Approximations of Shape and Configuration Space

Ernest Davis*
New York University
251 Mercer St.
New York, New York
DAVISE@CS.NYU.EDU

September 15, 1995

Abstract

We consider the issue of shape approximation in kinematic mechanical systems; that is, systems of rigid solid objects whose behavior can be characterized entirely in terms of the constraints that each object moves rigidly and that no two objects overlap, without considering masses or forces. The general question we address is the following: Suppose we have calculated the behavior of some kinematic system using ideal descriptions of the shapes of the objects involved. Does it then follow that a real mechanism, in which the shape of the objects approximates this ideal will have a similar behavior? In addressing this question, we present various possible definitions of what it means (a) for one shape to approximate another and (b) for the behavior of one mechanism to be similar to the behavior of another. We characterize the behavioral properties of a kinematic system in terms of its *configuration space*; that is, the set of physically feasible positions and orientations of the objects. We prove several existential theorems that guarantee that a sufficiently precise approximation of shape preserves significant properties of configuration space. In particular, we show that

- It is often possible to guarantee that the configuration space of system A is close to that of system B in terms of metric criteria by requiring that the shapes of A closely approximate those of B in terms of the dual-Hausdorff distance.
- It is often possible to guarantee further that the configuration space of A is topologically similar to that of B by requiring that the surface normals are close at corresponding boundary points of A and B.

In geometric computations for practical purposes, it is often desirable to use an approximation to the true shape. First, approximating a complex shape by a simpler one may substantially reduce computation costs (e.g. [Fleischer et al., 92]). Second, in many cases, the true shape cannot be fully determined. For example, it may not be possible to measure the shape with sufficient accuracy, or the shape may be an ideal to be manufactured using a process with some error tolerance, and so on (e.g. [Joskowicz and Taylor, 94], [Requicha, 83]). (Note that in this second category, the immediate question, "How can the approximation be computed from the exact shape?" does not arise; the approximation is given.) In such cases, it is important to know whether answers to problems computed using the approximations apply to the true situation.

In this paper, we consider the issue of shape approximation in the context of kinematic mechanical systems; that is, systems of rigid solid objects whose behavior can be characterized entirely in terms

*This research was supported by NSF grant #IRI-9300446. Thanks to Leo Joskowicz for many helpful comments and suggestions; also to Dana Angluin, Mladen Covic, Philip Davis, Drew McDermott, Bud Mishra, Igor Najfeld, Chee Yap, and Ken Yip.

of the constraints that each object moves rigidly and that no two objects overlap, without considering masses or forces. As shown by Joskowicz and Sacks [1991], a large number of mechanical devices of practical importance fall into this category. The general question we address is the following: Suppose we have calculated the behavior of some kinematic system using ideal descriptions of the shapes of the objects involved. Does it then follow that a real mechanism, in which the shape of the objects approximates this ideal will have a similar behavior?

To address this question, we must define what it means (a) for one shape to approximate another and (b) for the behavior of one mechanism to be similar to the behavior of another. As regards (a), we will consider two general criteria of shape approximation. The first, in terms of the dual-Hausdorff distance, considers shape **A** approximates **B** if every point in **A** is close to a point in **B** and vice versa, and every point outside **A** is close to a point outside **B** and vice versa. The second criterion, called "approximation in tangent", adds the further requirement that the tangents at corresponding points on the boundaries of **A** and **B** are close.

In addressing (b) above, we will characterize the behavioral properties of a kinematic system in terms of its *configuration space*; that is, the set of physically feasible positions and orientations of the objects. Thus, for the purposes of this paper, we will consider the behavior of two mechanical systems to be "similar" if their respective configuration spaces are close. Again, we will consider several different possible criteria of "closeness" between two configuration spaces; which, if any, of these is appropriate in a given circumstance depends on the particular application and the question being addressed.

Thus, we can reword the general question posed above: If the shapes of one system approximate those of another system, is the configuration space of the first close to the configuration space of the second, under the various definitions of "approximation" of shape and "closeness" of configuration space? In this paper, we shall prove several existential theorems that guarantee that a sufficiently precise approximation of shape preserves significant properties of configuration space. In particular, we show that

- It is often possible to guarantee that the configuration space of system A is close to that of system B in terms of metric criteria by requiring that the shapes of A closely approximate those of B in terms of the dual-Hausdorff distance.
- It is often possible to guarantee further that the configuration space of A is topologically similar to that of B by requiring that the shapes of A approximate those of B in tangent.

1 Regions and mappings

We write "$\mathrm{d}(\mathbf{p},\mathbf{q})$" to denote the Euclidean distance between points $\mathbf{p}$ and $\mathbf{q}$.

Definition 1.1: A *regular region* is a subset of $\Re^n$ that is non-empty, bounded, and equal to the closure of its interior.

Definition 1.2: The distance between two regions $\mathrm{d}(\mathbf{A},\mathbf{B})$ is defined, as usual, as the minimal distance between the two.

$$\mathrm{d}(\mathbf{A},\mathbf{B}) = \inf_{\mathbf{p}\in\mathbf{A},\mathbf{q}\in\mathbf{B}} \mathrm{d}(\mathbf{p},\mathbf{q})$$

Definition 1.3: Two regions *overlap* if their interiors have a non-empty intersection. The degree of overlap of regions **A** and **B**, denoted "$\mathrm{o}(\mathbf{A},\mathbf{B})$" is the radius of the largest sphere contained in $\mathbf{A}\cap\mathbf{B}$.

Lemma 1.1: Regular regions $\mathbf{A}$ and $\mathbf{B}$ overlap iff $\mathrm{o}(\mathbf{A}, \mathbf{B}) > 0$. If regular regions $\mathbf{A}$ and $\mathbf{B}$ overlap, then $\mathrm{d}(\mathbf{A}, \mathbf{B}) = 0$, but not conversely.

Proof: Immediate. ∎.

