Modes of Overlap

Antony Galton
Department of Computer Science
University of Exeter

May 2, 1996

1 Introduction

Representing and reasoning about spatial relationships has important applications in a number of areas, notably Artificial Intelligence and Geographical Information Systems. In the latter area particularly, the spatial entities of interest are two-dimensional regions representing portions of the earth's surface.

There exist a number of theories for handling the qualitative relationships amongst such regions, of which two of the best-known are the Regional Connection Calculus of Cohn *et al.* (Randell, Cui and Cohn 1992) and the 4-intersection and 9-intersection systems of Egenhofer (Egenhofer 1989, Egenhofer 1991, Egenhofer and Franzosa 1991). In Cohn's RCC-8 system, the qualitative relationship between two spatial regions is classified according to the degree of connection exhibited between the regions, e.g., totally disconnected, connected only at boundary points, and so on. In this paper we shall be particularly concerned with just one of these relations, designated PO (for 'partial overlap'): this is the relation in which each of the two regions has points both inside and outside the other.

The relation PO in fact covers a multitude of qualitatively different configurations, including, for example, both the examples in figure 1. Differences of this kind can be quite significant. Suppose, for example, that the regions represent two different features either of which renders the land uninhabitable by a particular species. Then in (b), unlike in (a), the land inhabitable by the species falls into two disjoint regions. The result may be that the two populations cannot mix, a situation liable to lead to a differentiation into two subspecies, and ultimately two species. In (a) this is less likely since the region habitable by the species forms a connected whole.

The RCC-8 system cannot readily handle distinctions of this kind. They can, to be sure, be expressed—e.g., in (b) the complement of the fusion of A and B is itself the fusion of two regions which are disconnected from each other. But the calculus does not lend itself readily to the systematic description of relations of this kind, and similar remarks apply to Egenhofer's 4-intersection and 9-intersection systems.

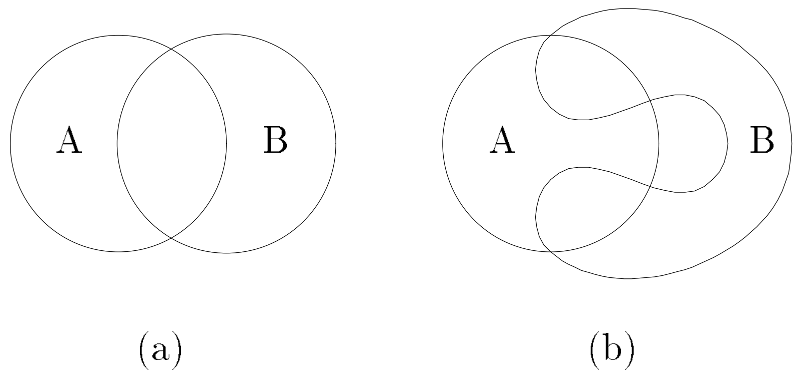

Figure 1: *Topologically distinct versions of 'partial overlap'.*

On the other hand, the recently-developed system of (Egenhofer and Franzosa 1995) can readily characterise the difference between the two figures in Figure 1: the figure in (a) can be denoted

$$\langle 0(0, \text{CrossingInto}, \text{Unbounded}); 1(0, \text{CrossingOutOf}, \text{Unbounded})\rangle,$$

whereas the one in (b) is represented by the longer expression

$$\langle 0(0, \text{CrossingOutOf}, \text{Unbounded}); 3(0, \text{CrossingInto}, \text{Unbounded}); \\ 2(0, \text{CrossingOutOf}, \text{Bounded}); 1(0, \text{CrossingInto}, \text{Bounded})\rangle.$$

This system of Egenhofer's (which we shall not explain here) is rich enough to characterise configurations of spatial regions up to topological equivalence, but it does so at the cost of requiring the complexity of the representations to increase indefinitely in step with the complexity of the configurations themselves. This is, of course, inevitable, and should not be regarded as a fault in Egenhofer's system. For some purposes, however, we might well be content with a less complete form of representation, while still requiring to discriminate between configurations such as those in Figure 1 which are lumped together in the RCC-8 system and Egenhofer's 4- and 9-intersection systems.

In this paper we shall present a new method for characterising the infinitely many qualitatively different ways in which spatial regions can overlap. This method can be regarded as a compromise between the simple, rather undiscriminating systems of Cohn and Egenhofer on the one hand, and the more complex topologically complete system of Egenhofer on the other. The method has some affinity with Egenhofer's matrix-based methods, though in its detailed working is quite different.

2 The overlap matrix

The essential difference between illustrations (a) and (b) in figure 1 is that the complement of $A \cup B$ is connected in the former case but has two connected components in the latter.

Our method for describing modes of overlap is to count the connected components, not just of the complement of $A \cup B$, but also of the region of overlap, $A \cap B$, as well as the 'left-over' regions $A \setminus B$ and $B \setminus A$[1].

The **overlap matrix** for two regions A and B is a 2×2 matrix with natural-number entries

$$[A, B] = \begin{pmatrix} x & a \\ b & o \end{pmatrix},$$

where

x is the number of connected components of $A \cap B$,

a is the number of connected components of $A \setminus B$,

b is the number of connected components of $B \setminus A$,

o is the number of connected components of $(A \cup B)^c$,

The regions A and B are assumed to have co-dimension zero, i.e., they are of the same dimension as the space in which they are considered to be embedded. In our illustrations we shall assume they are two-dimensional regions in a two-dimensional space (e.g., geographical areas on the surface of the globe).

The topological character of a single region A can to some extent be expressed by means of overlap matrices. Let $cv(A)$ be the convex hull of A, and let

c = the number of connected components of A,

c' = the number of connected components of the complement of A,

v = the number of connected components of $cv(A) \setminus A$,

v' = the number of connected components of the complement of $cv(A)$.

Clearly $A \setminus A$ is empty, and since any region is entirely contained in its convex hull, $A \setminus cv(A)$ is empty too. In addition, $A \cap A = A \cap cv(A) = A$. Hence we can put

$$[A, A] = \begin{pmatrix} c & 0 \\ 0 & c' \end{pmatrix}$$

$$[A, cv(A)] = \begin{pmatrix} c & 0 \\ v & v' \end{pmatrix}$$

Assuming A to be finite, then c' will be equal to one more than the number of interior cavities of A (for example, if A is the annular region bounded by two concentric circles, then the connected components its complement are, firstly, the region outside the outer

[1]The use of set-theoretic notations here is convenient, but nothing in what follows presupposes a commitment to regarding regions as sets of points in the style of point-set topology; and in particular, we shall not address the question of 'open' vs 'closed' regions.

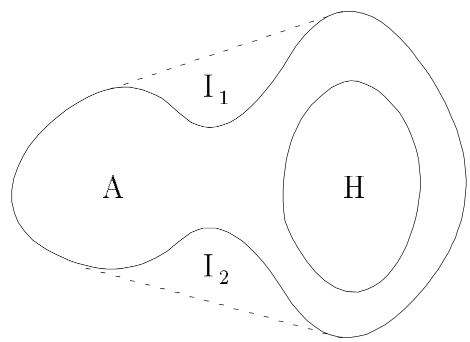

Figure 2: *Two kinds of concavity.*

circle, and secondly, the interior of the inner circle—the latter component being the sole interior cavity).

In two dimensions, if A is connected, then the connected components of $cv(A)\backslash A$ consist of, firstly, any interior cavities possessed by A, and secondly, any surface concavities, i.e., indentations in its boundary. The total number of such components is v, whereas the number of interior cavities is $c' - 1$. Hence the number of surface concavities must be $v - c' + 1$.

In figure 2, the connected region A has two surface concavities I_1 and I_2, and an internal cavity H. Thus

$$[A, A] = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \qquad [A, cv(A)] = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix},$$

3 Overlap space

The possible modes of overlap define a complexly-structured 'overlap space'. We shall begin by exploring in detail the restricted case in which A and B are finite connected regions of topological genus zero (i.e., one piece, no holes)—we shall call such regions *simple regions*. Later we shall discuss the implications of relaxing these constraints.

Not all matrices correspond to possible modes of overlap. There are a number of constraints, some of which apply generally, others only for our restricted class of regions. In the following list, we shall use the canonical notation $\begin{pmatrix} x & a \\ b & o \end{pmatrix}$, as explained above.

