arithmetic. See also [21], p. 93, theorem 103. ∎

Theorem 1 *In $\mathbb{E}^2$ or $\mathbb{E}^3$, a relation over $\mathcal{P}[\mathbb{Q}]$ is definable over $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$ if and only if it is arithmetic and invariant under rational PL-homeomorphism of the space to itself.*

Proof: The conjunctions of lemmas 25, 26 and 27.

Thus, suppose that you have in mind some particular relation Γ over regions — "$\mathbf{P}$ has a prime number of interior cavities", or "$\mathbf{P}$ and $\mathbf{Q}$ are interlocked toruses in $\mathbb{E}^3$", for example. To express this in $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$, you do the following:

Step 1. Figure out how you can express it as an arithmetic expression $\Psi(\hat{c}, \hat{w}^0 \ldots \hat{w}^k)$ over exact descriptions of the polyhedra $\langle \hat{c}, \hat{w}^0 \ldots \hat{w}^k \rangle$. If, as in this case, it is in fact computable, then you can write a program that takes as input exact descriptions of polyhedra and returns as an output "true" if the descriptions are well-formed and satisfy Γ and "false" otherwise. Then use standard computation theory to translate your program into the first-order theory of arithmetic. In any case, such an expression must exist if Γ is an arithmetic relation, because that is what it means for Γ to be an arithmetic relation (definition 17).

Step 2. Use Grzegorczyk's technique, described in section 2.3 to translate Ψ from step 1 into a formula over C, where the numbers are all represented in terms of numbers of interior-connected components.

Step 3. Using lemma 24, construct the relation `StructuralDescription(w0 ...wk, P1 ...Pk)`. (This depends only on k and is independent of Γ.)

Step 4. Return the formula

$$\exists_{\hat{c}, \hat{w}^0 \ldots \hat{w}^k} \mathtt{StructuralDescription}(\hat{w}^0 \ldots \hat{w}^k, \mathbf{P}^1 \ldots \mathbf{P}^k) \wedge \Psi(\hat{c}, \hat{w}^0 \ldots \hat{w}^k)$$

That is, you use the triangulation and labelling to translate from the topological relation Γ to a relation over numbers, and then you use Grzegorczyk's technique to translate back from the numbers to regions.

3.5 Extension to other domains of regions

Theorem 1 can be extended to some other domains of regions using the elementary equivalence results presented in [9]. For example, let $\mathcal{P}[\mathbb{R}]$ be the class of all polyhedra (not necessarily rational). The two following propositions follow directly from the analysis in [9].

Lemma 28 *Let* $\alpha(\mu^1 \ldots \mu^k)$ *be an open formula with k variables in the first order language with predicate* C *(or indeed any collection of topological predicates). Let* $\mathbf{P}^1 \ldots \mathbf{P}^k$ *be k rational polyhedra. Then* $\alpha(\mathbf{P}^1 \ldots \mathbf{P}^k)$ *is true over the domain* $\mathcal{P}[\mathbb{R}]$ *if and only if it is true over the domain* $\mathcal{P}[\mathbb{Q}]$.

Proof: See [9].

The point here is that the scope of the other quantified variables that are bound in α has changed from rational to real polyhedra, but that does not affect the truth of the formula.

Lemma 29 *Any k-tuple of polyhedra is PL-similarly situated to a k-tuple of rational polyhedra.*

Proof: See [9], lemma 20.

We also have an analogue of lemma 26:

Lemma 30 *Any relation definable in* $\langle \mathcal{P}[\mathbb{R}], C \rangle$ *is invariant under PL-homeomorphisms of* $\mathbb{E}^n$ *to itself.*

Proof: Immediate from the fact that any PL-homeomorphism maps $\mathcal{P}[\mathbb{R}]$ to itself and leaves `C(x,y)` invariant. ∎

Let $\mathcal{A}$ be the class of relations characterized by theorem 1; that is, the class of arithmetic relations that are invariant under rational PL-homeomorphisms of the space to itself. Combining these lemmas with theorem 1, we can characterize the expressive power of $\langle \mathcal{P}[\mathbb{R}], C \rangle$ as follows:

Corollary 31 *In $\mathbb{E}^2$ or $\mathbb{E}^3$, a k-place relation $\Gamma(\mathbf{P}^1 \ldots \mathbf{P}^k)$ is definable in $\langle \mathcal{P}[\mathbb{R}], \mathtt{C} \rangle$ if and only if*

- Γ *is invariant under real PL-homeomorphisms of the space to itself; and*

- *the restriction of Γ to $[\mathcal{P}[\mathbb{Q}]]^k$ is in $\mathcal{A}$.*

Proof: Left to right. Suppose that Γ is definable in $\langle \mathcal{P}[\mathbb{R}], \mathtt{C} \rangle$. Let $\alpha(\mu^1 \ldots \mu^k)$ be an open formula with k variables in the first order language with predicate $\mathtt{C}$ whose extension is Γ. By lemma 30, Γ is invariant over PL-homeomorphisms. We need to show that the restriction of Γ to $[\mathcal{P}[\mathbb{Q}]]^k$ is in $\mathcal{A}$. Let $\mathbf{Q}^1 \ldots \mathbf{Q}^k$ be rational polyhedra. Then $\Gamma(\mathbf{Q}^1 \ldots \mathbf{Q}^k)$ if and only if $\alpha(\mathbf{Q}^1 \ldots \mathbf{Q}^k)$ holds, interpreted over $\mathcal{P}[\mathbb{R}]$. By lemma 28 $\alpha(\mathbf{Q}^1 \ldots \mathbf{Q}^k)$ holds when interpreted over $\mathcal{P}[\mathbb{R}]$ if and only if it holds when interpreted over $\mathcal{P}[\mathbb{Q}]$; and this is a relation in $\mathcal{A}$.

Conversely, suppose that Γ satisfies the two conditions in the conclusion of the lemma. Let α be the formula in $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$ that expresses the restriction of Γ to $[\mathcal{P}[\mathbb{Q}]]^k$. Let $\mathbf{P}^1 \ldots \mathbf{P}^k$ be a k-tuple of polyhedra. Using lemma 29 let $\mathbf{Q}^1 \ldots \mathbf{Q}^k$ be a k-tuple of rational polyhedra that is similarly situated. Since Γ is invariant under PL-homeomorphisms, $\Gamma(\mathbf{P}^1 \ldots \mathbf{P}^k)$ if and only if $\Gamma(\mathbf{Q}^1 \ldots \mathbf{Q}^k)$. By construction of α this is equivalent to $\alpha(\mathbf{Q}^1 \ldots \mathbf{Q}^k)$ interpreted over $\mathcal{P}[\mathbb{Q}]$; but by lemma 28 that is equivalent to $\alpha(\mathbf{Q}^1 \ldots \mathbf{Q}^k)$ interpreted over $\mathcal{P}[\mathbb{R}]$. Since α, interpreted over $\mathcal{P}[\mathbb{R}]$, is invariant under PL-homeomorphisms, this is equivalent to $\alpha(\mathbf{P}^1 \ldots \mathbf{P}^k)$ interpreted over $\mathcal{P}[\mathbb{R}]$. That is, α defines Γ over $\mathcal{P}[\mathbb{R}]$. $\blacksquare$

An equivalent formulation of corollary 31 is as follows. Let α be a formula over $\mathtt{C}$. The extension of α over $\mathcal{P}[\mathbb{R}]$ is the image of its extension over $\mathcal{P}[\mathbb{Q}]$ under all PL-homeomorphisms of the space to itself. Therefore a relation is expressible in $\langle \mathcal{P}[\mathbb{R}], \mathtt{C} \rangle$ just if it is the image of some relation in $\mathcal{A}$ under all PL-homeomorphisms of the space to itself.

Using the other elementary equivalence results in [9] and an exactly analogous argument, we can prove the following:

Corollary 32 *Let $\mathcal{F}$ be a subfield of the reals, and let $\mathcal{P}[\mathcal{F}]$ be the domain of polyhedra over $\mathcal{F}$. In $\mathbb{E}^2$ or $\mathbb{E}^3$, a k-place relation $\Gamma(\mathbf{P}^1 \ldots \mathbf{P}^k)$ is definable in $\langle \mathcal{P}[\mathcal{F}], \mathtt{C} \rangle$ if and only if*

- Γ *is invariant under $\mathcal{F}$-PL-homeomorphisms of the space to itself; and*

- *the restriction of Γ to $\mathcal{P}[\mathbb{Q}]^k$ is in $\mathcal{A}$.*

Corollary 33 *Let $\mathcal{G}$ be the domain of bounded closed regular semi-algebraic regions. In $\mathbb{E}^2$ or $\mathbb{E}^3$, a k-place relation $\Gamma(\mathbf{P}^1 \ldots \mathbf{P}^k)$ is definable in $\langle \mathcal{G}, \mathtt{C} \rangle$ if and only if*

- Γ *is invariant under semi-algebraic homeomorphisms of the space to itself; and*

- *the restriction of Γ to $\mathcal{P}[\mathbb{Q}]^k$ is in $\mathcal{A}$.*

Corollary 34 *Let $\mathcal{O}$ be an o-minimal domain and let $\mathcal{D}$ be the class of bounded closed regular definable regions in $\mathcal{O}$. In $\mathbb{E}^2$ or $\mathbb{E}^3$, a k-place relation $\Gamma(\mathbf{P}^1 \ldots \mathbf{P}^k)$ is definable in $\langle \mathcal{O}, \mathtt{C} \rangle$ if and only if*

- Γ *is invariant under definable homeomorphisms of the space to itself; and*

- *the restriction of Γ to $[\mathcal{P}[\mathbb{Q}]]^k$ is in $\mathcal{A}$.*

See [9] for the details on all these.