Definition 1.4: We define three distance functions on regions. The *Hausdorff* distance from $\mathbf{A}$ to $\mathbf{B}$ is defined as the maximum of either the maximal distance from a point $\mathbf{p} \in \mathbf{A}$ to $\mathbf{B}$ or the maximal distance from a point $\mathbf{q} \in \mathbf{B}$ to $\mathbf{A}$.

$$\mathrm{d}_H(\mathbf{R}, \mathbf{S}) = \max(\sup_{q \in S} \inf_{p \in R} \mathrm{d}(\mathbf{p}, \mathbf{q}), \sup_{p \in R} \inf_{q \in S} \mathrm{d}(\mathbf{p}, \mathbf{q}))$$

The *complement-Hausdorff* distance from $\mathbf{A}$ to $\mathbf{B}$, denoted "$\mathrm{d}_{Hc}(\mathbf{A}, \mathbf{B})$", is the Hausdorff distance between the complements of $\mathbf{A}$ and $\mathbf{B}$. The *dual-Hausdorff* distance from $\mathbf{A}$ to $\mathbf{B}$, denoted "$\mathrm{d}_{Hd}(\mathbf{A}, \mathbf{B})$", is the maximum of the Hausdorff distance and the complement-Hausdorff distance.

Lemma 1.2. The Hausdorff distance, the complement-Hausdorff distance, and the dual-Hausdorff distance, are all metrics over the space of regular regions.

Proof: Straightforward. ∎

These three metrics define three different topologies on the space of regular regions. The dual-Hausdorff distance is strictly finer than the other two; the Hausdorff and complement-Hausdorff are incomparable. Therefore, if a function is continuous over the topology defined by the Hausdorff distance or over the topology defined by the complement-Hausdorff distance, it is also continuous over the topology defined by the dual-Hausdorff distance.

Lemma 1.3: For any regular regions $\mathbf{R}$, $\mathbf{R}'$ $\mathbf{S}$, $\mathbf{S}'$,

$$\mathrm{d}(\mathbf{R}, \mathbf{S}) \leq \mathrm{d}(\mathbf{R}', \mathbf{S}') + \mathrm{d}_H(\mathbf{R}', \mathbf{R}) + \mathrm{d}_H(\mathbf{S}', \mathbf{S})$$

Proof: Straightforward. ∎

Corollary 1.4: The distance function "$\mathrm{d}(\mathbf{A}, \mathbf{B})$" is a continuous function over the space of regular regions using the Hausdorff metric.

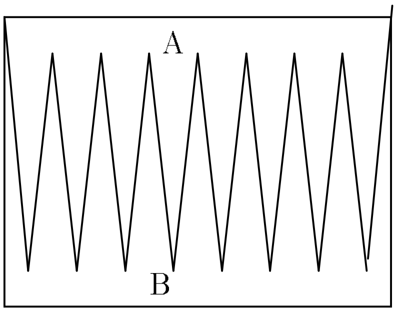
Lemma 1.5: For any regular regions $\mathbf{R}$, $\mathbf{R}'$ $\mathbf{S}$, $\mathbf{S}'$,

$$\mathrm{o}(\mathbf{R}, \mathbf{S}) \leq \mathrm{o}(\mathbf{R}', \mathbf{S}') + \max(\mathrm{d}_{Hc}(\mathbf{R}', \mathbf{R}), \mathrm{d}_{Hc}(\mathbf{S}', \mathbf{S}))$$

Proof: Let $\mathbf{O}$ be a sphere of radius $\mathrm{o}(\mathbf{R}, \mathbf{S})$ contained in $\mathbf{R} \cap \mathbf{S}$. Any point in the complement of $\mathbf{R}'$ must lie within $\mathrm{d}_{Hc}(\mathbf{R}', \mathbf{R})$ of the complement of $\mathbf{R}$. Hence any point in $\mathbf{O} - \mathbf{R}'$ must lie within $\mathrm{d}_{Hc}(\mathbf{R}', \mathbf{R})$ of the boundary of $\mathbf{O}$. Similarly, any point in $\mathbf{O} - \mathbf{S}'$ must lie within $\mathrm{d}_{Hc}(\mathbf{S}', \mathbf{S})$ of the boundary of $\mathbf{O}$. Hence the sphere concentric with $\mathbf{O}$ of radius $\mathrm{o}(\mathbf{R}, \mathbf{S}) - \max(\mathrm{d}_{Hc}(\mathbf{R}', \mathbf{R}), \mathrm{d}_{Hc}(\mathbf{S}', \mathbf{S}))$ must contain no points in the complement of $\mathbf{R}' \cap \mathbf{S}'$. So $\mathrm{o}(\mathbf{R}', \mathbf{S}')$ must be at least $\mathrm{o}(\mathbf{R}, \mathbf{S}) - \max(\mathrm{d}_{Hc}(\mathbf{R}', \mathbf{R}), \mathrm{d}_{Hc}(\mathbf{S}', \mathbf{S}))$. ∎

Corollary 1.6: The overlap function "$\mathrm{o}(\mathbf{A}, \mathbf{B})$" is a continuous function over the space of regular regions using the complement-Hausdorff metric.

Remark: The overlap function is not a continuous function of regions using the Hausdorff metric. For example, in figure 1, both combs $\mathbf{A}$ and $\mathbf{B}$ are close to $\mathbf{R}$ in the Hausdorff metric, but $\mathrm{o}(\mathbf{A}, \mathbf{B}) = 0$, while $\mathrm{o}(\mathbf{R}, \mathbf{R})$ is large. Note that $\mathbf{A}$ and $\mathbf{R}$ are not close in the complement-Hausdorff metric. The standard distance function between regions is not continuous over the complement-Hausdorff metric; the example is analogous. The distance between two regions and the Hausdorff distance between them are invariant under the operation of taking the closure of the argument, but not under the operation of taking the interior. The overlap and the complement-Hausdorff distance are invariant under the operation of taking the interior, but not under the operation of taking the closure.



$\mathbf{R} = \mathbf{A} \cup \mathbf{B}$
The Hausdorff distance from $\mathbf{R}$ to $\mathbf{A}$ is roughly half the width of a tooth.
The complement-Hausdorff distance from $\mathbf{R}$ to $\mathbf{A}$ is half the height of $\mathbf{R}$.