1. A region must have at least one connected component. Since $A = (A \cap B) \cup (A \setminus B)$, at least one of $A \cap B$ and $A \setminus B$ must have at least one component, and likewise with $A \cap B$ and $B \setminus A$. Hence we have

$$\begin{aligned} x + a &> 0 \\ x + b &> 0 \end{aligned}$$

Figure 3: *'Bridges' required when A is connected.*

or, put differently,

$$x = 0 \Rightarrow (a > 0 \wedge b > 0).$$

2. If A and B are both finite regions in infinite space, then $(A \cup B)^c$ must be non-empty, so

$$o > 0.$$

3. If A is connected, but $A \cap B$ is not, then A is not contained in B, so $A \setminus B$ must be non-empty. This is illustrated in figure 3(a), where two components of $A \cap B$ are shown, together with the 'bridge', part of $A \setminus B$, which is needed to ensure that A is connected. If B is connected too, then we need a separate bridge to form part of $B \setminus A$. Hence, if both A and B are connected, we have the constraint

$$x > 1 \Rightarrow a > 0 \wedge b > 0.$$

4. Assume now that A and B are both finite, connected, and of genus zero. If $(A \cup B)^c$ is disconnected, then since $A \cup B$ is finite, one of the connected components of $(A \cup B)^c$ must be a 'hole' in $A \cup B$. But if A is contained in B then $A \cup B = B$, so B has a hole; and likewise if B is contained in A then A must have a hole. Since neither A nor B has a hole (being of genus zero), it follows that neither is contained in the other, and hence both $A \setminus B$ and $B \setminus A$ are non-empty. We thus have the constraint

$$o > 1 \Rightarrow a > 0 \wedge b > 0.$$

5. Assume A is connected, and suppose that $A \cap B$ is empty. If $A \setminus B$ has more than one connected component, then as shown in figure 3(b), there must be a 'bridge', forming part of $A \cap B$, connecting the two components of $A \setminus B$ to ensure that A is connected. Since $A \cap B$ is empty, no such bridge can exist, so $A \setminus B$ cannot have more than one component. On the other hand, since $A = (A \cap B) \cup (A \setminus B)$, and $A \cap B$ is empty whereas A is not, it follows that $A \setminus B$ must have at least one component. Hence it has exactly one component. Similarly, if B is connected and $A \cap B$ is empty, then $B \setminus A$ must have exactly one component. We have the constraint

$$x = 0 \Rightarrow a = b = 1.$$

> Note that this implies our first constraint, but whereas the first constraint depends only on the assumption that A and B are non-null, this one assumes also that each region is connected.

Since there are infinitely many overlap matrices satisfying these constraints, we shall confine our attention to a small finite selection of them. We shall investigate all the overlap matrices satisfying these constraints with no entries greater than 2. Since o must be either 1 or 2, and each of the other three entries in the matrix is 0, 1, or 2, there are $2 \times 3^3 = 54$ *prima facie* possibilities. Not all of these satisfy the constraints.

If either a or b is 0, then we must have $x = o = 1$. this gives the five matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

Again, if $x = 0$, we must have $a = b = 1$, giving just the two matrices

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}.$$

In every other case all four entries must be either 1 or 2, giving $2^4 = 16$ possibilities. We have thus reduced the 54 *prima facie* cases to $5 + 2 + 16 = 23$. It turns out that all 23 correspond to possible modes of overlap for regions of the kind we are considering. These are shown in figure 4.

4 Change over time

We assume that all change is continuous. Even so, there must be discontinuities in the *qualitative* description. For example, consider the sequence illustrated in figure 5. Initially the two regions have overlap matrix

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

A 'finger' of region B comes round to meet A again. At the instant that it does so, the region outside both A and B becomes disconnected, with overlap matrix

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

If the finger of B continues its growth so that it penetrates into A, then the overlap matrix becomes

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

In this example, it will be noted that the change in the overlap matrix is 'quasi-continuous'

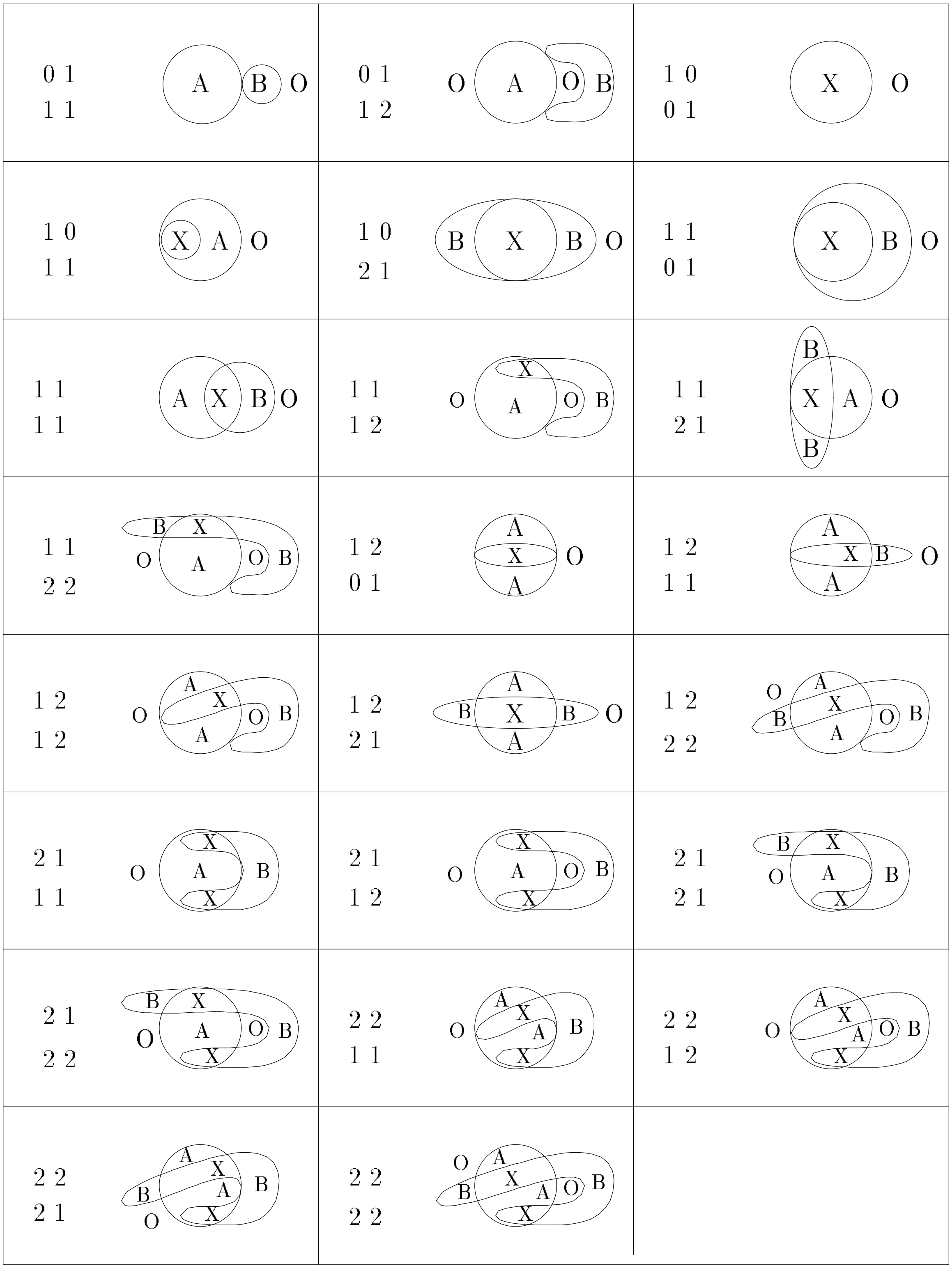

Figure 4: *The 23 simplest modes of overlap for simple regions*

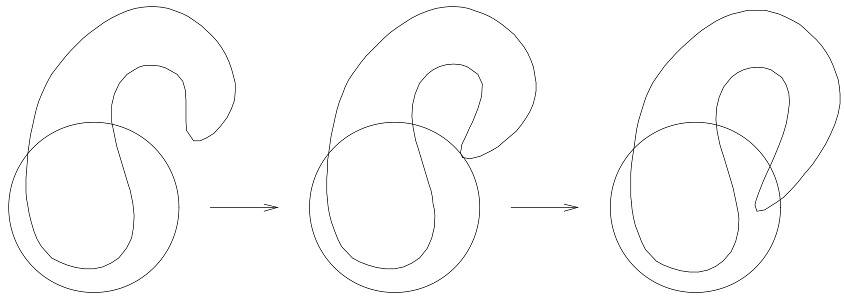

Figure 5: *A continuous transformation with a singularity.*

in that each change is as small as it can possibly be, given that the entries are integers: one entry in the matrix changes by 1 at each step.

A change in the overlap matrix corresponds to what Achille Varzi (pers. comm.) has called a "magic moment" of topology. One such moment occurs when two elements come into or break contact, thus altering the overall topological character of the configuration they form part of. The change is strictly instantaneous, in the sense that the 'contact' state holds at an instant which forms the lower or upper limit of a time interval throughout which the 'non-contact' state holds. Another kind of magic moment occurs when a component of a region, after a period of shrinking, finally disappears entirely. Again, the vanishing takes place at an instant. Changes of these kinds are *singularities* in the continuous deformation of the regions.