4 Representing the k-cube using regular regions

We now turn to the domain $\mathcal{R}$ of regular regions.

In this section, we show how the k-cube in $\mathbb{E}^n$ for $k < n$ can be represented in $\langle \mathcal{R}, \mathsf{C} \rangle$. This serves as a "warm-up" to the much more general theorem 3 in the next section; the same basic construction is used in the proof, but the restricted setting here makes it easier to follow.

As above, let $\mathcal{R}$ be the domain of all non-empty, bounded, regular regions in $\mathbb{E}^n$, and let $\mathcal{R}'$ be the set of iregions. Let $\mathcal{S}'$ be the structure with domain $\mathcal{R}'$ containing all of the relations over iregions defined in section 2; as we have shown, $\mathcal{S}$ is definable in $\mathcal{R}, \mathsf{C}$.

Notational comment: In this section and in section 5, we will be dealing with three kinds of enumeration:

a. The coordinate of a k-dimensional vector (i.e. k-tuple of real numbers). For this we will use a vector symbol with a subscript; e.g. $\vec{v}_i$ is the ith coordinate of vector $\vec{v}$.

b. The jth element in a sequence (usually infinite) of vectors. For this, we will use a hatted symbol to denote the sequence, subscripted with j. For example $\hat{v}_j$ is the jth vector of the sequence $\hat{v}$.

This can be combined with (a) by combining the two subscripts with a comma. For example $\hat{v}_{j,i} = (\hat{v}_j)_i$, the ith coordinate of the jth vector in sequence $\hat{v}$.

c. We will need multiple sequences. For this we will use superscripts. Thus $\hat{w}^k$ is the kth sequence. This can be combined with (a) and (b); e.g. $\hat{w}_j^k$ is the jth vector in the kth sequence, and $\hat{w}_{j,i}^k$ is its ith coordinate.

The first question is, what is the class of iregions in $\mathcal{R}'$? The answer is pleasingly simple and general: Every compact point set is an iregion (lemma 36).

Lemma 35 *Let U be an open set and let point $\mathbf{p} \in \mathrm{Bd}(U)$. Then there exist two open sets V and W such that $V^- \setminus \{\mathbf{p}\} \subset U$, $W^- \setminus \{\mathbf{p}\} \subset U$, V and W are disjoint, and $\mathrm{Bd}(V) \cap \mathrm{Bd}(U) = \mathrm{Bd}(W) \cap \mathrm{Bd}(U) = \mathrm{Bd}(V) \cap \mathrm{Bd}(W) = \{\mathbf{p}\}$.*

Thus, the closures of V and of W except point $\mathbf{p}$ are disjoint and are inside the interior of U; and the boundaries of any pair of V, W, and U all meet at the single point $\mathbf{p}$.

Proof: Since U is open and $\mathbf{p} \in \mathrm{Bd}(U)$, we can find a sequence of points $\hat{u} = \langle \mathbf{u}_1, \mathbf{u}_2 \ldots \rangle$ in U that converges to $\mathbf{p}$. Let $d_i = d(\mathbf{u}_i, \hat{u} \setminus \{\mathbf{u}_i\})$ the distance from $\mathbf{u}_i$ to all the other points in $\hat{u}$; since $\mathbf{u}_i$ is not an accumulation point of $\hat{u}$, it follows that $d_i > 0$. Let $\epsilon_i = \min(d_i/2, 1/i)$ and let $U_i = B(\mathbf{u}_i, \epsilon_i) \cap U$; thus, all of the U_i are disjoint. Let V_i and W_i be two open sets such that $V_i^- \subset U_i$, $W_i^- \subset U_i$, and V_i^- and W_i^- are disjoint. Finally, let $V = \cup_i V_i$ and $W = \cup_i W_i$. It is then straightforward to verify that V and W meet the conditions of the lemma. (Figure 7). ∎

Lemma 36 *Every compact set in $\mathbb{E}^n$ is an iregion in $\mathcal{R}'$.*

Proof: Let C be a compact set in $\mathbb{E}^n$. We need to construct two regular regions $\mathbf{X}, \mathbf{Y}$ such that $\mathbf{X} \cap \mathbf{Y} = C$. We do this in two parts: first, the points that are in the regularization of C, which are easy, and then the points that are not, which require more work.

Let $\mathbf{R} = (C^\circ)^-$, the regularization of C. $\mathbf{R}$ is itself regular, unless it is empty, and $\mathbf{R} \subset C$ so we will simply include it in both $\mathbf{X}$ and $\mathbf{Y}$.
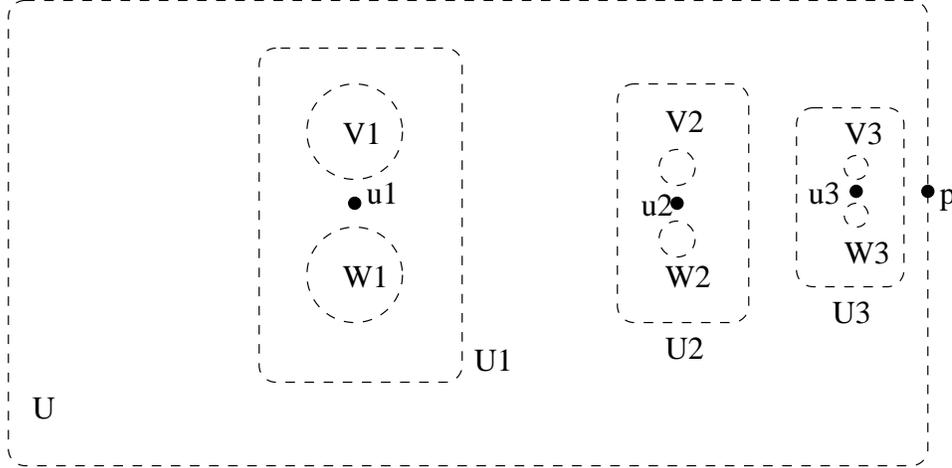
Figure 7: Construction for lemma 35

Let $\mathbf{p}_1, \mathbf{p}_2 \ldots$ be a countable sequence of distinct points that is dense in $C \setminus \mathbf{R}$. We will construct three sequences of open sets, U_i, X_i and Y_i. The U's will be setwise decreasing, and the X's and Y's will be increasing.

- Let S be a bounded open set containing C.

- Let $U_0 = S \setminus C$. Note that U_0 is open and that all the points $\mathbf{p}_i$ are on $\mathrm{Bd}(U_0)$.

- $X_0 = Y_0 = \emptyset$.

- for $i \geq 1$,

  - Let $Z_i = U_{i-1} \cap B(\mathbf{p}_i, 1/i)$. Using lemma 35 let V_i and W_i be open sets such that $V_i \subset Z_i$, $W_i \subset Z_i$, $\mathrm{Bd}(V_i) \cap \mathrm{Bd}(Z_i) = \mathrm{Bd}(W_i) \cap \mathrm{Bd}(Z_i) = \mathrm{Bd}(V_i) \cap \mathrm{Bd}(W_i) = \{\mathbf{p}_i\}$.

  - Set $X_i = X_{i-1} \cup V_i$; $Y_i = Y_{i-1} \cup W_i$; and $U_i = S \setminus (X_i^- \cup Y_i^- \cup C)$. Note that $X_i^- \cap C = Y_i^- \cap C = \{\mathbf{p}_j | j \leq i\}$ so all of the points $\mathbf{p}_j$ with $j > i$ are in U_i^-. Thus, the iteration can continue forever (or, if $C \setminus \mathbf{R}$ is a finite set of points, until all the points are used up.)

Now, let $\mathbf{X} = \mathrm{Closure}(\cup_i X_i)$ and let $\mathbf{Y} = \mathrm{Closure}(\cup_i Y_i)$. Then $\mathbf{X}$ and $\mathbf{Y}$ are regular as they are each the closure of an open set.