Figure 1: Hausdorff and complement-Hausdorff distances

Definition 1.5: A *placement* is a pair $\langle \mathbf{R}, M \rangle$ of a regular region $\mathbf{R}$ and a rigid mapping M. Intuitively, $\mathbf{R}$ is the region occupied by an object in some standard position, and M is the displacement of the object from that standard. Thus, in placement $\langle \mathbf{R}, M \rangle$ the object occupies position $M(\mathbf{R})$.

Definition 1.6: Let $P_1 = \langle \mathbf{R}, M_1 \rangle$ and $P_2 = \langle \mathbf{R}, M_2 \rangle$ be placements over the same region $\mathbf{R}$. The *placement distance* between P_1 and P_2 is defined as the maximal displacement of any point in $\mathbf{R}$ in going from M_1 to M_2.

$$\mathrm{l}(P_1, P_2) = \max_{\mathbf{r} \in \mathbf{R}} \mathrm{d}(M_1(\mathbf{r}), M_2(\mathbf{r}))$$

Lemma 1.7: For any fixed regular region $\mathbf{R}$, the function $\mathrm{l}(\langle \mathbf{R}, M_1 \rangle, \langle \mathbf{R}, M_2 \rangle)$ has the following properties:

A. $\mathrm{l}(\langle \mathbf{R}, M_1 \rangle, \langle \mathbf{R}, M_2 \rangle)$ is a metric on the space of placements over region $\mathbf{R}$.

B. $\mathrm{l}(\langle \mathbf{R}, M_1 \rangle, \langle \mathbf{R}, M_2 \rangle)$ is invariant under change of coordinate system and under inversion. That is, let T be a linear mapping; and let M_i^{-1} be the inverse of M_i. Then

$$\mathrm{l}(\langle T(\mathbf{R}), M_1 \rangle, \langle T(\mathbf{R}), M_2 \rangle) = \mathrm{l}(\langle \mathbf{R}, M_1 \circ T \rangle, \langle \mathbf{R}, M_2 \circ T \rangle).$$
$$\mathrm{l}(\langle \mathbf{R}, M_1 \rangle, \langle \mathbf{R}, M_2 \rangle) = \mathrm{l}(\langle \mathbf{R}, M_1^{-1} \rangle, \langle \mathbf{R}, M_2^{-1} \rangle).$$

Proof: Immediate. ▮.

Though the metric $\mathrm{l}(\langle \mathbf{R}, M_1 \rangle, \langle \mathbf{R}, M_2 \rangle)$ depends on the region $\mathbf{R}$, the topology generated on the space of mappings is independent of $\mathbf{R}$.

Lemma 1.8: For any placements $\langle \mathbf{R}_A, M_{A1} \rangle$, $\langle \mathbf{R}_A, M_{A2} \rangle$, $\langle \mathbf{R}_B, M_{B1} \rangle$, $\langle \mathbf{R}_B, M_{B2} \rangle$, we have

$$\mathrm{d}(M_{A1}(\mathbf{R}_A), M_{B1}(\mathbf{R}_B)) \leq$$
$$\mathrm{d}(M_{A2}(\mathbf{R}_A), M_{B2}(\mathbf{R}_B)) + \mathrm{l}(\langle \mathbf{R}_A, M_{A1} \rangle, \langle \mathbf{R}_A, M_{A2} \rangle) + \mathrm{l}(\langle \mathbf{R}_B, M_{B1} \rangle, \langle \mathbf{R}_B, M_{B2} \rangle)$$
$$\mathrm{o}(M_{A1}(\mathbf{R}_A), M_{B1}(\mathbf{R}_B)) \leq$$
$$\mathrm{o}(M_{A2}(\mathbf{R}_A), M_{B2}(\mathbf{R}_B)) + \mathrm{l}(\langle \mathbf{R}_A, M_{A1} \rangle, \langle \mathbf{R}_A, M_{A2} \rangle) + \mathrm{l}(\langle \mathbf{R}_B, M_{B1} \rangle, \langle \mathbf{R}_B, M_{B2} \rangle)$$

Proof: Straightforward. ▮

Corollary 1.9: For fixed regular regions $\mathbf{A}, \mathbf{B}, \mathbf{R}$ the functions $\mathrm{d}(M_1(\mathbf{A}), M_2(\mathbf{B}))$ and $\mathrm{o}(M_1(\mathbf{A}), M_2(\mathbf{B}))$ are continuous functions of M_1 and M_2 relative to the metric $\mathrm{l}(\langle \mathbf{R}, M_1\rangle, \langle \mathbf{R}, M_2\rangle)$.

Proof: Immediate from Lemma 1.8. ∎

It is of some interest to define a metric over the space of all placements with the symmetry properties described in lemma 1.7. This can be done in the following steps:

Lemma 1.10: For any two regular regions $\mathbf{A}, \mathbf{B}$ and mappings M_1, M_2,
$\mathrm{l}(\langle \mathbf{A}, M_1\rangle, \langle \mathbf{A}, M_2\rangle) \leq \mathrm{l}(\langle \mathbf{B}, M_1\rangle, \langle \mathbf{B}, M_2\rangle) + 2\mathrm{d}_H(\mathbf{A}, \mathbf{B})$

Proof: Let $\mathbf{a}$ be any point in $\mathbf{A}$ and let $\mathbf{b}$ be the nearest point to $\mathbf{a}$ in $\mathbf{B}$. Then

$$\mathrm{d}(M_1(\mathbf{a}), M_2(\mathbf{a})) \leq \mathrm{d}(M_1(\mathbf{a}), M_1(\mathbf{b})) + \mathrm{d}(M_1(\mathbf{b}), M_2(\mathbf{b})) + \mathrm{d}(M_2(\mathbf{b}), M_2(\mathbf{a})) \leq$$
$$\mathrm{d}_H(\mathbf{A}, \mathbf{B}) + \mathrm{l}(\langle \mathbf{B}, M_1\rangle, \langle \mathbf{B}, M_2\rangle) + \mathrm{d}_H(\mathbf{A}, \mathbf{B}).$$