Some changes in mode of overlap can involve more than one singularity simultaneously (a 'multiply magic' moment). An example is shown in figure 6. Here two regions are initially in a configuration with overlap matrix

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Region B shrinks in such a way that the two connected components of $B \setminus A$ vanish simultaneously, yielding overlap matrix

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

An entry in the matrix has changed by 2, corresponding to the two simultaneous singularities.

Sometimes two simultaneous singularities can cancel each other out in the sense that the overall topological character, as specified by the overlap matrix, remains unchanged. An example is shown in figure 7. Here an indentation develops in the boundary of one of the components of $B \setminus A$; at the same time the other component of $B \setminus A$ starts to shrink. The tip of the indentation makes contact with A at exactly the same moment as

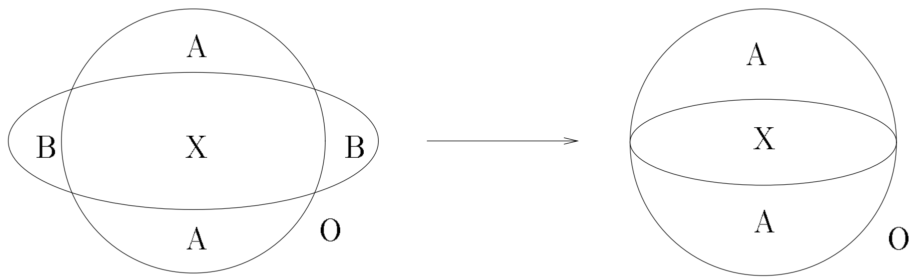

Figure 6: *Transformation with two simultaneous singularities.*

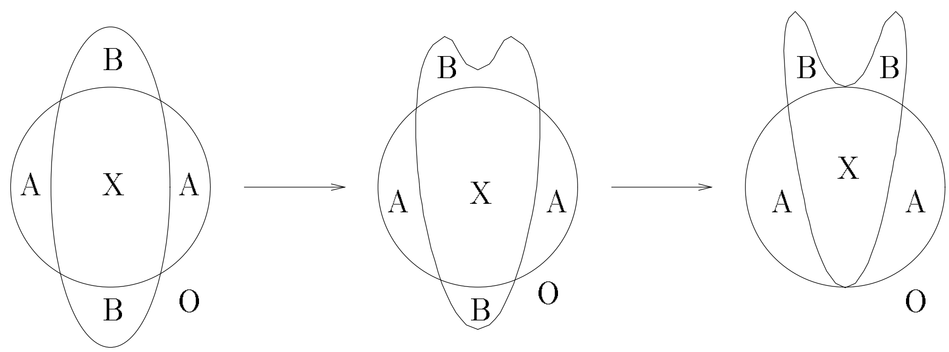

Figure 7: *Simultaneous singularities preserving mode of overlap.*

the shrinking component finally vanishes. But at that moment, the remaining component becomes two. Thus throughout the change, the overlap matrix is

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

5 Perturbation

Following (Galton 1993), we call a qualitative state S' a *perturbation* of qualitative state S so long as at least one of the following situations is possible:

1. S' holds at an instant that limits an interval over which S holds.
2. S holds at an instant that limits an interval over which S' holds.

Perturbation is clearly a reflexive and symmetric, but not in general transitive relation. To exclude the reflexive case, we shall call S' a *proper perturbation* of S so long as S' is a perturbation of S and $S' \neq S$.

States which are perturbations of each other are neighbours in state-space. So long as continuity is preserved, all changes of state must be from a state to one of its perturbations. Knowledge of the perturbation relationship can thus provide an important tool in predicting future changes or reconstructing past ones. A diagram showing perturbations is an *envisionment* in the sense of de Kleer (see for example (de Kleer and Brown 1984)). Singularities are akin to the attainment of *landmark values* (Kuipers 1986).

Here we shall be concerned with the perturbation structure of overlap space, as specified by the set of possible overlap matrices. The smallest possible change to an overlap matrix is when just one entry increases or decreases by 1. The matrices correspond to nearest neighbours in overlap space. The full diagram of the nearest neighbour relation will be a subset of the four-dimensional hypercube lattice (or *tesseract* lattice), occupying that one-sixteenth part of the four-dimensional space for which the coordinates are non-negative. Unlike the RCC and Egenhofer models, we are here dealing with infinitely many qualitatively distinct states.

Amongst the 23 modes of overlap presented in Figure 4, which pairs are perturbations of each other? It can be verified by inspection that whenever the overlap matrices differ by at most 1, the corresponding modes of overlap are mutual perturbations. If the matrices are set out on the nodes of the tesseract lattice, we obtain a diagram such as in figure 8.

In some cases, two neighbouring transitions can occur simultaneously, corresponding to perturbations involving multiple singularities. An example is the case shown in Figure 6, in which there is a transition from $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ to $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ which does not pass through the natural intermediate state $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$.

Another example, this time involving a shift along a diagonal of the tesseract lattice, is the transition from $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ to $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ shown in Figure 9.

Not all diagonals in the tesseract lattice correspond to possible direct transitions, however. For example, there is no way of getting from $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ to $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ without passing through either $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. A full understanding of the structure of overlap space therefore requires us to be able to predict which 'diagonal' transitions are possible and which are not. The key to this is the notion of tangency, which will be introduced in the next section.

6 Tangency

We shall gain a clearer insight into the structure of overlap space by examining more closely exactly what happens when one mode of overlap is directly transformed to another. Whenever this happens, a connected component is created or destroyed. The fundamental sequence governing this process is illustrated in Figure 10. In (a), two boundary segments