I claim that $C \setminus \mathbf{R} = \mathbf{X} \cap \mathbf{Y}$. On the one hand, each of the points $\mathbf{p}_i$ is in both $\mathbf{X}$ and $\mathbf{Y}$; hence $\mathbf{X} \cap \mathbf{Y} \supset \mathrm{Closure}(\{\mathbf{p}_1, \mathbf{p}_2 \ldots\}) = C \setminus \mathbf{R}$. Conversely, let $\mathbf{q}$ be a point in $\mathbf{X} \cap \mathbf{Y}$. Then

- $\mathbf{q}$ is either in V_i^- for some i or it is an accumulation point of a sequence $\mathbf{v}_i \in V_i, i = 1, 2 \ldots$

- $\mathbf{q}$ is either in W_i^- for some i or it is an accumulation point of a sequence $\mathbf{w}_i \in W_i, i = 1, 2 \ldots$

However: V_i^- and W_i^- intersect only at $\mathbf{p}_i$. For $i \neq j$, V_i^- and W_j^- do not intersect. And $V_i \subset B(\mathbf{p}_i, 1/i)$, so any accumulation point of the $\mathbf{v}_i$ must lie in $C \setminus \mathbf{R}$. Thus, one way and another, $\mathbf{q} \in C \setminus \mathbf{R}$.

Finally, letting $\mathbf{X} = \mathbf{X}' \cup \mathbf{R}$ and $\mathbf{Y} = \mathbf{Y}' \cup \mathbf{R}$, we have $C = \mathbf{X} \cap \mathbf{Y}$.

∎

Z is the union of the triangles. N is the union of the squares. D is the union of the circles.
This is not drawn to scale.
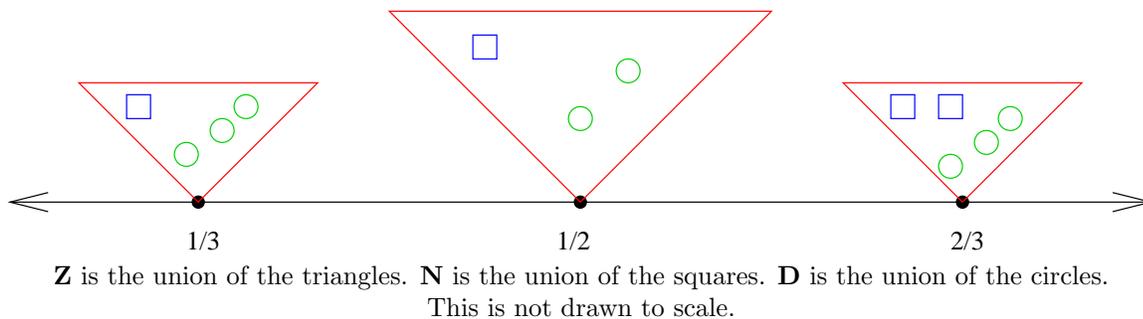
Figure 8: Labelled Comb for the 1-cube

Next we will define a general construction called a "labelled comb" which we can associate with any compact, lower-dimensional, point set. To illustrate how this works, we will first show how a particular comb can be created for the standard embedding of the $n-1$ cube in $\mathbb{E}^n$ and then generalize it.

Let $\mathbb{F}^{n-1}$ be the standard embedding in $\mathbb{E}^n$ of the $n-1$-dimensional cube: $\mathbb{F}^{n-1} = [0,1]^{n-1} \times \{0\}$. We begin by describing a certain region $\mathbf{Z}$, which we will call a "comb", consisting of infinitely many little cones, each of which touches $\mathbb{F}^{n-1}$ in a point with rational coefficients. The cones are right circular cones rising in the nth dimension. For any rational point $\mathbf{p} = \langle p_1 \ldots p_{n-1}, 0 \rangle$, let q be the largest denominator of any of the p_i expressed in lowest terms; then the cone rising from $\mathbf{p}$ goes to height $1/4q^2$. That is, the cone $\mathbf{C}$ from $\mathbf{p}$ is defined by the expression

$$\mathbf{C} = \{\mathbf{c} = \langle c_1 \ldots c_n \rangle \mid 0 \leq c_n \leq 1/4q^2 \wedge d(\mathbf{c}, \langle p_1 \ldots p_{n-1}, c_n \rangle) \leq c_n\}$$

One constructs such a cone over every rational point in $\mathbb{F}^{n-1}$ and then $\mathbf{Z}$ is the union of all these cone. The intersection of $\mathbf{Z}$ with a "horizontal" hyperplane $z = h$ consists of a collection of $n-1$ dimensional spheres of radius h centered over a rational point whose highest denominator q satisfies $1/4q^2 \geq h$. It is easily shown that no two such points can be closer than $1/q(q-1)$; hence, no two spheres at the same height meet; hence, no two cones meet. Therefore, each cone is an interior-connected component of $\mathbf{Z}$.

We now "label" each cone $\mathbf{C}$ with the numerators and denominators of the corresponding point $\{\langle p_1 \ldots p_{n-1}, 0 \rangle\} = \mathbf{C} \cap \mathbb{F}^{n-1}$ by creating $2(n-1)$ regions $\mathbf{N}_1 \ldots \mathbf{N}_{n-1}$ and $\mathbf{D}_1 \ldots \mathbf{D}_{n-1}$ such that the number of interior-connected components of $\mathbf{N}_i \cap \mathbf{C}$ and of $\mathbf{D}_i \cap \mathbf{C}$ are the numerator and denominator of p_i. Figure 8 shows the cones of the labelled comb with $q \leq 3$ for $n = 2$.

In order to attach a similar construction to an arbitrary lower-dimensional set, we need to generalize it slightly. The problem is that if C is a lower-dimensional compact set and $\mathbf{p} \in C$, there is not necessarily a interior-connected open subset of $\mathbb{E}^n \setminus C$ whose boundary contains $\mathbf{p}$. We fix this by replacing the "cone" meeting $\mathbf{p}$ by an infinite sequence of interior-connected regions in $\mathbb{E}^n \setminus C$ that converge on $\mathbf{p}$ and we label *each* of these with the desired label. Figure 9 illustrates the basic construction. Note that no connected open region with $\mathbf{p}$ on the boundary avoids C, but the union of the dotted regions does.

Definition 22 *Let $\mathbf{Z}$ be a region in $\mathcal{R}$ and let $\mathbf{L}_1 \ldots \mathbf{L}_{2k}$ be regions in $\mathcal{R}$. Let $\vec{m}$ be a k-dimensional vector of rationals; $\vec{m} = \langle r_1/d_1, r_2/d_2 \ldots r_k/d_k \rangle$, where r_i/d_i is in lowest terms. An interior-connected component $\mathbf{U}$ of $\mathbf{Z}$ is labelled by $\mathbf{L}_1 \ldots \mathbf{L}_{2k}$ with $\vec{v}$ if, for $i = 1 \ldots k$, $\mathbf{L}_i \cap \mathbf{U}$ has r_i interior-connected components and $\mathbf{L}_{k+i} \cap \mathbf{U}$ has d_i interior-connected components.*
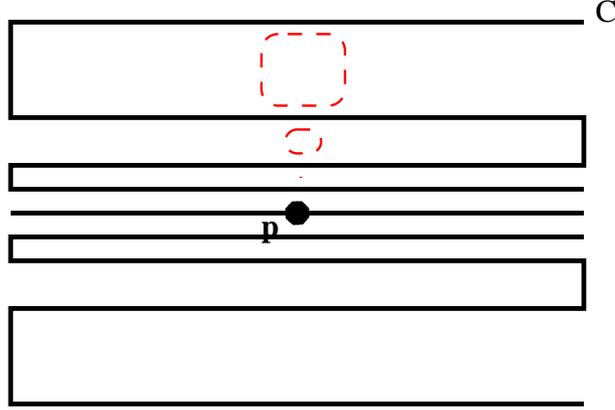
Figure 9: Constructing a labelled comb at points where no connected cone can be placed

The sector of $\mathbf{Z}$ labelled with $\vec{m}$ *is the closure of the union of the interior-connected components of* $\mathbf{Z}$ *labelled with* $\vec{m}$.

Lemma 37 *Let S be a lower-dimensional compact point set in $\mathbb{E}^n$. Let $\mathbf{p}_1, \mathbf{p}_2 \ldots$ be a countable sequence of disjoint points in S. Let $\hat{q}$ be a sequence of k-dimensional rational vectors. Then there exist regions $\mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{2k} \in \mathcal{R}$ satisfying the following: for each i, let $\mathbf{Y}(\hat{q}_i)$ be the sector of $\mathbf{Z}$ labelled by $\mathbf{L}_1 \ldots \mathbf{L}_{2k}$ with $\hat{q}_i$. Then $\mathbf{Y}(\hat{q}_i) \cap S = \mathbf{p}_i$.*

Proof: Carry out the same construction as in lemma 36. As constructed, the sequence of open sets $V_1, V_2 \ldots$ has the properties that

- for $i \neq j$, Closure(V_i) and Closure(V_j) are disjoint.