Taking the maximum over $\mathbf{a} \in \mathbf{A}$ gives the result. ∎

Definition 1.7: The *raw distance* from placement $P_1 = \langle \mathbf{R}_1, M_1\rangle$ to $P_2 = \langle \mathbf{R}_2, M_2\rangle$ is defined as the sum of twice the dual-Hausdorff distance from $\mathbf{R}_1$ to $\mathbf{R}_2$ plus the placement distance between mappings M_1 and M_2 over domain $\mathbf{R}_1 \cup \mathbf{R}_2$.

$$\mathrm{p}_{raw}(\langle \mathbf{R}_1, M_1\rangle, \langle \mathbf{R}_2, M_2\rangle) = 2\mathrm{d}_{Hd}(\mathbf{R}_1, \mathbf{R}_2) + \mathrm{l}(\langle \mathbf{R}_1 \cup \mathbf{R}_2, M_1\rangle\langle \mathbf{R}_1 \cup \mathbf{R}_2, M_2\rangle)$$

Lemma 1.11 Over the space of placements with compact domains, the function p_{raw} has the following properties:

A. It is non-negative, and equal to zero iff $P_1 = P_2$.

B. It is symmetric.

C. It is invariant under linear transformations and inversion.

D. $\mathrm{p}_{raw}(\langle \mathbf{R}_1, M_1\rangle, \langle \mathbf{R}_1, M_2\rangle) \geq 2\mathrm{d}_{Hd}(\mathbf{R}_1, \mathbf{R}_2)$

E. $\mathrm{p}_{raw}(\langle \mathbf{R}_1, M_1\rangle, \langle \mathbf{R}_1, M_2\rangle) \geq \mathrm{l}(\langle \mathbf{R}_1, M_1\rangle, \langle \mathbf{R}_1, M_2\rangle)$

Proof: Immediate from the definition and lemma 1.7. ∎

However, p_{raw} does not obey the triangle inequality. This can be fixed as follows:

Definition 1.8: The *placement distance* from placement $P_A = \langle \mathbf{R}_A, M_A\rangle$ to $P_B = \langle \mathbf{R}_B, M_B\rangle$ is defined as the minimum over all sequences of the form $P_0 = P_A, P_1, P_2 \ldots P_k = P_B$ of the sum $\mathrm{p}_{raw}(P_0, P_1) + \mathrm{p}_{raw}(P_1, P_2) + \ldots + \mathrm{p}_{raw}(P_{k-1}, P_k)$

$$\mathrm{p}(P_A, P_B) = \inf_{P_0 = P_A, P_1, P_2 \ldots P_k = P_B} \sum_{i=1}^{k} \mathrm{p}_{raw}(P_{i-1}, P_i)$$

Theorem 1.12: The function $\mathrm{p}(P_1, P_2)$ is a metric on the space of all placements with regular domains. Moreover, it satisfies properties (C), (D), and (E) of lemma 1.11.

Proof: Symmetry and invariance under linear transformation and inversion follow immediately from lemma 1.11 and the definition. The triangle inequality follows from the construction of definition 1.8; the triangle inequality always holds on the shortest path from A to B. Property (D) holds from the fact that, for any sequence $P_0 \ldots P_k$,

$$\sum_{i=1}^{k} \mathrm{p}_{raw}(P_{i-1}, P_i) \geq \sum_{i=1}^{k} 2\mathrm{d}_{Hd}(P_{i-1}, P_i) \geq 2\mathrm{d}_{Hd}(P_0, P_k)$$

where the second inequality follows from the fact that the dual-Hausdorff distance obeys the triangle inequality.

Property (E) is established as follows: Let $P_0 = \langle \mathbf{R}_0, M_0 \rangle, \ldots P_k = \langle \mathbf{R}_k, M_k \rangle$ be any series of placements from P_0 to P_k. For each placement P_i in this series, we define the cost function Cost(P_i) = $\mathrm{l}(\langle \mathbf{R}_i, M_i \rangle, \langle \mathbf{R}_i, M_k \rangle)$. Now

$$\begin{array}{l}
\mathrm{Cost}(P_{i+1}) + \mathrm{p}_{raw}(P_i, P_{i+1}) = \\
\mathrm{l}(\langle \mathbf{R}_{i+1}, M_{i+1} \rangle, \langle \mathbf{R}_{i+1}, M_k \rangle) + 2\mathrm{d}_{Hd}(\mathbf{R}_i, \mathbf{R}_{i+1}) + \mathrm{l}(\langle \mathbf{R}_i \cup \mathbf{R}_{i+1}, M_i \rangle \langle \mathbf{R}_i \cup \mathbf{R}_{i+1}, M_{i+1} \rangle) \geq \\
\mathrm{l}(\langle \mathbf{R}_{i+1}, M_{i+1} \rangle, \langle \mathbf{R}_{i+1}, M_k \rangle) + 2\mathrm{d}_H(\mathbf{R}_i, \mathbf{R}_{i+1}) + \mathrm{l}(\langle \mathbf{R}_i, M_i \rangle \langle \mathbf{R}_i, M_{i+1} \rangle) \geq \text{(by lemma 1.10)} \\
\mathrm{l}(\langle \mathbf{R}_i, M_{i+1} \rangle, \langle \mathbf{R}_i, M_k \rangle) + \mathrm{l}(\langle \mathbf{R}_i, M_i \rangle \langle \mathbf{R}_i, M_{i+1} \rangle) \geq \text{(by lemma 1.7A)} \\
\mathrm{l}(\langle \mathbf{R}_i, M_i \rangle, \langle \mathbf{R}_{i+1}, M_k \rangle) = \mathrm{Cost}(P_i).
\end{array}$$

Since Cost(P_k)=0, it follows that Cost$(P_0) = \mathrm{l}(P_0, P_k) \leq \sum_{i=1}^{k} \mathrm{p}_{raw}(P_i, P_{i-1})$. Since this holds for every series from $P_A = P_0$ to $P_B = P_k$ it follows that $\mathrm{p}(P_A, P_B) \geq \mathrm{l}(P_A, P_B)$. ∎

Definition 1.8 is non-constructive, and I do not have any strong constructive feeling for the function $\mathrm{p}(P_1, P_2)$, nor any idea how to compute it. I do not even know whether there is, necessarily, a least cost sequence for definition 1.8, nor how long it can be. My hunch is that the minimal cost can be achieved with a quite short sequence (possibly as low as $k = 2$ or $k = 3$) but I have no proof.