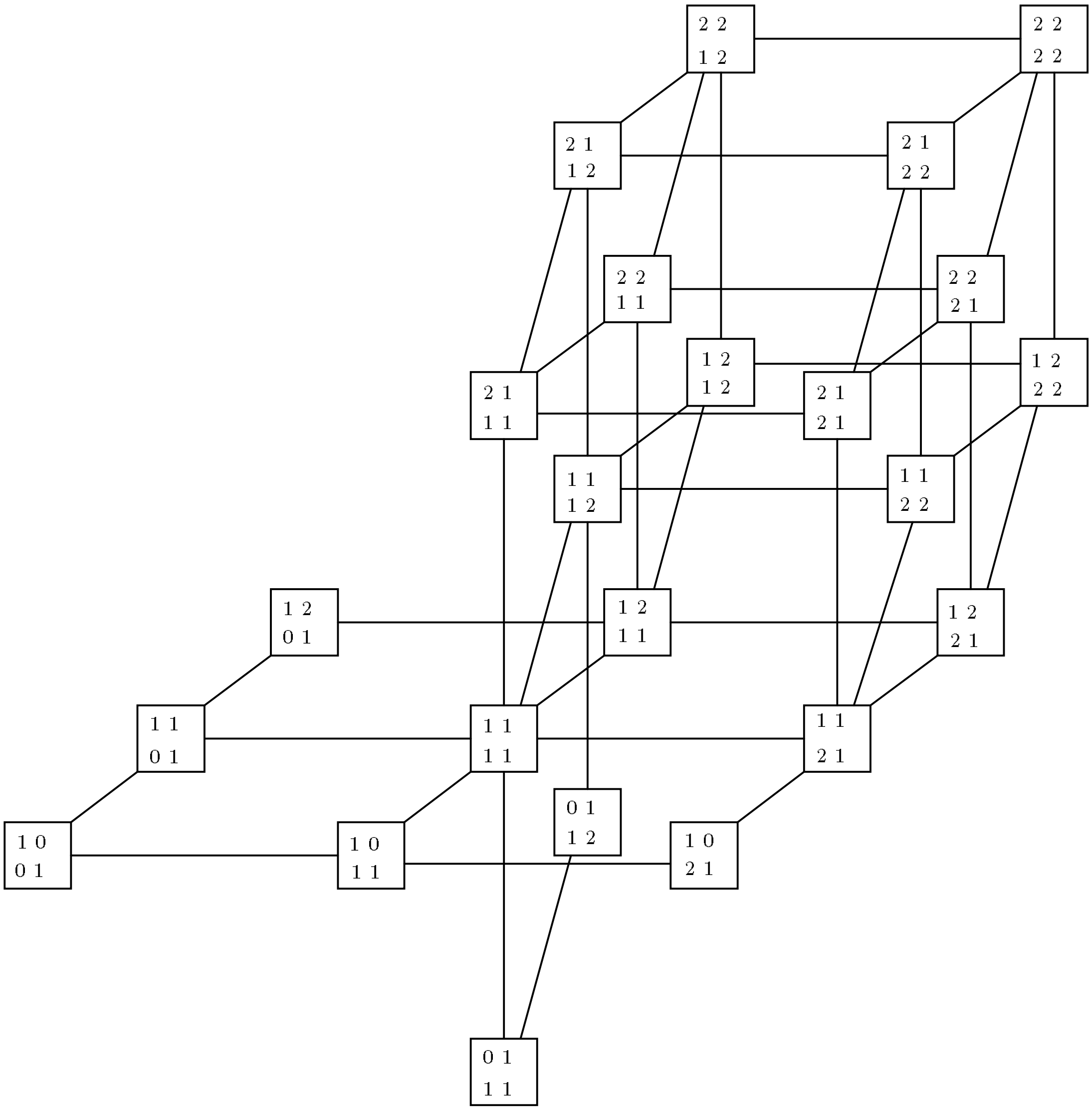

Figure 8: *23 overlap matrices for simple regions.*

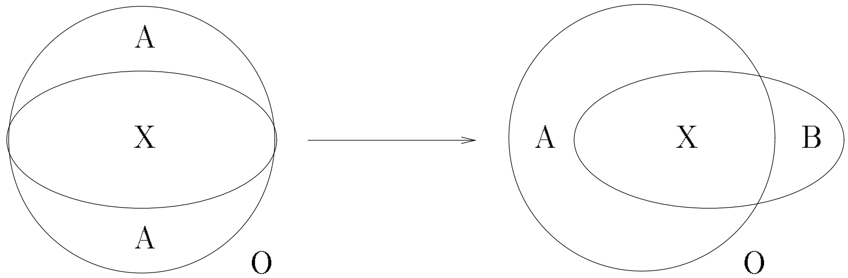

Figure 9: *Direct transition along a diagonal of the tesseract lattice.*

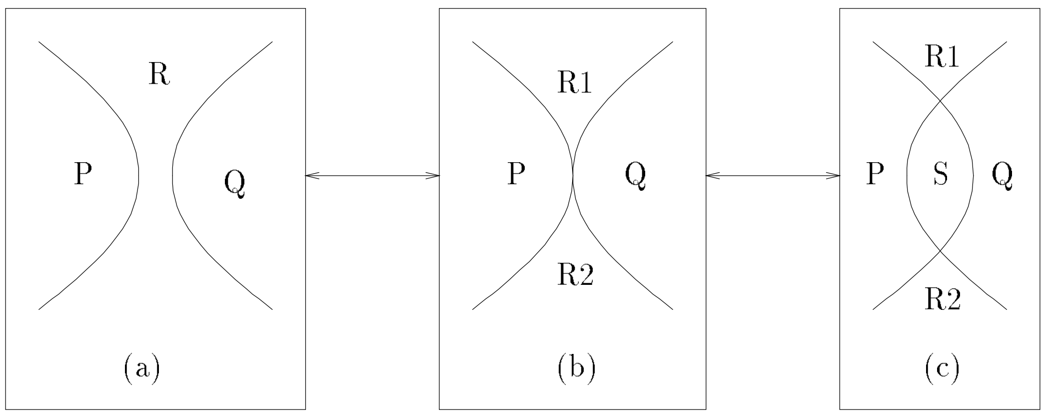

Figure 10: *Component creation by contact and penetration.*

of regions P and Q are separated by part of region R. At (b), the two segments become mutually tangential (*contact*), thus (in general) dividing R into two components $R1$ and $R2$. At (c) the two segments have crossed one another (*penetration*), leading to the creation of a new component S. This process can be clearly seen in the sequence illustrated in Figure 5.

The situation is complicated by the fact that the boundary segments may be variously oriented with respect to the primary regions A and B. If the left-hand segment forms part of the boundary of A, for example, it makes an important difference whether it is convex or concave with respect to A, in other words whether region P is part of A or not. Taking all the possibilities into consideration, we have the eight cases illustrated in Figure 11

Taking the figure one row at a time, we have:

(i) The transition Oc results in the division of the region O into two components $O1$ and $O2$ as a result of contact between regions A and B. The transition Xp results in the creation of a new overlap component X as a result of the interpenetration of A

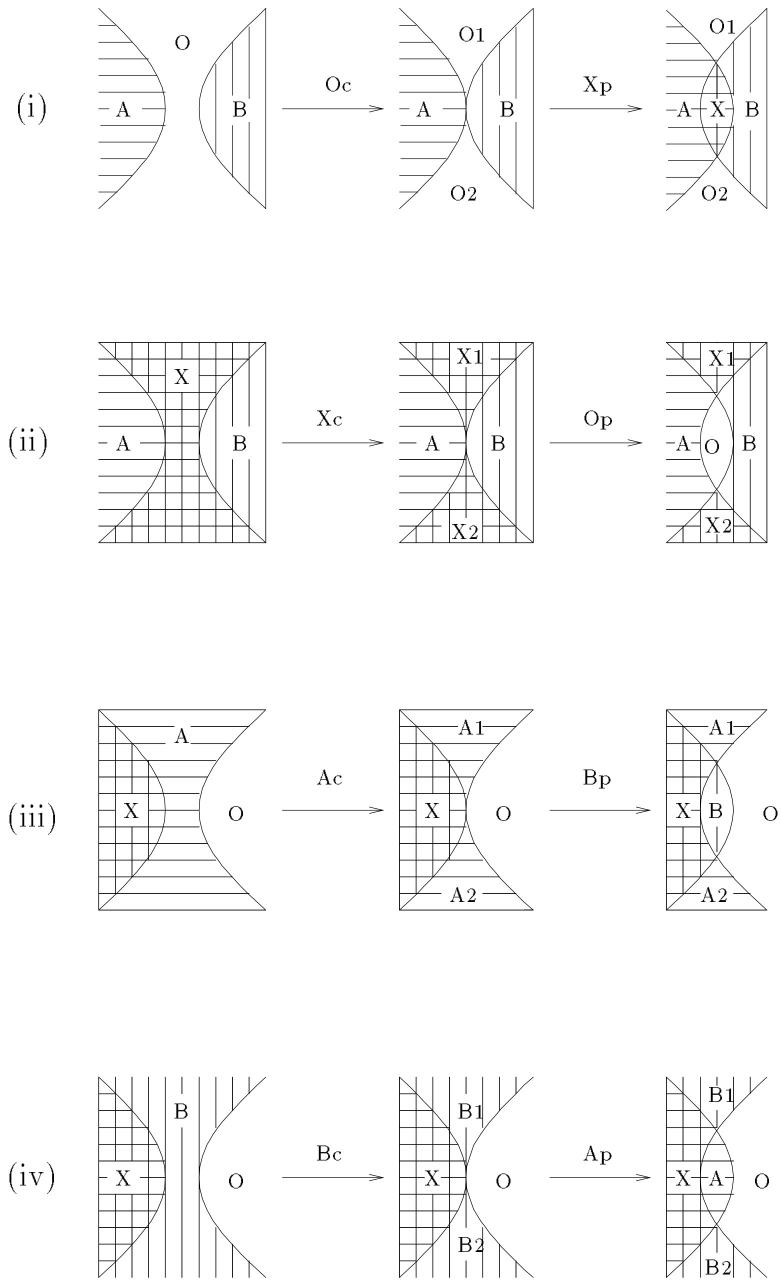

Figure 11: *The eight types of component creation.*

and B. The whole sequence may be illustrated by the overlap modes

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

(ii) Xc creates a new X-component by contact while Op creates a new O-component by penetration. The sequence is illustrated by the overlap modes

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

(iii) Ac creates a new A-component by contact while Bp creates a new B-component by penetration. The sequence is illustrated by the overlap modes

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

(iv) Bc creates a new B-component by contact while Ap creates a new A-component by penetration. The sequence is illustrated by the overlap modes

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

We can take a pair of adjacent elements in the lattice diagram (Figure 8), say $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, and consider the nature of the transition between them. The transition must involve the creation of a new component to region B. This can only occur by means of one of the transitions Bc and Bp, the former involving the creation of a tangency, the latter the destruction of one. In fact both types of transition are possible, as shown in Figure 12. The two versions of the resulting overlap mode can be converted into each other, without departing from that mode, as shown in Figure 7. This conversion involves the simultaneous operation of transitions Bc and Bp^{-1}, resulting in no net change in the number of B-components, but a gain of two points of tangency.

It should be noted that there are a few occasions when the transitions shown in Figure 11 do not result in the creation of a new component, as follows:

1. A and B are separated from one another and then come into contact. This is a transition of type Oc, but the number of components of region O remains equal to 1. The overlap matrix—which is $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ throughout—cannot detect this change. In the RCC notation, the change is from DC to EC. As hinted at the start, overlap matrices make fine discriminations within the RCC relation PO, but are relatively undiscriminating outside it.