- for each i, Closure(V_i) $\cap S = \mathbf{p}_i$.

Let $\mathbf{Z} = \mathrm{Closure}(\cup_{i=1}^{\infty})V_i$.

Let $\hat{q}_{i,j} = r_{i,j}/d_{i,j}$ in lowest terms. For each interior-connected component U of V_i and for each $j = 1 \ldots k$, let $\mathbf{K}(U, j)$ and $\mathbf{K}(U, k+j)$ be a regular region that are subsets of U and that have $r_{i,j}$ components and $d_{i,j}$ components respectively. Let $\mathbf{L}_j$ be the closure of the unions of $\mathbf{K}(U, j)$ over all such U. Then the sector of $\mathbf{Z}$ labelled with $\hat{q}_i$, $\mathbf{Y}(\hat{q}^i)$ is just Closure(V_i). ∎

We now present a topological characterization of the labelling of the k-cube discussed above.

Definition 23 *Let S be a lower-dimension iregion in $\mathcal{R}'$. Let $\mathbf{Z}$ be a region in $\mathcal{R}$ and, for $i = 1 \ldots 2k$ let $\mathbf{L}_i$ be regions in $\mathcal{R}$.*

Let $\mathcal{Q} = \mathbb{C}^k \cap \mathbb{Q}^k$ be the set of k-tuples of rationals between 0 and 1.

We say that the tuple $\langle S, \mathbf{Z}, \mathbf{L}_1, \ldots, \mathbf{L}_{2k} \rangle$ is an iregion with a k-cube rational labelled comb *if the following hold.*

   A. *Let $\vec{m}$ be a vector in $\mathcal{Q}$. Let $\mathbf{Y}(\vec{m})$ be the sector of $\mathbf{Z}$ labelled by $\mathbf{L}_1 \ldots \mathbf{L}_{2k}$ with $\vec{m}$. Then:*

      A.i $\mathbf{Y}(\vec{m})$ *is non-empty.*

      A.ii $\mathbf{Y}(\vec{m}) \cap S$ *is a single point $\{\mathbf{p}\}$, which we will denote $\mathbf{p}(\vec{m})$*

For brevity we will say that $\mathbf{Y}(\vec{m})$ *and its components and* $\mathbf{p}(\vec{m})$ *are labelled with* $\vec{m}$.

B. $\mathbf{Z} = \cup_{\vec{m} \in \mathcal{Q}} \mathbf{Y}(\vec{m})$

C. *Let* $\mathbf{x}$ *be a point in* $\mathbb{E}^n$, *and suppose that for every neighborhood* $\mathbf{U}$ *of* $\mathbf{x}$ *there exists a vector* $\vec{m} \in \mathcal{Q}$ *such that* $\mathbf{p}(\vec{m}) \in \mathbf{U}$. *Then* $\mathbf{x} \in S$. *That is, S contains the closure of the set* $\{\mathbf{p}(\vec{m}) \mid \vec{m} \in \mathcal{Q}\}$.

D. *For every* $\vec{m} \in \mathcal{Q}$ *and every rational* $\delta > 0$, *there exists a region* $\mathbf{H}$ *such that for every* $\vec{q} \in \mathcal{Q}$, $\mathbf{p}(\vec{q}) \in \mathbf{H}^{\circ}$ *if and only if* $d(\vec{m}, \vec{q}) < \delta$. *We will call this region* $\mathbf{H}(\vec{m}, \delta)$.

E. *For every point* $\mathbf{x} \in S$ *and neighborhood* $\mathbf{U}$ *of* $\mathbf{x}$ *there exists a vector* $\vec{m} \in \mathcal{Q}$ *and a rational* $\delta > 0$ *such that* $\mathbf{x} \in \mathbf{H}(\vec{m}, \delta) \subset \mathbf{U}$.

F. *For any* $\vec{u} \neq \vec{v} \in \mathcal{Q}$ *there exist rational* $\delta > 0$ *such that* $\mathbf{H}(\vec{u}, \delta)$ *is disjoint from* $\mathbf{H}(\vec{v}, \delta)$.

Lemma 38 *For any fixed k, the relation "iregion with a k-cube rational labelled comb" is definable in the structure $\mathcal{S}'$.*

Proof: Let S be an iregion and let $\mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_k$ be regions.

Let $\vec{m}$ be a vector in $\mathcal{Q}$, represented as a $2k$-tuple of integers. The condition that all the components of $\vec{m}$ are between 0 and 1 is an arithmetic constraint. The relation "$\mathbf{Y}(\vec{m})$ is the sector of $\mathbf{Z}$ labelled with $\vec{m}$ by $\mathbf{L}_1 \ldots \mathbf{L}_k$" — i.e. "for every region $\mathbf{B}$, $\mathbf{B}$ is an interior connected component of $\mathbf{Y}(\vec{m})$ if and only if $\mathbf{B}$ is an interior-connected component of $\mathbf{Z}$ and $\mathbf{B}$ is labelled with $\vec{m}$ by $\mathbf{L}_1 \ldots \mathbf{L}_k$" — is definable in $\mathcal{S}'$ since interior-connected component and equality of natural numbers are definable.

The relation "point $\mathbf{p}$ in S is labelled with $\vec{m}$ by $\mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_k$" can be defined as "Point $\mathbf{p}$ is the intersection of S with the sector of $\mathbf{Z}$ labelled with $\vec{m}$ by $\mathbf{L}_1 \ldots \mathbf{L}_k$".

The region "$\mathbf{H}(\vec{m}, \delta)$" can be defined in terms of $S, \mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_m$ as follows: "$\mathbf{H}(\vec{m}, \delta)$ is a region; and for every vector $\vec{q} \in \mathcal{Q}$, $\mathbf{p}(\vec{q})$ is in the interior of $\mathbf{H}(\vec{m}, \delta)$ if and only if $d(\vec{q}, \vec{m}) < \delta$." Note that the inequality "$d(\vec{q}, \vec{m}) < \delta$" is just an arithmetic relation over rational numbers, representable as in section 2.3. This is not, of course, the geometric distance between points in $\mathbb{E}^n$, which is not expressible in our language; it is an arithmetic relation between two $2k$-tuples of natural numbers, which we instantiate in terms of numbers of interior-connected components, and interpret as k-tuples of rational numbers.

Using these relations, the translation of definition 23 into a first-order formula over $\mathcal{S}'$ is entirely straightforward. ∎

Definition 24 *S is a k-cube combable surface if there exists $\mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{2k}$ such that $\langle S, \mathbf{Z}, \mathbf{L}_1, \ldots, \mathbf{L}_{2k} \rangle$ is a surface with a k-cube rational labelled comb.*

Lemma 39 *If iregion S is homeomorphic to $\mathbb{C}^k$ then S is a k-cube combable surface.*

Proof: Let h be a homeomorphism from $\mathbb{C}^k$ to S. By lemma 37 there exist $\mathbf{Z}, \mathbf{L}^1 \ldots \mathbf{L}^{2k}$ such, for every rational vector $\vec{q} \in \mathbb{C}^k$ the point $h(\vec{q}) \in S$ is labelled with $\vec{q}$ by $\mathbf{Z}, \mathbf{L}^1 \ldots \mathbf{L}^{2k}$. Properties A-F of definition 23 are immediate from the fact that h is a homeomorphism. ∎

Lemma 40 *If S is a k-cube combable surface, then S is homeomorphic to the k-cube $\mathbb{C}^k$.*

Proof: Let $\langle S, \mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{2k} \rangle$ be a surface with a labelled comb. We construct the homeomorphism $h(\vec{x})$ from $\mathbb{C}^k$ to S as follows:

- If $\vec{x}$ is a rational vector, then define $h(\vec{x}) = \mathbf{p}(\vec{x})$

- Else, let $\hat{m}$ be a sequence of rational vectors converging to $\vec{x}$. Then $h(\vec{x})$ is the unique accumulation point of $h(\hat{m})$.