2 Configuration space

We will indicate the ith component of a tuple V as $V[i]$, reserving subscripts to distinguish different tuples.

Definition 2.1: A *display* is a k-tuple of regular regions. Intuitively, these are the shapes of k objects. Hence, we will often refer to the indices $1 \ldots k$ as "objects".

Definition 2.2: A display D' is a *contraction* of display D if, for $i = 1 \ldots k$, $D'[i] \subset D[i]$. D' is an *expansion* of D if D is a contraction of D'.

Definition 2.4: A *configuration* is a k-tuple of rigid mappings C. Intuitively, these are the displacements of each object from its standard position as given in a display. The *configuration space on k objects* is the set of all such k-tuples.

Definition 2.5: A *scenario* is a pair of a display and a scenario, or, equivalently, a tuple of placements. If $\langle D, C \rangle$ is a scenario, then, slightly abusing notation, we will write $CD[i]$ for $C[i](D[i])$, the region occupied by the ith object in the scenario.

We extend the metrics d_H, d_{Hc}, d_{Hd} to displays and the metric p to scenarios by taking the max of the function over indices. That is, $\mu(x, y) \equiv \max_i \mu(x[i], y[i])$, where μ is one of the above functions and x and y are displays or scenarios.

For any display D, we define the metric p^D over configurations as $\mathrm{p}^D(C_1, C_2) = p(\langle D, C_1 \rangle, \langle D, C_2 \rangle)$.

The *clearance* of a scenario is the minimal distance between places of two different objects in the scenario. The *maximal overlap* of a scenario is the maximal overlap of two different objects in the scenario.

$$\begin{array}{l}
\mathrm{clearance}(D, C) = \min_{i \neq j} \mathrm{d}(CD[i], CD[j]) \\
\mathrm{overlap}(D, C) = \max_{i \neq j} \mathrm{o}(CD[i], CD[j])
\end{array}$$

Lemma 2.1: Let D and D' be two displays. Then for any configuration C,

clearance$(D, C) \leq$ clearance$(D', C) + 2\mathrm{d}_H(D, D')$.
overlap$(D, C) \leq$ overlap$(D', C) + \mathrm{d}_{Hc}(D, D')$.

Proof: Immediate from lemmas 1.3 and 1.4. ▮

Definition 2.5: A scenario $\langle D, C\rangle$ is *feasible* if overlap$(D, C) = 0$. $\langle D, C\rangle$ is *contact-free* if clearance$(D, C) > 0$. $\langle D, C\rangle$ is *forbidden* if overlap$(D, C) > 0$. For any display D, the set of configurations C such that $\langle D, C\rangle$ is feasible is denoted "free(D)"; the set of configurations C such that $\langle D, C\rangle$ is forbidden is denoted "forbidden(D)"; the set of configurations C such that $\langle D, C\rangle$ is contact-free is denoted "cfree(D)"; and the set of configurations C such that $\langle D, C\rangle$ is not contact-free is denoted "contact(D)".

We now generalize the Hausdorff construction to define a distance function over subsets of any metric space.

Definition 2.6: Let μ be a metric over space $\mathcal{O}$. Let $\mathcal{S}$ and $\mathcal{T}$ be subsets of $\mathcal{O}$. Then the function $\mu_H(\mathcal{S}, \mathcal{T})$ is defined as

$$\mu_H(\mathcal{S}, \mathcal{T}) = \max(\sup_{s \in \mathcal{S}} \inf_{t \in \mathcal{T}} \mu(s, t),\ \sup_{t \in \mathcal{T}} \inf_{s \in \mathcal{S}} \mu(s, t))$$

Thus, the domain of μ_H is $2^{\mathcal{O}} \times 2^{\mathcal{O}}$ and the range is the non-negative reals union infinity. It is easily verified that μ_H is a metric over the space of closed subsets of $\mathcal{O}$ (in the extended sense of "metric" that allows infinite values.)

In particular, applying the Hausdorff construction to the metric p^D on configurations gives a metric p^D_H over the space of closed regions in configuration space.

Theorem 2.2: For any displays D, D', p^D_H(forbidden(D),forbidden(D')) $\leq \mathrm{d}_H(D, D')$.

Proof: Let C be a configuration in forbidden(D). Then there exist objects $I \neq J$ and points $\mathbf{p}, \mathbf{q}$ such that $\mathbf{p} \in$interior$(D[I])$, $\mathbf{q} \in$interior$(D[I])$, and $C[I](\mathbf{p}) = C[J](\mathbf{q})$. By definition of the Hausdorff distance between the two displays, there must exists points $\mathbf{p}' \in$interior$(D'[I])$ and $\mathbf{q}' \in$interior$(D'[J])$ such that $\mathrm{d}(\mathbf{p}, \mathbf{p}') \leq \mathrm{d}_H(D, D')$ and $\mathrm{d}(\mathbf{q}, \mathbf{q}') \leq \mathrm{d}_H(D, D')$.

Let $T[I]$ be the translation mapping point $\mathbf{x}$ to $\mathbf{x}+C[I](\mathbf{p})-C[I](\mathbf{p}')$ and let $T[J]$ be the translation mapping point $\mathbf{x}$ to $\mathbf{x}+C[J](\mathbf{q})-C[J](\mathbf{q}')$. Let C' be the configuration such that $C'[I] = T[I] \circ C[I]$; $C'[J] = T[J] \circ C[J]$; and $C'[K] = C[K]$ for $K \neq I, J$. Then $C'[I](\mathbf{p}') = C'[J](\mathbf{q}')$, so $C' \in$ forbidden(D'). Also, clearly, $\mathrm{p}(C, C') \leq \max(\mathrm{d}(\mathbf{p}, \mathbf{p}'), \mathrm{d}(\mathbf{q}, \mathbf{q}')) \leq \mathrm{d}_H(D, D')$.