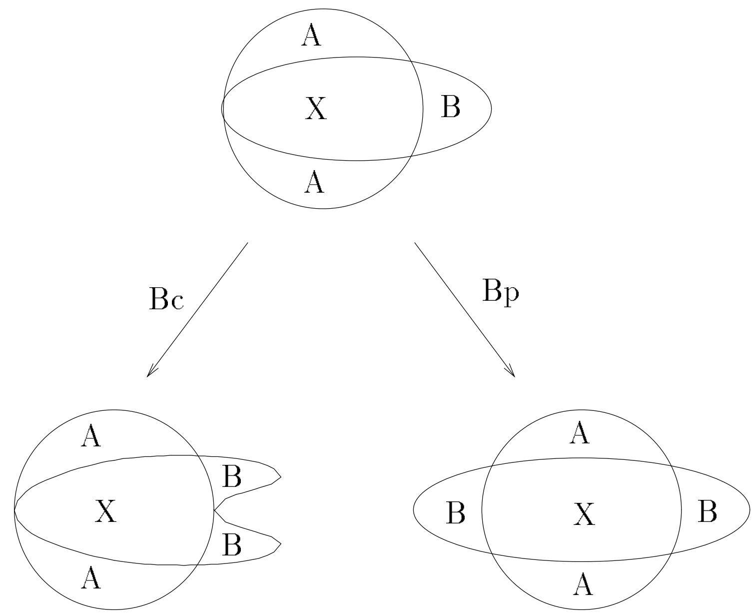

Figure 12: *Two ways of obtaining overlap mode* $(1,2;2,1)$ *from* $(1,2;1,1)$.

2. A is a non-tangential part of B (NTPP in the RCC system), and moves out to make contact with boundary of B in one place, thus becoming a tangential proper part (TPP) of it. This is a transition of type Bc, but the number of components of B remains unchanged, the overlap matrix being $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

3. The same, with A and B swapped round.

Finally, there is one other type of case which is not well described in terms of Figure 11. This is the case when regions A and B are initially equal—overlap matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$—and then either A becomes a tangential (or non-tangential) proper part of B or *vice versa*. Here we have a creation of a component of either $B \setminus A$ or $A \setminus B$, but none of the transitions Ap, Ac, Bp, Bc seems to be an adequate description of what happens here. The boundaries of A and B, initially coincident, become separated from one another along part of their length, causing a new region to appear between them. In general, this kind of phenomenon can be assimilated to component creation by means of a contact transition Xc, Ac, Bc, or Oc, as shown in Figure 13; it is just in the cases considered here that this way of treating it becomes unviable.

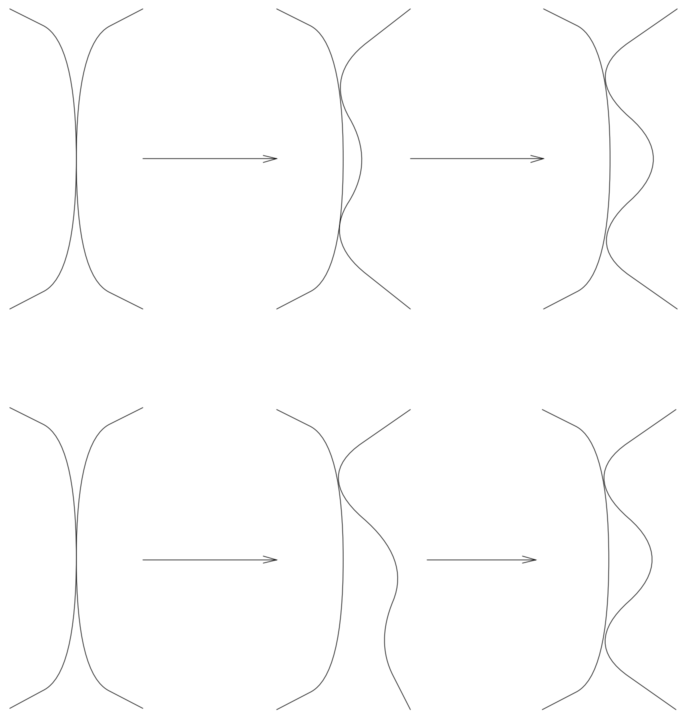

Figure 13: *Boundary separation (top) is usually equivalent to tangency gain by contact (bottom).*

7 Multiple singularities

Each transition in Figure 11 is a singularity ('magic moment'). There are 16 different singularities, since each of the eight named transitions has an inverse. The 'c' transitions involve gain of a tangency and gain of a component, their inverses involve loss of a tangency and loss of a component. The 'p' transitions involve loss of a tangency and gain of a component, their inverses involve gain of a tangency and loss of a component.

Any two or more singularities can occur simultaneously, but the effect of this on the overlap matrix can be various. For simplicity's sake we shall confine our attention to the case of two simultaneous singularities. We shall write K for the overlap matrix before the double singularity, L for the matrix at the instant of singularity, and M for the matrix after the singularity. Each singularity involves either loss or gain of a tangency. There are thus three different types of simultaneous double singularity:

(I) Both singularities involve gain of tangency.

(II) Both singularities involve loss of tangency.

(III) One singularity involves gain, and the other loss, of tangency.

Under continuous deformation, any neighbourhood of a state of tangency must include states of non-tangency, so in order to get from a state of non-tangency to a state of tangency it is necessary to pass through intermediate positions of non-tangency. Thus we have the principle that *whenever a tangency is gained or lost, the tangency itself is present at the instant of transition.* We shall call this the **tangency principle**. It is related to the 'continuity rule' of (Williams 1990), with tangency taking the place of value zero.

Using this principle, we can see that in a double singularity of type I we must have $L = M$, while for a double singularity of type II we must have $K = L$. The third matrix may or may not be equal to the other two. Only in the case of type III can we have L distinct from both K and M, and in this case we may or may not have $K = M$. We thus have the possibilities shown in the following table:

	$t < t_0$	$t = t_0$	$t > t_0$	tangency change
Ia	K	M	M	$+2$
Ib	K	K	K	$+2$
IIa	K	K	M	-2
IIb	K	K	K	-2
IIIa	K	L	M	0
IIIb	K	L	K	0

Here t_0 is the time at which the double singularity occurs. The possibilities shown in the table are explained as follows.

Ia This case involves the simultaneous gain of two tangencies, with a change in mode of overlap. An example is the transition

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix},$$

in which regions $A \setminus B$ and $(A \cup B)^c$ simultaneously each lose one component with the onset of two tangencies. The singularities here are Op^{-1} and Ap^{-1}.

Ib Here the overlap matrix remains constant throughout, with the simultaneous gain of two tangencies. This is the case already illustrated in Figure 7, the two singularities in this case being Bc and Bp^{-1}.

IIa This is the time-reversal of type Ia. An example is shown in Figure 9, with singularities Ac^{-1} and Bp.

IIb This is the time-reversal of type Ib.

IIIa This requires simultaneous tangency loss and gain, with overall change in overlap mode. An example is the sequence

$$\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

which is illustrated as the upper sequence in Figure 14.

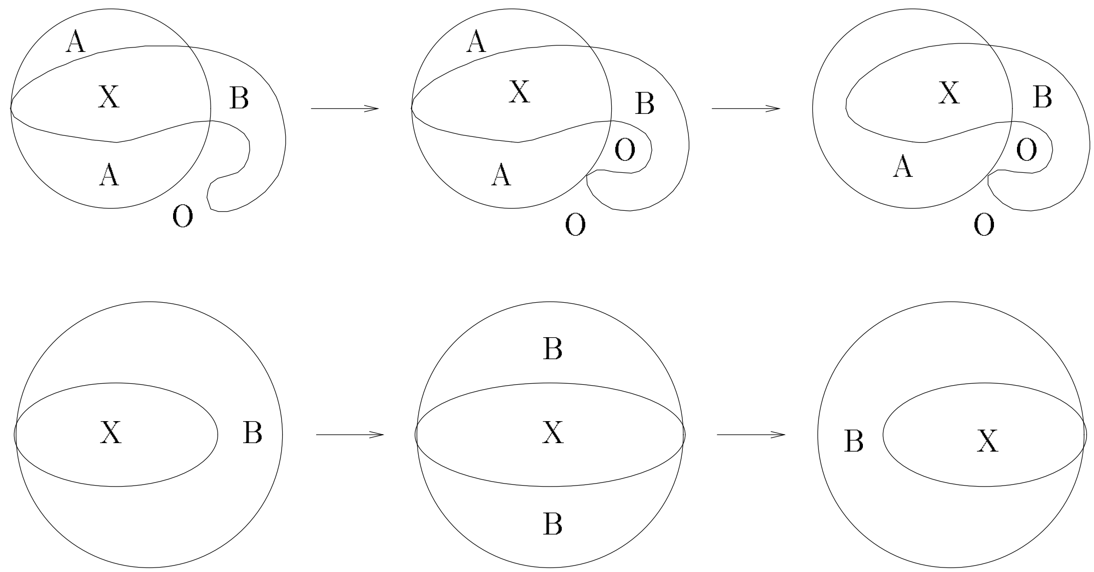

Figure 14: *Double singularities with intermediate overlap mode.*

IIIb Here the overlap matrix ends up the same as it started, but at the instant of singularity it is different: this also requires simultaneous gain and loss of tangency. An example is shown as the lower sequence in Figure 14, in which the left-hand extremity of A detaches itself from the boundary of B at the same instant as the right hand extremity of A makes contact with the boundary of B. Just at that instant the overlap mode is different; the transition is