Essentially, the conditions in definition 23 assert directly that h is a homeomorphism. To be specific: to prove that h is a homeomorphism, we need to show that it is uniquely defined, a bijection, and continuous. Since $\mathbb{C}^k$ is compact, it follows that h is a homeomorphism.

h **is uniquely defined:**. Let $\hat{m}$ and $\hat{q}$ be sequences of rational vectors in $\mathcal{Q}$ that both converge to the same value $\vec{x}$. Since S is compact (all iregions are compact), the sequences $h(\hat{m})$ and $h(\hat{q})$ have accumulation points $\mathbf{a}$ and $\mathbf{b}$ respectively. Suppose that $\mathbf{a} \neq \mathbf{b}$. Choose disjoint neighborhoods A of $\mathbf{a}$ and B of $\mathbf{b}$. Using part (E) of definition 23, choose $\vec{u}, \delta_u, \vec{v}, \delta_v$ so that $\mathbf{a} \in H(\vec{u}, \delta_u) \subset A$ and $\mathbf{b} \in H(\vec{v}, \delta_v) \subset B$. Let $0 < \delta < \min(\delta_u, \delta_v, d(\vec{u}, \vec{v})/4)$. Then either $d(\vec{x}, \vec{u})$ or $d(\vec{x}, \vec{v})$ is greater than 2δ; assume w.l.o.g that $d(\vec{x}, \vec{v}) > 2\delta$. Since all but finitely many of the values $\hat{q}_1, \hat{q}_2$ are within δ of $\vec{x}$, it follows that all but finitely many are more than δ from $\vec{v}$. But then by definition of $\mathbf{H}(\vec{v}, \delta)$ (part D), all but finitely many of $h(\hat{q}_1), h(\hat{q}_2) \ldots$ are outside $\mathbf{H}(\vec{v}, \delta)$, contradicting the assumption that $\mathbf{b}$ is an accumulation point.

h **is an injection:** Let $\vec{x} \neq \vec{y} \in \mathbb{C}^k$. Choose rational vectors $\vec{u}, \vec{v}$ so that $d(\vec{x}, \vec{u}) < d(\vec{x}, \vec{y})/4$ and $d(\vec{y}, \vec{v}) < d(\vec{x}, \vec{y})/4$. Choose δ to satisfy part (F) of definition 23. Let $\hat{m}$ and $\hat{q}$ be sequences of rational vectors converging to $\vec{x}$ and $\vec{y}$ respectively. Then all but finitely many of the $\hat{m}_i$ are more than δ from $\vec{v}$ and all but finitely many of the $\hat{q}_i$ are more than δ from $\vec{u}$. Since $h(\vec{x}) = \lim_{i \to \infty} h(\hat{m}_i) \in \mathbf{H}(\vec{u}, \delta)$ and $h(\vec{y}) = \lim_{i \to \infty} h(\hat{q}_i) \in \mathbf{H}(\vec{v}, \delta)$ and since $\mathbf{H}(\vec{u}, \delta)$ and $\mathbf{H}(\vec{v}, \delta)$ are disjoint, we must have $h(\vec{x}) \neq h(\vec{y})$.

h **is a surjection:** Let $\mathbf{p} \in S$. Using part (C), we can find a sequence $\hat{m}$ of rational vectors in $\mathcal{Q}$, such that $h(\hat{m})$ converges to $\mathbf{p}$. Let $\vec{x}$ be an accumulation point in $\mathbb{C}^k$ of $\hat{m}$ then $\mathbf{p} = h(\vec{x})$.

h **is continuous.** Let $\mathbf{p} \in S$ and let $\mathbf{U}$ be a neighborhood of $\mathbf{p}$. Using part (E) choose rational $\vec{v}$ and δ such that $\mathbf{p} \in \mathbf{H}(\vec{v}, \delta) \subset \mathbf{U}$. Let V be the open ball in $\mathbb{C}^k$, $V = \vec{B}(\vec{v}, \delta)$. Then by definition of $\mathbf{H}$, $h(V) \subset \mathbf{H}(\vec{v}, \delta) \subset U$.

∎

Theorem 2 *The property "Iregion P is homeomorphic to the k-cube" is definable in $\mathcal{S}'$.*

Proof: Immediate from lemmas 38, 39, and 40. ∎

Corollary 41 takes this back to the original structure $\langle \mathcal{R}, \mathtt{C} \rangle$.

Corollary 41 *The property "The intersection of regions $\mathbf{A}$ and $\mathbf{B}$ is homeomorphic to the k-cube," is definable in $\langle \mathcal{R}, \mathtt{C} \rangle$.*

5   Representing analytical relations

In this section we prove a very strong generalization of theorem 2.

Theorem 3: *Any analytical relation over lower-dimensional compact point sets that is invariant over homeomorphisms can be expressed in $\langle \mathcal{R}', \mathtt{C} \rangle$.*

We begin by defining the terminology of theorem 3.

Definition 25 *Let* $(\mathbb{R}^n)^{\mathbb{N}}$ *be the set of all countable sequences of vectors in* $\mathbb{R}^n$. *Let* $\Gamma(\hat{v}^1 \ldots \hat{v}^k)$ *be a k-place relation over* $(\mathbb{R}^n)^{\mathbb{N}}$. *Let* $\alpha(\hat{v})$ *be the injection mapping a sequence of vectors* $\hat{v}$ *to a sequence of integers described in section 2.3. Let* Γ' *be the k-place relation over* $\mathbb{N}^{\mathbb{N}}$ *defined by*

$$\Gamma'(\hat{w}^1 \ldots \hat{w}^k) \equiv \exists_{\hat{v}^1 \ldots \hat{v}^k} \ \hat{w}^1 = \alpha(\hat{v}^1) \wedge \ldots \wedge \hat{w}^k = \alpha(\hat{v}^k) \wedge \Gamma(\hat{v}^1 \ldots \hat{v}^k)$$

Then Γ *is said to be analytical if* Γ' *is analytical.*

We now define a mapping $\nabla(\Gamma)$ which maps a k-place relation Γ over the space of compact geometric region to a k-place relation over sequences of real numbers. The symbol $\mathbb{W}$ will denote the collection of all compact point sets of $\mathbb{E}^n$.

Definition 26 *Let* $e : \mathbb{E}^n \mapsto \mathbb{R}^n$ *be the standard coordinate system. Let* Γ *be a k-place relation over* $\mathbb{W}$. *We define* $\nabla(\Gamma)$ *to be the k-place relation* Δ *over* $(\mathbb{R}^n)^{\mathbb{N}}$ *satisfying the following. For any sequences of n-vectors* $\hat{w}^1 \ldots \hat{w}^k$, $\Delta(\hat{w}^1 \ldots \hat{w}^k)$ *holds if and only* $\Gamma(\text{Closure}(\hat{w}^1) \ldots \text{Closure}(\hat{w}^k))$.

Definition 27 *A k-place relation* Γ *over* $\mathbb{W}$ *is* analytical *if* $\nabla(\Gamma)$ *is analytical.*

For example we can show that the relation over compact point sets "A is homeomorphic to B" is analytical. The construction is analogous to the proof (p. 16) that PL-homeomorphism over polyhedra is arithmetic. Let h be a homeomorphism from compact A to B. Then h can be specified by specifying the values of $\vec{x}$ and $h(\vec{x})$ for $\vec{x} \in S$ where S is a countable dense subset of A. Thus, effectively, h can be construed as an element of $\mathbb{R}^{\mathbb{N}}$, a countable sequence of reals. Therefore given $\hat{w}^A$, a sequence of points dense in A, and $\hat{w}^B$, a sequence of points dense in B, the relation "A is homeomorphic to B" corresponds to the following condition on $\hat{w}^A$ and $\hat{w}^B$: There exists a sequence $\hat{w}^h$ (intuitively, the values of $h(\vec{w})$ for $\vec{w} \in \hat{w}^A$) satisfying the following:

1. Let $p_1, p_2 \ldots$ and $q_1, q_2, \ldots$ be sequences of indices. Then $\lim_{i \to \infty} \hat{w}^A_{p_i} = \lim_{j \to \infty} \hat{w}^A_{q_j}$. if and only if $\lim_{i \to \infty} \hat{w}^h_{p_i} = \lim_{j \to \infty} \hat{w}^h_{q_j}$. (This corresponds to the conditions that h is one-to-one and continuous and has a continuous inverse.)

2. The closure of $\hat{w}^h$ is equal to the closure of $\hat{w}^B$ (viz. B)

It is easy to show that both (1) and (2) are definable in the structure $\langle \mathbb{R} \cup \mathbb{R}^{\mathbb{N}}, +, *, [] \rangle$.

Definition 28 *A k-place relation* Γ *over* $\mathbb{W}$ *is* invariant under homeomorphisms *if the following holds: Let* $C^1 \ldots C^k \in \mathbb{W}$ *be compact sets satisfying* Γ. *Let h be a homeomorphism from* $C^1 \cup \ldots \cup C^k$ *to some iregion* $U \in \mathbb{W}$. *Then* $\Gamma(h(C^1) \ldots h(C^k))$.

We now show how to construct a labelled comb for any lower-dimensional, compact point set S in $\mathbb{R}^n$. We will choose an countable dense subset of points in S, and label each point with its n-dimensional coordinates.

Definition 30 generalizes definition 23. It defines a "labelled comb" that can be constructed on any lower-dimensional region S and that characterizes S up to homeomorphism.