Thus every configuration in forbidden(D) is within $\mathrm{d}_H(D, D')$ of some configuration in forbidden(D'), and vice versa, by symmetry. The Hausdorff distance from forbidden(D) to forbidden(D') is therefore at most $\mathrm{d}_H(D, D')$. ▮.

Thus, to calculate the forbidden space of a display to an accuracy ϵ, it is at worst necessary to use shape approximations of accuracy ϵ. The forbidden space is a continuous function of the shapes of the objects involved, and the maximum ratio of the change in space to the change in shape is 1.

The same applies to contact space.

Corollary 2.3: For any displays D, D', p^D_H(contact(D),contact(D')) $\leq \mathrm{d}_H(D, D')$.

Proof: Because the shapes in the display are required to be regular, the space contact(D) is the closure of the space forbidden(D), and the Hausdorff distance between the closure of two sets is equal to the Hausdorff distance between the sets. ▮

Unfortunately, this result does not apply to the free space, which is usually more important than the forbidden space. Free space is discontinuous under small changes in shape. The most we can say is the following:

Starting from any display, a sufficiently small shrinkage in the object shapes, as measured in the complement-Hausdorff distance, gives rise to a small expansion of free space. That is, the free space of a display can be approximated arbitrarily accurately by using a sufficiently precise (in the sense of complement-Hausdorff distance) inscribed approximation of the shapes.

Starting from any display, a sufficiently small expansion in the object shapes, as measured in the Hausdorff distance, gives rise to a small shrinkage of contact-free space. That is, the contact-free space of a display can be approximated arbitrarily accurately by using a sufficiently precise (in the sense of Hausdorff distance) circumscribed approximation of shape.

The remainder of this section formalizes and proves these claims. First we need a few definitions and lemmas:

Definition 2.3: A function $f(D)$ from the space of displays to a topological space $\mathcal{O}$ is *inward (outward) continuous* if the following holds: Let D be a display, and let U be a neighborhood of $f(D)$ in $\mathcal{O}$. Then there exists a neighborhood V of D, such that, for every $D' \in V$, if D' is a contraction (expansion) of D, then $f(D') \in U$. Intuitively, if f is inward continuous and you shrink all the shapes in D by a small amount, then you remain close to $f(D)$. If f is outward continuous and you expand all the shapes in D by a small amount, then you remain close to $f(D)$.

Lemma 2.4: The functions "free(D)" and "cfree(D)" are monotonically non-increasing. That is, if D' is an expansion of D, then free(D') is a subset of free(D) and cfree(D') is a subset of cfree(D).

Proof: Immediate. ∎

Lemma 2.5: Let U be a compact metric space with metric μ. Let f be a non-negative continuous function over U. Then for any $\epsilon > 0$ there exists a $\delta > 0$ such that, for all X, if $f(X) < \delta$ then there exists a Y such that $\mu(X,Y) < \epsilon$ and $f(Y) = 0$. That is, any X with a very small value of f lies close to some Y where $f = 0$.

Proof: Let $B = \{Z \in U \mid \exists_Y \, f(Y) = 0 \land \mu(Z,Y) < \epsilon\}$ and let $U' = U - B$. That is, B is all points within ϵ of some zero of f, and U' is the complement of B in U. Clearly, B is open, so U' is compact. Therefore, f attains a minimum value $\delta > 0$ on U'. This δ then satisfies the condition of the lemma. ∎

Lemma 2.6: Let D be a display, and let U be a compact region of configuration space. For any $\epsilon > 0$ there exists a $\delta > 0$ such that, for every display D', if $\mathrm{d}_{Hc}(D, D') < \delta$ then any configuration in free(D') $\cap\, U$ is within ϵ of some configuration in free(D).

Proof: For any display D and configuration C, let $\mathrm{overlap}^D(C) = \mathrm{overlap}(\langle D, C\rangle)$. By lemma 2.1, if $C' \in$free(D') and $\mathrm{d}_{Hc}(D, D') < \delta$, then $\mathrm{overlap}^D(C') < \delta$. Applying lemma 2.5, with $f(x)$ being the function $\mathrm{overlap}^D$ and with μ being the metric p^D, we infer that we can choose δ so that, for any C' in free(D) $\cap\, U$, if $\mathrm{overlap}^D(C') < \delta$ then there exists a configuration C such that $\mathrm{p}^D(C, C') < \epsilon$ and $\mathrm{overlap}^D(C) = 0$, so $C \in$free(D). ∎

If T and C are two configurations, we will write "$T \circ C$" to mean the configuration $\langle T[1] \circ C[1], \ldots, T[k] \circ C[k]\rangle$.

Definition 2.7: Let f be a function over configuration space; let U be an open region in configuration space; and let T be a tuple of rigid mappings. We say that T preserves f over U if for every configuration $C \in U$, $f(T \circ C) = f(C)$.

Definition 2.8: Let D be a display in n-dimensional space. Let $\Delta = \sum_{i=1}^{k}$ (diameter($D[i]$) + 3). The *basic configuration region* of D is the set of all configurations C such, for every i, $CD[i]$ lies inside the box $[0, \Delta]^n$.

Lemma 2.7: The basic configuration region of D is compact.

Proof: A rigid mapping on n-dimensional Euclidean space can be expressed in a standard way as an $(n+1)^2$ matrix. Hence, a configuration on a k-object display D can be viewed as a $k \cdot (n+1)^2$ vector dimensional vector and the configuration space as a whole is a $kn(n-1)/2$ dimensional manifold of such vectors. It is easily verified (a) that the standard topology over the manifold is the same as the topology defined by the metric $\mathrm{p}^D(C_1, C_2)$; (b) that a set of configurations is bounded in the manifold if and only if it is bounded relative to the metric $\mathrm{p}^D(C_1, C_2)$; (c) that the basic configuration region is closed and bounded relative to the metric $\mathrm{p}^D(C_1, C_2)$. Hence, the basic configuration region is closed and bounded in the manifold; hence it is compact. ∎

We now define two functions over configuration space. $\Gamma^D(C)$ maps each configuration C into the basic configuration region, in a way that preserves the relative position of nearby objects. $\Theta^D(C)$ is the tuple of rigid mappings such that $\Theta^D(C) \circ C = \Gamma^D(C)$.