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

We now have an answer to our earlier question concerning the diagonals of the tesseract lattice. Take any path along the orthogonal connections in the tesseract lattice: this represents a possible transformation in overlap space, each step of the transformation corresponding to one link in the path. It is always possible to carry out these steps in sequence; the question is whether they can be carried out simultaneously so as to lead directly from the first mode in the sequence to the last in one step. The answer is that this will be possible if and only if either all the steps involve an increase in the number of tangencies or all the steps involve a decrease in the number of tangencies. Thus for example, in Figure 6 there is a direct transition involving simultaneous gain of two tangencies, while Figure 9 shows a direct transition involving the simultaneous loss of two tangencies.

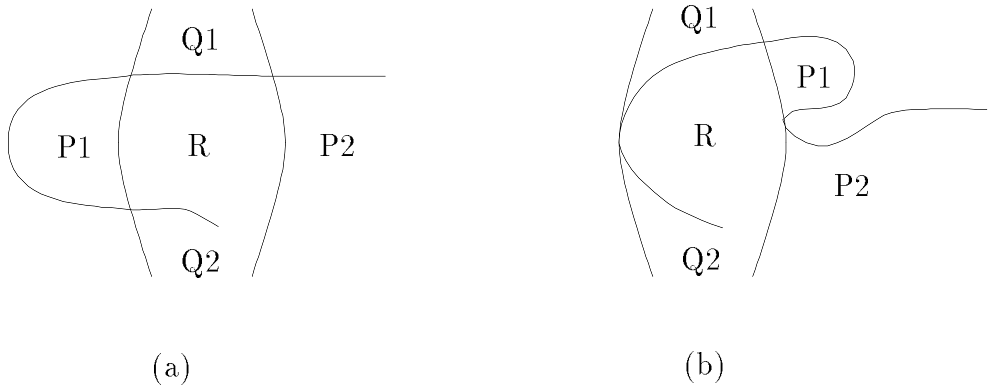

Figure 15: *Isotope formation.*

8 Isotopy

In Figure 7 we saw that there are two topologically distinct configurations having overlap matrix $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. They are distinguished by the fact that in one of them there are no tangencies whereas in the other there are two tangencies. Since these two configurations occupy the same place in the tesseract lattice, it is appropriate to call them *isotopes* of the overlap mode represented by the matrix.

Not all overlap modes possess multiple isotopes. As far as I have discovered at present, there are only two ways of producing isotopes. The simplest way involves the simple addition of a tangency. In certain cases this can be done without changing the overlap mode. We have already seen two cases of this: they are, first, the DC and EC isotopes of $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, and second, the NTPP and TPP isotopes of $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Other examples of this kind require non-simple regions and will be discussed in section 10. This kind of isotopy will be called *Type I isotopy*.

The second way of generating isotopes, *Type II isotopy*, is by converting a configuration of the type shown in Figure 15(a) into a configuration of the type shown in Figure 15(b). This transformation preserves the overlap mode but introduces two extra tangencies. For this kind of transformation to apply, at least two of the entries in the initial overlap matrix must be greater than 1.

In Figure 16 are shown all the cases of Type II isotopy that can be produced by this means beginning with one of the 23 overlap matrices with all entries at most equal to 2. In each case the 'shifted' region, represented by P1 in Figure 15, is indicated by the letter of the region (A, B, O, or X) which it is a component of. In four cases there are two candidates for the 'shifted' region—in three of the four cases, choice of either candidate

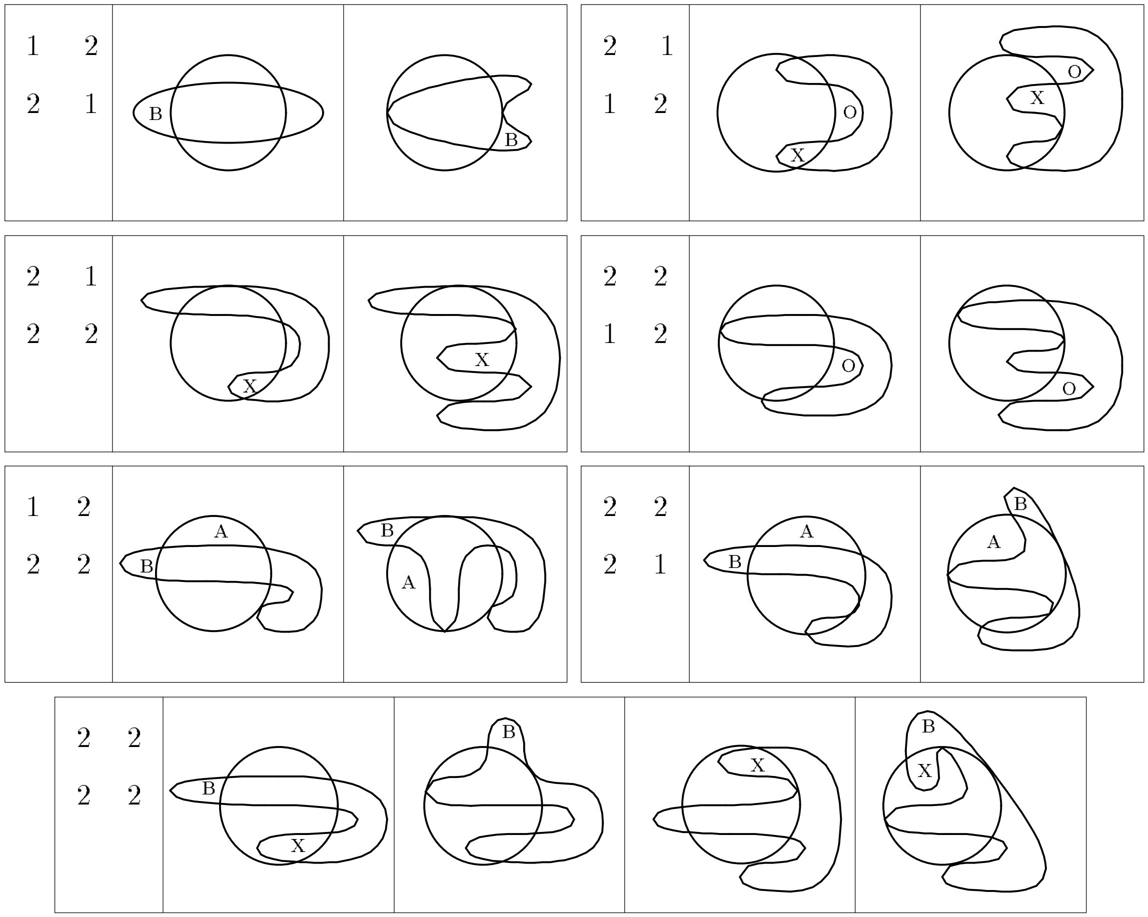

Figure 16: *Isotopes for matrices with entries* ≤ 2.

leads to topologically identical results, but in the case of $\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$ it makes a difference whether the B-component or the X-component is selected; if both are selected, an extra four tangencies arise, giving the fourth isotope shown for this matrix.

If a matrix M has an isotope with n tangencies, we may denote it ${}^{n}M$. Note that in some cases this will be ambiguous since the matrix may have more than one topologically distinct isotope with n tangencies—it is simply less ambiguous than the unadorned M.

We are now in a position to illustrate a transition involving more than two simultaneous singularities. Consider the sequence

$${}^{0}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \longrightarrow {}^{1}\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \longrightarrow {}^{2}\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \longrightarrow {}^{3}\begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix} \longrightarrow {}^{4}\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

Since each step involves a gain of one tangency, it should be possible to perform the steps simultaneously, thereby transforming the first matrix in the sequence directly to the last.