Definition 29 *Let* $\mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{3n}$ *be regions in* $\mathcal{R}$ *and let* $\vec{x} \in \mathbb{R}^n$ *be an n-dimensional vector. An interior-connected component* $\mathbf{U}$ *of* $\mathbf{Z}$ *is labelled by* $\mathbf{L}_1 \ldots \mathbf{L}_{3n}$ *with* $\vec{x}$ *if, for* $i = 1 \ldots n$, *the triple* $\mathbf{L}_i \cap \mathbf{U}$, $\mathbf{L}_{i+n} \cap \mathbf{U}$, $\mathbf{L}_{i+2n} \cap \mathbf{U}$ *is an encoding of the real number* $\vec{x}_i$ *as described in section 2.3.*

The sector *of* $\mathbf{Z}$ *labelled with* $\vec{x}$ *is the closure of the union of the interior-connected components of* $\mathbf{Z}$ *labelled with* $\vec{x}$.

Lemma 42 *Let S be a lower-dimensional compact point set in $\mathbb{E}^n$. Let $\mathbf{p}_1, \mathbf{p}_2 \ldots$ be a countable sequence of disjoint points in S. Let $\hat{q}$ be a sequence of n-dimensional vectors. Then there exist regions $\mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{3n} \in \mathcal{R}$ satisfying the following: For each i, let $\mathbf{Y}(\hat{q}_i)$ be the sector of $\mathbf{Z}$ labelled by $\mathbf{L}_1 \ldots \mathbf{L}_{3n}$ with $\hat{q}_i$; then $\mathbf{Y}(\hat{q}_i) \cap S = \mathbf{p}_i$.*

Proof: Exactly analogous to the proof of lemma 37.

Definition 30 *Let S be a lower-dimensional point set in $\mathbb{E}^n$. Let $\hat{w}$ be a countable sequence of n-dimensional vectors. Let $\mathbf{Z}$ be a region in $\mathcal{R}$ and, for $i = 1 \ldots 3n$ let $\mathbf{L}_i$ be regions in $\mathcal{R}$.*

We say that the tuple $\langle S, \mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{3n} \rangle$ is an iregion with a labelled comb based on $\hat{w}$ if the following conditions hold.

- A. *Let $\vec{m}$ be a vector in $\hat{w}$. Let $\mathbf{Y}(\vec{m})$ be the sector of $\mathbf{Z}$ labelled by $\vec{m}$. Then:*

    A.i $\mathbf{Y}(\vec{m})$ *is non-empty.*

    A.ii $\mathbf{Y}(\vec{m}) \cap S$ *is a single point $\{\mathbf{p}\}$, which we will denote $\mathbf{p}(\vec{m})$.*

    For brevity we will say that $\mathbf{Y}(\vec{m})$ and its components and $\mathbf{p}(\vec{m})$ are labelled with $\vec{m}$.

- B. $\mathbf{Z} = \cup_{\vec{m} \in \hat{w}} \mathbf{Y}(\vec{m})$

- C. *Let $\mathbf{x}$ be a point in $\mathbb{E}^n$, and suppose that for every neighborhood $\mathbf{U}$ of $\mathbf{x}$ there exists a vector $\vec{m} \in \hat{w}$ such $\mathbf{p}(\vec{m}) \in \mathbf{U}$. Then $\mathbf{x} \in S$. That is, S contains the closure of the set $\{\mathbf{p}(\vec{m}) \,|\, \vec{m} \in \hat{w}\}$.*

- D. *For every $\vec{m} \in \hat{w}$ and every $\delta > 0$, there exists a region $\mathbf{H}$ such that for every $\vec{q} \in \hat{w}$, $\mathbf{p}(\vec{q}) \in \mathbf{H}^\circ$ if and only if $d(\vec{m}, \vec{q}) < \delta$. We will call this region $\mathbf{H}(\vec{m}, \delta)$.*

- E. *For every point $\mathbf{x} \in S$ and neighborhood $\mathbf{U}$ of $\mathbf{x}$ there exists a vector $\vec{m} \in \hat{w}$ and a $\delta > 0$ such that $\mathbf{x} \in \mathbf{H}(\vec{m}, \delta) \subset \mathbf{U}$.*

- F. *For any $\vec{u} \neq \vec{v} \in \hat{w}$ there exist $\delta > 0$ such that $\mathbf{H}(\vec{u}, \delta)$ is disjoint from $\mathbf{H}(\vec{v}, \delta)$.*

Definition 31 *Let $S^1 \ldots S^k$ be compact lower-dimensional subsets of $\mathbb{E}^n$. Let $\hat{w}^1 \ldots \hat{w}^k \in (\mathbb{R}^n)^{\mathbb{N}}$ be countable sequences of n-vectors. Let $\mathbf{Z}$, $\mathbf{L}_i$ for $i = 1 \ldots 3n$ and $\mathbf{B}^j$ for $j = 1 \ldots k$ be regions in $\mathcal{R}$. (The region $\mathbf{B}^j$ serve as labels on each labelled point to specify whether it is part of S^j.) Let $\mathbf{Z}^j$ be the closure of interior-connected components U of $\mathbf{Z}$ such that $U \cap \mathbf{B}^j$ is non-empty.*

We say that the tuple $\langle S^1 \ldots S^k, \mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{3n}, \mathbf{B}^1 \ldots \mathbf{B}^k \rangle$ is a set of iregions with labelled combs based on $\hat{w}^1 \ldots \hat{w}^k$ if:

- $\mathbf{Z} = \cup_{j=1}^{k} \mathbf{Z}^j$.

- *for $j = 1 \ldots k$, $\langle S^j, \mathbf{Z}^j, \mathbf{L}_1 \ldots \mathbf{L}_{3n} \rangle$ is an iregion with a labelled comb based on $\hat{w}^j$.*

Definition 32 *Let Φ be a k-place relation over $(\mathbb{R}^n)^{\mathbb{N}}$. Let $S^1 \ldots S^k$ be lower-dimensional point sets in $\mathbb{E}^n$. We say that $S^1 \ldots S^k$ is Φ-combable if there exist $\hat{w}^1 \ldots \hat{w}^k \in (\mathbb{R}^n)^{\mathbb{N}}$, $\mathbf{Z}$, $\mathbf{L}_i$ for $i = 1 \ldots 3n$ and $\mathbf{B}^j$ for $j = 1 \ldots k$ such that $\Phi(\hat{w}^1 \ldots \hat{w}^k)$ and $\langle S^1 \ldots S^k, \mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{3n}, \mathbf{B}^1 \ldots \mathbf{B}^k \rangle$ is a set of iregions with labelled combs based on $\hat{w}^1 \ldots \hat{w}^k$.*

Lemma 43 *Let Φ be an analytical relation over $(\mathbb{R}^n)^{\mathbb{N}}$. Then the k-place relation over compact lower-dimensional point sets, "$\langle S^1 \ldots S^k \rangle$ is Φ-combable" is expressible in $\mathcal{S}'$.*

Proof: Exactly analogous to the proof of lemma 38.

Lemma 44 *Let $\langle S^1 \ldots S^k, \mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{3n}, \mathbf{B}^1 \ldots \mathbf{B}^k \rangle$ be a set of iregions with labelled combs based on $\hat{w}^1 \ldots \hat{w}^k$. Let $S = \cup_{i=1}^k S^i$ and let $\hat{w} = \cup_{i=1}^k \hat{w}^i$. Then $\langle S, \mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_n \rangle$ is an iregion with labelled comb over $\hat{w}$.*

Proof: Immediate from the definition. ∎

Lemma 45 *Let $\hat{w}$ be a bounded countable subset of $\mathbb{R}^n$ and let $S \subset \mathbb{E}^n$ homeomorphic to $\mathrm{Closure}(\hat{w})$. Then there exist $\mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{3n}$ such that $\langle S, \mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{3n} \rangle$ is an iregion with labelled comb based on $\hat{w}$.*

Proof: Let h be a homeomorphism from $\mathrm{Closure}(\hat{w})$ to S. Using lemma 42, construct a comb $\langle S, \mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{3n} \rangle$ where for every $\vec{m} \in \hat{w}$, the point $h(\vec{m})$ is labelled with $\vec{m}$. Then it is straightforward to check that the conditions of definition 30 are satisfied. ∎

Lemma 46 *Let $\hat{w}^1 \ldots \hat{w}^k \in (\mathbb{R}^n)^{\mathbb{N}}$ be sequences of vectors. Let $S^1 \ldots S^k$ be compact point sets in $\mathbb{E}^n$. Let $\hat{w} = \cup_{i=1}^k \hat{w}^i$ and let $S = \cup_{i=1}^k S^i$. If there exists a homeomorphism h from S to $\mathrm{Closure}(\hat{w})$ such that $h(S^i) = \mathrm{Closure}(\hat{w}^i)$ for for $i = 1 \ldots k$, then there exist $\mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{3n}, \mathbf{B}^1 \ldots \mathbf{B}^k$ such that $\langle S^1 \ldots S^k, \mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{3n}, \mathbf{B}^1 \ldots \mathbf{B}^k \rangle$ is a set of iregions with labelled combs based on $\hat{w}_1 \ldots \hat{w}_k$.*

Proof: Immediate from lemma 45.