Definition 2.9: Let C be any configuration over display D. We say that C has a *gap* in direction $\hat{u}$ from A to B, $A < B$ if

- There is an object I such that maximal coordinate of $CD[I]$ in the $\hat{u}$ direction is equal to A;
- There is an object J such that minimal coordinate of $CD[J]$ in the $\hat{u}$ direction is equal to B; and
- There is no object containing any point whose $\hat{u}$ coordinate is between A and B.

The size of this gap is $B - A$.

Definition 2.10: Let C be any configuration over display D. We define the configuration $\Theta^D(C)$ using the following algorithm: In each coordinate direction $\hat{u}$, we begin by translating the whole configuration so that the lowermost point in the $\hat{u}$ direction in the scenario has coordinate 1. We then look for gaps $[A, B]$ of size greater than 3. We reduce any such gap to being of size exactly 3 by translating all the objects above the gap by a distance $B-(A+3)$, while leaving all the objects below the gap where they are. We repeat until all such gaps are eliminated. We carry out this procedure in the $\hat{x}$, $\hat{y}$, and $\hat{z}$ directions. When all this is complete, the final configuration is $\Gamma^D(C)$ and the combined transformations give $\Theta^D(C)$. Table 1 displays this algorithm in pseudo-code. It is easily seen that this procedure gives a unique result; in particular, that the result does not depend on the order in which gaps are reduced or coordinate directions are considered.

Definition 2.11: The function "$\mathrm{bclear}^D(C)$" is defined as $\min(\mathrm{clearance}(D, C), 1)$.

Lemma 2.8: The functions $\Gamma^D(C)$ and $\Theta^D(C)$ have the following properties:

a. $\Gamma^D(C) = \Theta^D(C) \circ C$.

b. Both $\Gamma^D(C)$ and $\Theta^D(C)$ are continuous functions of C.

c. If two objects are within distance 3 in C then their relative position is the same in $\Gamma^D(C)$ as in C.

d. The distance between two objects is greater than or equal to 3 in C if and only if it is greater than or equal to 3 in $\Gamma^D(C)$.

e. $\Gamma^D(C)$ is in the basic configuration region, a distance 1 from the boundaries of that region.

f. Let U be the open ball of radius 1 around C in configuration space. Then the function $\lambda(C')(\Theta^D(C)) \circ C'$ maps U into the basic configuration region and it preserves the functions $\mathrm{overlap}^D$ and bclear^D over U.

```
begin T = the $k$-tuple of identity mappings;
    for $\hat{u}$ in $\{\hat{x}, \hat{y}, \hat{z}\}$ do /* $\hat{u}$ loops over the coordinate directions */
        begin let $L$ be the minimum coordinate over all objects $I$
                        of the $\hat{u}$ coordinate of $TCD[I]$;
            let $W$ be the configuration all of whose elements are the translation
                        $\lambda(\vec{P})(\vec{P} + (1 - L)\hat{u})$;
            $T := W \circ T$;
            loop let $[A, B]$ be any gap in $T(C)$ in the $\hat{u}$ direction of size $> 3$;
                if there is no such gap, exitloop
                 else let $W$ be the configuration defined as follows:
                        for $I = 1 \dots K$ do
                            if $TC[I](D[I])$ has $\hat{u}$ coordinates less than $A$
                              then $W[I]$ is the identity
                              else $W[I]$ is the mapping $\lambda(\vec{V})(\vec{V} - (B - (A + 3))\hat{u})$;
                    $T := W \circ T$;
            endloop
        end
    $\Theta^D(C) := T$; $\Gamma^D(C) = T \circ C$
end.
```

Table 1: Computing $\Theta^D(C)$

Proof:

a. Immediate by construction.

b. Within a space of configurations all of which have the same system of large gaps (gaps of size greater than 3 in the same coordinates between the same objects), the continuity of Θ^D is immediate. Moreover, in any such space, as the size of a particular gap approaches 3, the transformation associated with closing the gap approaches the identity, which is its value once the size of the gap becomes 3. Thus Θ^D is continuous. The continuity of Γ^D follows immediately.

c. If two objects are within distance 3 in C, then there cannot be any large gaps between them. Hence, all the intermediate transformations W move them together, thus preserving their relative position.

d. If the two objects are originally on opposite sides of some large gap, that gap will be reduced to size 3, so they are still on opposite sides of a gap of size 3, and hence at least 3 apart. If they were distance 3 apart but not on opposite sides of any gap in C, then the transformations will move them together, their relative positions will be unchanged, and their distance will be unchanged. The converse argument hold the same way.

e. The range from the minimal coordinate of any point in any object in the x direction in $\Gamma^D(C)$ to the maximal coordinate is at most the sum of the diameters of the objects plus 3 times $K-1$, where K is the number of objects. Proof: Order the objects by increasing order of their lowest x coordinate. The lowest x coordinate of object J is at most 3 greater than the maximal x coordinate of some preceding object $I < J$, and hence at most 3 + diameter($D[I]$) greater than the minimal x coordinate of I. Moreover, the objects are placed so that the minimal coordinate in each direction is 1. Therefore, all the objects in $\Gamma^D(C)$ lie in the box $[1, \Delta - 1]^3$. Thus,

$\Gamma^D(C)$ lies in the basic configuration region, as does any configuration within distance 1 of $\Gamma^D(C)$.

f. The fact that U is in the basic configuration region is immediate from (e). The fact that overlap^D and bclear^D are preserved is immediate from (c) and (d).