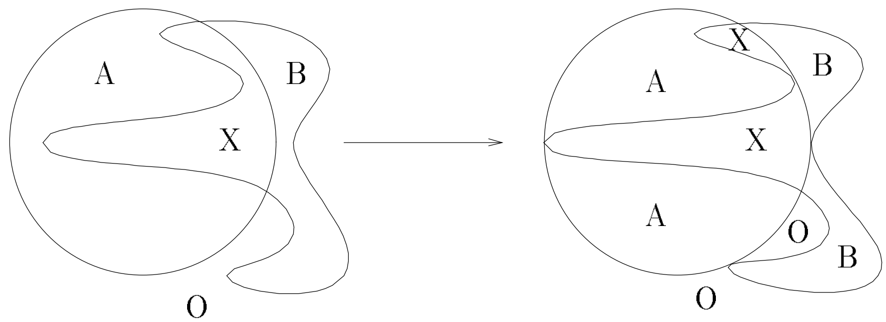

Figure 17: *Simultaneous acquisition of four tangencies.*

This transition is shown in Figure 17. The left-hand illustration, for which the overlap mode is still $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, shows the configuration when it is just about to undergo the transition by simultaneously acquiring 4 tangencies.

9 Instantaneous tenure

If the mode of overlap of two regions A and B undergoes a sequence of changes as a result of the movement, growth, or deformation of one or both regions, then for each mode in the sequence we can consider the time for which that mode is realised. We shall call this time the *tenure* of the mode in the sequence. In general this tenure will be an *interval*, i.e., a period of time having positive duration, but in some cases it may be only a durationless instant. We have seen some examples of such *instantaneous tenure* already: it occurs for the middle mode in both the sequences shown in Figure 14. (Note that this middle mode *may* be realised instantaneously, but it is not forced to—it depends on whether the loss of tangency occurs at the same instant as the gain, or after an interval.)

The results of the previous sections enable us to make some general remarks about the possibility of instantaneous tenure for modes of overlap. If there is a direct transition between distinct isotopes ${}^mM$ and ${}^nN$, then we know that $m \neq n$, since net loss or gain of tangency is involved. We shall say that ${}^mM$ *dominates* ${}^nN$ so long as $m > n$, i.e., the transition from ${}^mM$ to ${}^nN$ involves loss of tangency. By the Tangency Principle, we know that when ${}^mM$ is tranformed to ${}^nN$ or *vice versa*, at the instant of transition it is ${}^mM$ that holds. Suppose now that we have a sequence

$${}^lL \longrightarrow {}^mM \longrightarrow {}^nN$$

and that ${}^mM$ has instantaneous tenure, so that ${}^lL$ and ${}^nN$ hold over contiguous intervals with ${}^mM$ holding at the instant of transition. By the Tangency Principle, ${}^mM$ must dominate both ${}^lL$ and ${}^nN$, i.e., $m > l$ and $m > n$. Thus we have the rule that *instantaneous*

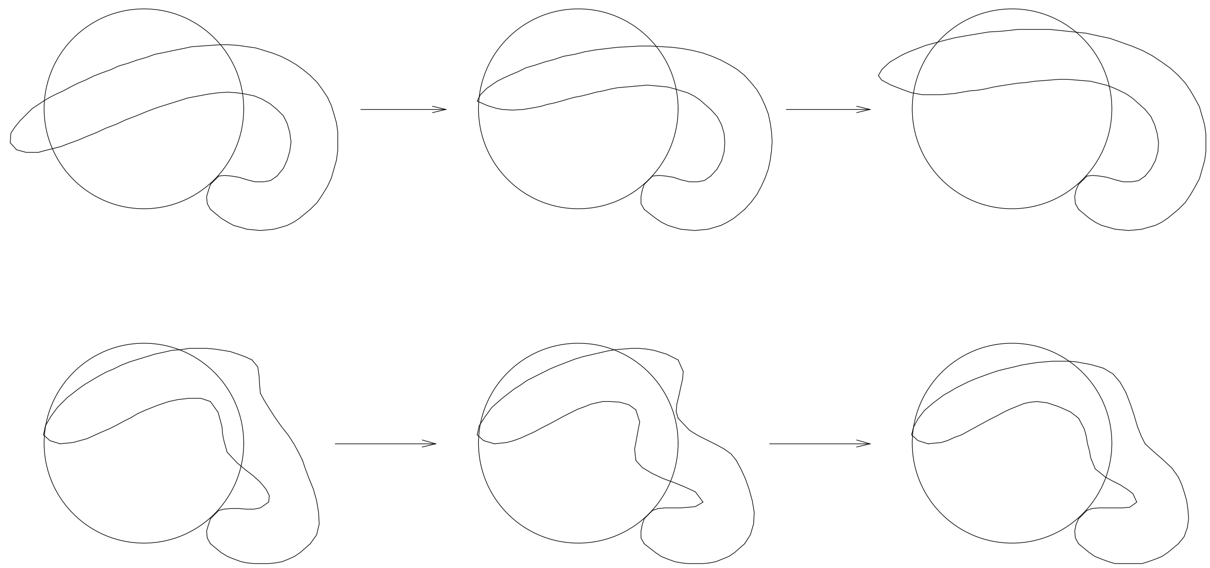

Figure 18: *'Mutual dominance' of two overlap modes.*

tenure is only possible for a given isotope in the context of a transition sequence in which it dominates both its immediate neighbours. (cf. (Galton 1994a)).

We have to express this in terms of isotopes rather than overlap matrices since the tangency number plays a part in determining dominance (in the terminology of (Galton 1994b), the overlap matrices do not constitute a dominance space, but the isotopes do). This results in a curious possibility, namely that we can have two overlap modes such that each may hold instantaneous tenure in the context of a deviation from the other. One might describe this as a case of 'mutual dominance', though this would be something of a misnomer since dominance is really a relation on isotopes rather than on overlap modes. This is illustrated for the matrices $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ and $\begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$ in Figure 18.

10 Complex regions

By a complex region we mean any region which is not simple in the sense introduced in section 3. Thus a complex region might be scattered or contain one or more holes. Some modes of overlap are only possible if one or both of the regions is complex; and for those modes of overlap which are possible with simple regions, the introduction of complex regions can result in further isotopy.

We shall not attempt to survey the overlap modes for complex regions in anything like the same degree of detail as we have applied to the modes for simple regions. We shall merely present a few examples in order to give some idea of the possibilities.

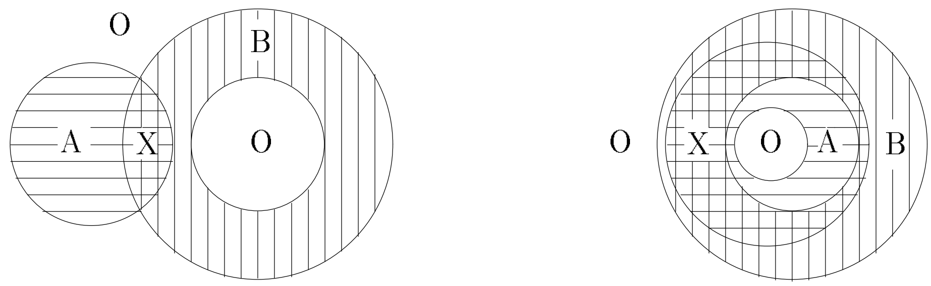

Figure 19: *Additional isotopes involving complex regions.*

As an example of additional isotopy arising from the introduction of complex regions, we take the overlap matrix $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. Figure 19 shows two additional isotopes for this overlap mode, one involving a simple region with a one-hole region, the other involving two one-hole regions.

In Figure 20 we show some examples of overlap modes for complex regions which are not possible at all if both regions are simple.

In the top row are three configurations in which A and B are in fact the same region: in the first case, a one-hole region, in the second a region with two simple components, and in the third a region with a simple component and a one-hole component.

In the second row are three configurations in which A and B are disjoint, with at least one of them having more than one connected component. In the first two cases there is the possibility of adding tangencies without changing the overlap mode: for example, in the second case, we could have one of the components of A touching both components of B, and we could have the other component of A touching one of the components of B. These give Type I isotopes of the original configuration. With the addition of a further tangency, between the second component of A and the other component of B, we change the overlap mode to the one shown in the third example on this row.

The remaining rows show a selection of more complicated cases, two with A being a simple region, four with A a torus, B being variously simple or in two parts. Again, with most of these, there are Type I isotopes obtainable by simply adding a tangency.