Lemma 47 *Let $\langle S^1 \ldots S^k, \mathbf{Z}, \mathbf{L}_1 \ldots \mathbf{L}_{3n}, \mathbf{B}^1 \ldots \mathbf{B}^k \rangle$ be a set of iregions with labelled combs based on $\hat{w}^1 \ldots \hat{w}^k$. Then there is a homeomorphism h from $\mathrm{Closure}(\cup_{i=1}^k \hat{w}^i)$ to $\cup_{i=1}^k S^i$ such that for $i = 1 \ldots k$, $h(\mathrm{Closure}(\hat{w}^i)) = S^i$.*

Proof: Analogous to the proof of lemma 40. ∎

Theorem 3 *Any analytical relation Γ over compact lower-dimensional regions that is invariant over homeomorphisms can be expressed in $\mathcal{S}'$.*

Proof: This is simply a matter of putting the pieces together.

Let $\Delta = \nabla(\Gamma)$, as defined in definition 26. By definition 26 Δ is analytic, so by lemma 43, the relation "$\langle Q^1 \ldots Q^k \rangle$ is Δ-combable" is definable in $\mathcal{S}'$.

Let $e : \mathbb{E}^n \to \mathbb{R}^n$ be the standard coordinate system. Let $Q^1 \ldots Q^k$ be any k iregions. Let $Q = \cup_{i=1}^k Q^i$. Then the following are equivalent:

A. $\Gamma(Q^1 \ldots Q^k)$.

B. There exist $\hat{w}^1 \ldots \hat{w}^k \in (\mathbb{R}^n)^{\mathbb{N}}$ such that for $i = 1 \ldots k$, $Q^i = \mathrm{Closure}(e^{-1}(\hat{w}^i))$ and $\Delta(\hat{w}^1 \ldots \hat{w}^k)$.

C. There exist $\hat{v}^1 \ldots \hat{v}^k \in (\mathbb{R}^n)^{\mathbb{N}}$ and a homeomorphism h from $\mathrm{Closure}(\cup_{i=1}^k \hat{v}^i)$ to Q such that $h(\mathrm{Closure}(\hat{v}^i)) = Q^i$ and $\Delta(\hat{v}^1 \ldots \hat{v}^k)$.

D. $\langle Q^1 \ldots Q^k \rangle$ is Δ-combable

(B) is equivalent to (A) by definition 26. (B) implies (C) because one can choose h to be the restriction of e to $\mathrm{Closure}(\cup_i \hat{v}^i)$.

To show that (C) implies (B), assume that we are given such h and $\hat{v}^1 \ldots \hat{v}^k$. Choose $\hat{w}_i = e(h(\hat{v}^i))$. For $i = 1 \ldots k$ let $D^i = \mathrm{Closure}(e^{-1}(\hat{v}_i)) = e^{-1}(h^{-1}(Q^i))$. Then the composition $e^{-1} \circ h^{-1}$ is a homeorphism from Q to $\cup_i D^i$. Using the definition of Δ and the fact that Γ is invariant under homeomorphism, we have that $\Delta(\hat{w}^1 \ldots \hat{w}^k)$ if and only if $\Gamma(Q^1 \ldots Q^k)$ if and only if $\Gamma(D^1 \ldots D^k)$ if and only if $\Delta(\hat{v}^1 \ldots \hat{v}^k)$.

(D) follows from (C) by lemmas 47 and 46.

Thus the condition $\Gamma(Q^1 \ldots Q^k)$ is definable in the structure $\langle \mathcal{R}', \mathtt{C} \rangle$. ∎

6 Other languages

Combining the analysis developed in this paper with that in [8] and [9], we are able to give a substantial part of an answer to a challenge recently raised by Hahmann and Gruninger [15]. They write (p. 47)

> Of particular interest is the question whether other theories [i.e. first-order theories with quantification over regions] exists between mereotopology and mereogeometries. The RCC with convexity is so far the weakest theory extending pure mereotopology with some morphological predicate, but perhaps, other predicates such as relative size can be used to extend mereotopology without reaching full mereogeometry or even the equivalent of affine geometry. This question remains unexplored as far as we know and needs to be looked at more carefully in the future.

We can approach this problem along the following lines. Suppose that Ω is some "reasonable" domain of regions, e.g. $\mathcal{P}[\mathbb{Q}]$, $\mathcal{P}[\mathbb{R}]$ or $\mathcal{R}$. Suppose further that we have some collection of relations over regions $\Phi_1, \Phi_2 \ldots \Phi_k$; we will assume that Φ_1 is in fact the relation $\mathtt{C}(\mathtt{x}, \mathtt{y})$ since we agree with the implicit assumption of the quoted passage that theories that do not at least include mereotopology are not worth considering. Suppose finally that all these relations are not off the charts in terms of computational/logical complexity; specifically, if the domain of regions is either $\mathcal{P}[\mathbb{Q}]$ then these are all arithmetic relations and if the domain is $\mathcal{P}[\mathbb{R}]$ or $\mathcal{R}$ then they are all analytical relations. It will also be convenient to define Φ_0 as the unary relation of being an element of Ω: $\Phi_0 = \{\langle x \rangle | x \in \Omega\}$. Let Δ be the structure $\langle \Omega, \mathtt{C}, \Phi_2 \ldots \Phi_k \rangle$. The problem then is to characterize the relations that can be defined in Δ.

Following [9] let $\Theta(\Phi_0 \ldots \Phi_k)$ be the class of all homeomorphisms of $\mathbb{E}^n$ to itself for which the relations $\Phi_0 \ldots \Phi_k$ are all invariants; this is clearly a *subgroup* of the space of all homeomorphisms of $\mathbb{E}^n$ to itself. Then we can certainly make the following claims:

1. Any relation that is definable in Δ is invariant under $\Theta(\Phi_0 \ldots \Phi_k)$. The proof is analogous to that of lemma 26.

2. If $\Omega = \mathcal{P}[\mathbb{Q}]$ and $\Phi_0 \ldots \Phi_k$ are all arithmetic relations, then any relation that is definable in Δ is arithmetic. The proof is analogous to lemma 27.

3. If $\Omega = \mathcal{P}[\mathbb{R}]$ or $\Omega = \mathcal{R}$ and $\Phi_0 \ldots \Phi_k$ are all analytical relations, then any relation that is definable in Δ is analytical. The proof is essentially analogous to lemma 27.

Let $\Delta' = \langle \Omega', \mathtt{C}, \Phi'_2 \ldots \Phi'_m \rangle$ be some different structure over $\mathbb{E}^n$.. If Φ'_i or membership in Ω' is not invariant under $\Theta(\Phi_0 \ldots \Phi_k)$ then, by the contrapositive of (1) there are relations definable in Δ'

that are not definable in Δ (viz. Φ'_i itself), so Δ is certainly not more expressive than Δ'.

For example, following the suggestion of Hahmann and Gruninger, let us consider the relation `LargerVolume(x,y)`, meaning that the volume of x is greater than the volume of y. It is easily shown that the corresponding group of homeomorphisms $\Theta(\Omega, \mathtt{C}, \mathtt{LargerVolume})$ is the class of homeomorphisms that multiply the volume of every region by a fixed constant $c > 0$. Over the domains $\mathcal{P}[\mathbb{Q}]$ and $\mathcal{P}[\mathbb{R}]$, this is the class of PL-homeomorphisms where either every cell has determinant c or every cell has determinant $-c$. Over the domain $\mathcal{R}$, this is harder to characterize; it certainly includes, but is not limited to, all homeomorphisms of $\mathbb{E}^n$ to itself that are piecewise differentiable, and such that the Jacobian on the interior of every cell is either always c or always $-c$. Therefore, any predicate definable in the structure $\langle \Omega, \mathtt{C}, \mathtt{LargerVolume} \rangle$ is invariant under the corresponding group of volume-ratio preserving homeomorphisms.

In [8] we were able to completely characterize the expressivity of the structures $\langle \Omega, \mathtt{C}, \mathtt{Convex} \rangle$ and $\langle \Omega, \mathtt{C}, \mathtt{Closer} \rangle$, where Ω can be either $\mathcal{P}[\mathbb{Q}]$, $\mathcal{P}[\mathbb{R}]$ or $\mathcal{R}$, in $\mathbb{E}^n$ for $n \geq 2$. `Closer(x,y,z)` is the ternary relation "x is closer to y than to z". `Convex(x)` is the unary relation "x is convex".

- In $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C}, \mathtt{Closer} \rangle$, a relation is definable iff it is arithmetic and invariant under Euclidean transformations.

- In $\langle \mathcal{P}[\mathbb{R}], \mathtt{C}, \mathtt{Closer} \rangle$, and in $\langle \mathcal{R}, \mathtt{C}, \mathtt{Closer} \rangle$, a relation is definable iff it is analytical and invariant under Euclidean transformations.