Allowing complex regions introduces further complications to the discussion of continuous change. With simple regions, transformation of overlap mode can occur as a result of relative movement of the two regions (as for example in Figure 9), of growth or shrinkage of one or both regions (as in Figure 6), of deformation of a region (as in Figure 5), or any combination of these. Essentially the same holds for complex regions, except that now we must allow that deformation can change the topological character of a region. An example of this is shown in Figure 21. Here a one-hole region is continuously transformed first into a simple region, then into a two-piece region (cf. (Casati and Varzi 1994, Chapter 6)). The

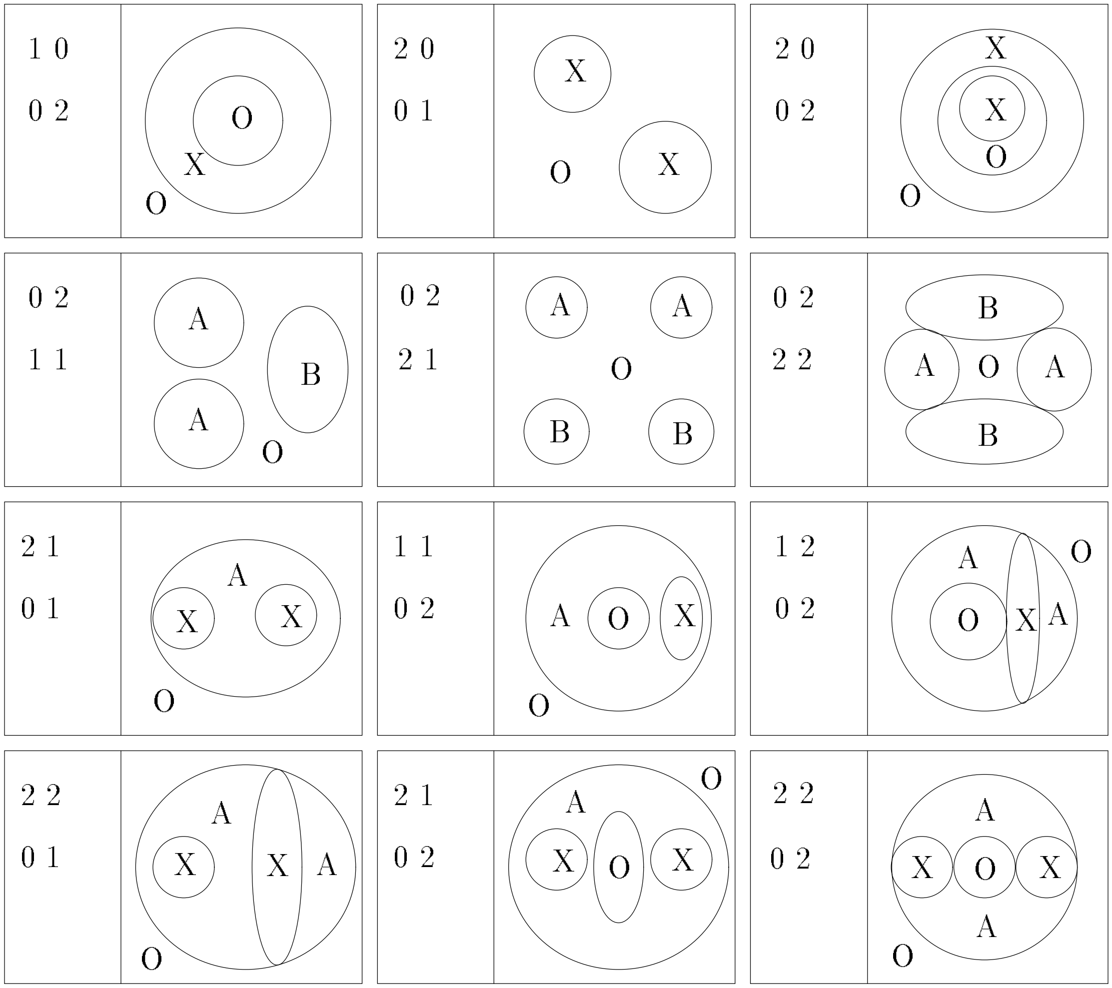

Figure 20: *Overlap modes which require at least one complex region.*

overlap mode of the region A with itself, $[A, A]$, passes through the sequence

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

It is easy to see that A's mode of overlap with a fixed second region B will similarly undergo changes. Although this kind of deformation is in an obvious sense continuous, it should be noted that it is *not* continuous in the topological sense, i.e., the deformed region is not homeomorphic to the original. More precisely, it is continuous with respect to a metric based on distances within the plane as a whole, but discontinuous with respect to a metric based on distances within the region itself.

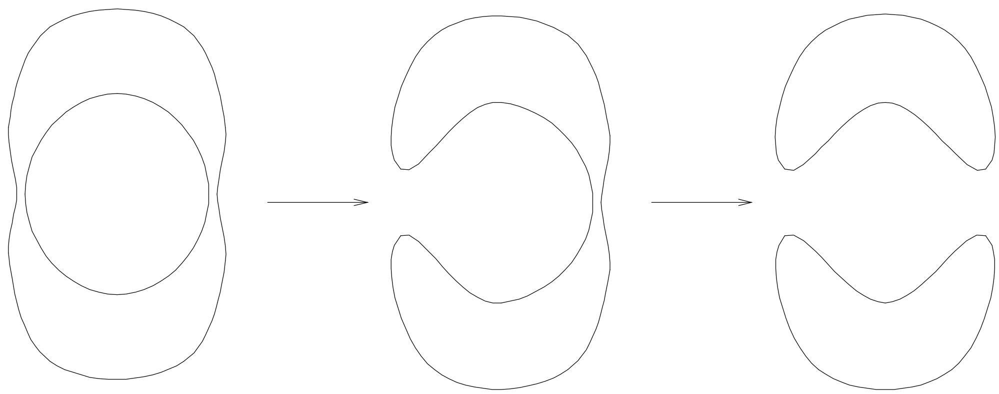

Figure 21: *Continuous deformation of a region with change of topological character.*

11 Concluding remarks

We have devised a system for representing the possible modes of overlap of spatial regions, using an *overlap matrix* inspired by the intersection matrices of Egenhofer. Overlap matrices enable us to describe qualitative spatial configurations at a level of detail hard to emulate in the spatial relation systems of Egenhofer and Cohn *et al.*, without going so far as to embrace the complexities of the topologically complete system of (Egenhofer and Franzosa 1995).

We have attended closely to the abstract *overlap space* determined by the modes of overlap together with the natural neighbourhood relation, here called *perturbation*. This enables us to specify the ways in which the mode of overlap can change through time. By attending closely to the character of the transition between two modes we have identified *tangency* as the key feature determining not only which transitions are possible, but also the possible patterns of temporal incidence, e.g., whether or not the tenure of a given overlap mode can be instantaneous, and in which contexts. The notion of *isotopes*—topologically distinct configurations having the same mode of overlap—proved to be of importance here.

We have largely confined our attention to simple two-dimensional regions in a two-dimensional space, but there is no reason why the same ideas should not be applied to regions of higher dimension or different codimension. The two-dimensional case remains of particular interest in the light of potential applications to the qualitative description of geographical space.

Acknowledgments

I should like to thank Achille Varzi for pointing out an number of errors in an earlier draft of this paper.

References

Casati, R. and Varzi, A. (1994). *Holes and Other Superficialities*, MIT Press, Cambridge, MA, and London.

de Kleer, J. and Brown, J. S. (1984). A qualitative physics based on confluences, *Artificial Intelligence* **24**: 7–83.

Egenhofer, M. J. (1989). A formal definition of binary topological relationships, *in* W. Litwin and H.-J. Schek (eds), *Third International Conference on Foundations of Data Organisation and Algorithms (FODO), Paris, France*, Springer-Verlag, New York, pp. 271–286. Lecture Notes in Computer Science, Vol. 367.

Egenhofer, M. J. (1991). Reasoning about binary topological relations, *in* O. Günther and H.-J. Schek (eds), *Advances in Spatial Databases*, Springer-Verlag, pp. 143–160.

Egenhofer, M. J. and Franzosa, R. D. (1991). Point-set topological spatial relations, *Int. Jour. of Geographical Information Systems* **5**: 161–174.

Egenhofer, M. J. and Franzosa, R. D. (1995). On the equivalence of topologicl relations, *Int. Jour. of Geographical Information Systems* **9**: 133–152.

Galton, A. (1993). Towards an integrated logic of space, time, and motion, *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*.

Galton, A. (1994a). Instantaneous events, *in* H. J. Ohlbach (ed.), *Temporal Logic: Proceedings of the ICTL Workshop*, Max-Planck-Institut für Informatik, Saarbrücken.

Galton, A. (1994b). Perturbation and dominance in the qualitative representation of continuous state-spaces, *Technical Report 270*, Department of Computer Science, University of Exeter.

Kuipers, B. J. (1986). Qualitative simulation, *Artificial Intelligence* **29**: 289–338.

Randell, D. A., Cui, Z. and Cohn, A. G. (1992). A spatial logic based on regions and connection, *Proceedings of the Third International Conference on Knowledge Representation and Reasoning*, pp. 165–176. Cambridge, Massachusetts, October 1992.

Williams, B. C. (1990). Temporal qualitative analysis: Explaining how physical systems work, *in* D. S. Weld and J. de Kleer (eds), *Qualitative Reasoning about Physical Systems*, Morgan Kaufmann, San Mateo, California, pp. 133–177.