- In $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C}, \mathtt{Convex} \rangle$, a relation is definable iff it is arithmetic and invariant under affine transformations.

- In $\langle \mathcal{P}[\mathbb{R}], \mathtt{C}, \mathtt{Convex} \rangle$, and in $\langle \mathcal{R}, \mathtt{C}, \mathtt{Convex} \rangle$, a relation is definable iff it is analytical and invariant under affine transformations.

Since volume ratios are preserved under affine transformations, but affine properties such as `Convex` are not preserved under volume-ratio preserving transformations, it follows that $\langle \Omega, \mathtt{C}, \mathtt{Convex} \rangle$ is strictly more expressive than $\langle \Omega, \mathtt{C}, \mathtt{LargerVolume} \rangle$, for any of the domains Ω under consideration.

We can carry out a similar analysis for other relations. For instance, consider the relation `Smooth(x)`, meaning that region x has an everywhere differentiable boundary. Over the domain of polyhedra, this is the null relation, since no polyhedra are smooth, but over the domain $\mathcal{R}$, the associated group of homeomorphisms $\Theta(\mathcal{R}, \mathtt{C}, \mathtt{Smooth})$ is just the class of diffeomorphisms of $\mathbb{E}^n$ to itself. The relation `Smooth` is not invariant under under volume-ratio preserving transformations, though it is invariant under affine transformations; and conversely the relation `LargerVolume` is not invariant under diffeomorphisms. We can conclude, therefore that expressivity of the structure $\langle \mathcal{R}, \mathtt{C}, \mathtt{Smooth} \rangle$ is incomparable to the expressivity of the structure $\langle \mathcal{R}, \mathtt{C}, \mathtt{LargerVolume} \rangle$ and strictly less than the expressivity of the structure $\langle \mathcal{R}, \mathtt{C}, \mathtt{Convex} \rangle$.

We conclude this section by posing a new challenge in the spirit of Hahmann and Gruninger's:

> Can we give general conditions for relations $\Phi_2 \ldots \Phi_k$ that suffice to guarantee the statement, "A relation is definable in the structure $\langle \Omega, \mathtt{C}, \Phi_2 \ldots \Phi_k \rangle$ if it is [arithmetic/analytical] and invariant under $\Theta(\Phi_0 \ldots \Phi_k)$?" We have proven this for the particular relations listed above; are these special cases of a more general theorem?

7 Conclusions

We have given necessary and sufficient conditions for a property over polyhedra to be definable over $\langle \mathcal{P}[\mathbb{Q}], \mathtt{C} \rangle$ in two- and three-dimensional space, and we have proven a strong result about the

expressivity of $\langle \mathcal{R}, \texttt{C} \rangle$ in characterizing relations over lower-dimensional point sets in n-dimensional space.

Certainly, there is room for improvement. As regards theorem 1, the most obvious improvement would be to generalize to $\mathbb{E}^n$ for $n > 3$. The main difficulty is generalizing lemmas 14 and 15 to the n-dimensional case. This seems to be difficult; the topology of $\mathbb{E}^n$ for $n > 3$, especially for $n = 4$ is full of strange anomalies, and the techniques used to prove these theorems do not at all generalize.

As regards theorem 3, it would obviously be desirable to have a theorem that includes fully dimensional regions, and that strengthens "homeomorphic" to "similarly situated", and that gives a biconditional. These objectives are, in fact, rather closely related. Since a regular region is largely, though not entirely, determined by its boundary, and since theorem 3 allows us to characterize the boundary of a region up to homeomorphism, it seems as though there might be some way to strengthen this from a characterization of the boundary to a characterization of the whole region, but we have not found it.

A final source of discontent is that, though definition 17 of an arithmetic relation over rational polyhedra and definition 27 of an analytical relation compact point sets each seems reasonable enough individually, they are rather *ad-hoc* and different in flavor. One would prefer to have a more principled or systematic way of associating a complexity category with a property of point set relations.

References

[1] B. Bennett. A Categorical Axiomatisation of Region-Based Geometry. *Fundamenta Informaticae,* **46**, 145-158. 2001.

[2] B. Bennett, A.G. Cohn, P. Torini, and S. Hazarika. A foundation for region-based geometry. *Proceedings of the Fifteenth European Conference on Artificial Intelligence.* 204-208. 2000.

[3] B. Bennett, A.G. Cohn, P. Torini, and S. Hazarika. Describing rigid-body motion in a qualitative theory of spatial regions. *Proceedings, Seventeenth National Conference on Artificial Intelligence,* 503-509. 2000.

[4] S. Borgo, N. Guarino, and C. Masolo. A Pointless Theory of Space Based on Strong Connection and Congruence. *Fifth International Conference on Principles of Knowledge Representation and Reasoning.* 220-229. 1996.

[5] S. Borgo and C. Masolo. Mereogeometries: a Semantic Comparison. Tech. Report 02/10, LADESB-CNR, Padova, Italy, January 2001.

[6] B.L. Clarke. A Calculus of Individuals Based on "Connection". *Notre Dame Journal of Formal Logic,* **23**(3), 204-218. 1981.

[7] A.G. Cohn and J. Renz, Qualitative Spatial Reasoning, in F. van Harmelen, V. Lifschitz, and B. Porter (eds.) *Handbook of Knowledge Representation,* Elsevier, 2007.

[8] E. Davis. The Expressivity of Quantifying over Regions. *Journal of Logic and Computation,* vol. 16, 2006, 891-916.

[9] E. Davis, Elementary Equivalent Structures for Topological Languages over Regions in Euclidean Space. *Journal of Logic and Computation,* to appear.

[10] T. de Laguna. Point, line, and surface as sets of solids. *Journal of Philosophy,* **19**, 449-461. 1922.

[11] C. Dornheim. Undecidability of plane polygonal mereotopology. *Sixth International Conference on Principles of Knowledge Representation and Reasoning.* 342-353. 1998.

[12] G. Gerla. Pointless Geometries. In F. Buekenhout (ed.) *Handbook of Incidence Geometry,* Elsevier, 1995, 1015-1031.

[13] N.M. Gotts. How far can we 'C'? Defining a doughnut using connection alone. *Fourth International Conference on Principles of Knowledge Representation and Reasoning.* 246-257. 1994.

[14] A. Grzegorczyk. Undecidability of some topological theories. *Fundamenta Mathematicae.* **38**, 137-152. 1951.

[15] T. Hahmann and M. Gruninger. Region-Based Theories of Space: Mereotopology and Beyond. In S. Harazika (ed.) *Qualitative Spatio-Temporal Representation and Reasoning: Trends and Future Direction,* IGI Global, 2012, 1-62,

[16] P.G. Hinman. *Recursion Theoretic Hierarchies.* Springer-Verlag, Berlin. 1978.

[17] W. Hodges, *Model Theory,* Cambridge University Press, Cambridge, 1993.

[18] A.A. Markov, "Insolubility of the problem of homeomorphy," *Prof. International Congress of Math.,* 1960.

[19] J. Milnor, "Two complexes which are homeomorphic but combinatorially distinct," *Annals of Mathematics,* **74**:3, 1961, 575-590.

[20] E. Moise, *Geometric Topology in Dimensions 2 and 3,* Springer-Verlag, 1977.

[21] Y. Nenov, *Computational Logic of Euclidean Spaces,* Ph.D. Thesis, University of Manchester, 2011.

[22] I. Pratt. First-Order Qualitative Spatial Representation Languages with Convexity. *Spatial Cognition and Computation,* **1**, 181-204. 1999.

[23] I. Pratt and D. Schoop. Expressivity in polygonal, plane mereotopology. *Journal of Symbolic Logic.* **65**(2), 822-838. 2000.

[24] I. Pratt-Hartmann. First-Order Mereotopology. In M. Aiello, I. Pratt-Hartmann, and J. van Benthem, *Handbook of Spatial Logics,* Springer. 2007.

[25] D.A. Randell, Z. Cui, and A.G. Cohn. A spatial logic based on regions and connection. *Third International Conference on Principles of Knowledge Representation and Reasoning.* 165-176. 1992.

[26] C.P. Rourke and B.J. Sanderson, *Introduction to piecewise-linear topology,* Springer-Verlag, 1982.

[27] M. Schaefer and D. Stefankovic. Decidability of string graphs. *Journal of Computer and System Sciences.* **68**, 319-334. 2004.

[28] A. Tarski. Foundations of the Geometry of Solids. In A. Tarski, *Logic, Semantics, Metamathematics.* Oxford Clarendon Press. 1956.

[29] C. Thomassen. "The Jordan-Schonflies Theorem and the Classification of Surfaces," *American Mathematical Monthly,* **99**, 1992, 116-131.

[30] A.N. Whitehead. *Process and Reality.* Macmillan, New York. 1929.