▮

Lemma 2.9: Let D be a display. For any $\epsilon > 0$ there exists a $\delta > 0$ such that, for every display D', if $\mathrm{d}_{Hc}(D, D') < \delta$ then any configuration in free(D') is within ϵ of some configuration in free(D). (This is the same as lemma 2.6, but with the restriction to a compact region dropped.)

Proof: Let Q be the basic configuration region of D. Let $\epsilon > 0$. Let δ_0 be chosen to satisfy lemma 2.6 for the value $\min(\epsilon, 1)$, and let $\delta < \min(\delta_0, 1, \min_i \mathrm{d}_{Hc}(D[i], \emptyset))$. (The complement-Hausdorff distance from a region $D[i]$ to the empty set is well-defined and finite; it is equal to the radius of the largest circle that can be inscribed in $D[i]$.) Let display D' be then chosen so that $\mathrm{d}_{Hd}(D, D') < \delta$ and let $C' \in$free(D'). Let E be the display such that $E[i] = D'[i] \cap D[i]$; then clearly $C' \in$free(E). It is immediate that $\mathrm{d}_{Hc}(E, D) \leq d_{Hc}(D', D) < \delta$. $E[i]$ is non-empty, because of the constraint that $\delta < \mathrm{d}_{Hc}(D[i], \emptyset)$.

Let U be the ball of radius 1 around C'. By lemma 2.8, any two objects in E that do not overlap in C' also do not overlap in $\Gamma^D(C')$, so $\Gamma^D(C') \in$free(E). By lemma 2.6, there is a configuration $C_1 \in$free(D) such that $p^D(C_1, C') < \min(\epsilon, 1)$. Thus, $C_1 \in \Theta^D(C') \circ U$. Let $C = (\Theta^D(C'))^{-1} \circ C_1$. Then since $C \in U$, $\Theta^D(C')$ preserves overlap^D on C, and hence $C \in$free(D). ▮

Theorem 2.10: The function "free(D)", mapping a display to a region in configuration space, is inward continuous, using the metric d_{Hc} on displays and the metric p_H on regions of configuration space.

Proof: What this says, unwrapping the definitions, is that, if D' is a contraction of D and is close enough to D in the complement-Hausdorff metric then

a. Every configuration in free(D) is close to some configuration in free(D'). This is trivial, since free(D) is a subset of free(D') (lemma 2.4).

b. Every configuration in free(D') is close to some configuration in free(D). This is lemma 2.9.

▮

Lemma 2.11: Let U be a compact metric space with metric μ and let f be a continuous, real-valued function over U. Then, for any $\epsilon > 0$, there exists a $\delta > 0$, such that every point $X \in U$ where $f(X) > 0$ is within ϵ of a point Y where $f(Y) > \delta$.

Proof: Fix a value $\epsilon > 0$. For any $\delta > 0$, define the set $Q_\delta = \{Y \mid \exists_X f(X) > \delta \wedge \mu(X, Y) < \epsilon\}$. Thus, for any $b < a$, $Q_b \supset Q_a$. Let W be the closure of the set $\{X \mid f(X) > 0\}$; then clearly $\cup_{\delta > 0} Q_\delta \supset W$. Since W is compact, there must be a finite subcover of W, $Q_{\delta 1}, Q_{\delta 2} ... Q_{\delta k}$. Choosing δ to be the smallest of these δ_i, we infer that $Q_\delta \supset W$, which is the desired result.

Lemma 2.12: Let D be a display, and let U be a compact region of configuration space. For any $\epsilon > 0$ there exists a $\delta > 0$ such that, for every display D', if $\mathrm{d}_H(D, D') < \delta$ then any configuration in cfree$(D) \cap U$ is within ϵ of some configuration in cfree(D').

Proof: Applying lemma 2.11, with $f(x)$ being the function bclear^D and with μ being the metric p^D, we infer that we can choose δ so that, for any C in cfree$(D) \cap U$, since bclear$^D(C) > 0$, there exists a configuration C' such that $\mathrm{p}^D(C, C') < \epsilon$ and bclear$^D(C') > \delta$. By lemma 2.1, if bclear$^D(C') > \delta$ and $\mathrm{d}_H(D, D') < \delta$, then bclear$^{D'}(C') > 0$, so $C' \in$cfree(D'). ▮

Lemma 2.13: Let D be a display. For any $\epsilon > 0$ there exists a $\delta > 0$ such that, for every display D', if $d_H(D, D') < \delta$ then any configuration in cfree(D) is within ϵ of some configuration in cfree(D'). (This is the same as lemma 2.12, dropping the restriction to a compact region of configuration space.)

Proof: The proof is exactly analogous to the proof of lemma 2.9 (in fact, slightly simpler): Start with any configuration $C \in$cfree(D). Use lemma 2.8 to find the mapping $\Theta^D(C)$ taking C into the basic configuration region. Use lemma 2.12 to find a configuration to find a configuration $C_0 \in$cfree(D') within ϵ of $\Gamma^D(C)$. Then mapping back $C' = (\Theta^D(C))^{-1} \circ C_0$ gives the desired answer. ∎

Theorem 2.14: The function "cfree(D)", mapping a display to a region in configuration space, is outward continuous, using the metric d_H on displays and the metric p_H on regions of configuration space.

Proof: Unwrapping the definitions, this says that, if D' is an expansion of D and is close enough to D in the Hausdorff metric then

a. Every configuration in cfree(D) is close to some configuration in cfree(D'). This is lemma 2.13.

b. Every configuration in cfree(D') is close to some configuration in cfree(D). This is trivial, since cfree(D') is a subset of cfree(D) (lemma 2.4).

∎

If free(D) space is equal to the closure of cfree(D), which is true for almost every display D, we can allow both expansion and contraction as long as the dual-Hausdorff distance is small.

Theorem 2.15: If free(D) is the closure of cfree(D), then both free($\cdot$) and cfree($\cdot$) are continuous in a neighborhood of D using the metric d_{Hd} on displays and the metric p_H on regions of configuration space.

Proof: Immediate from lemmas 2.9 and 2.13. ∎

3 Paths

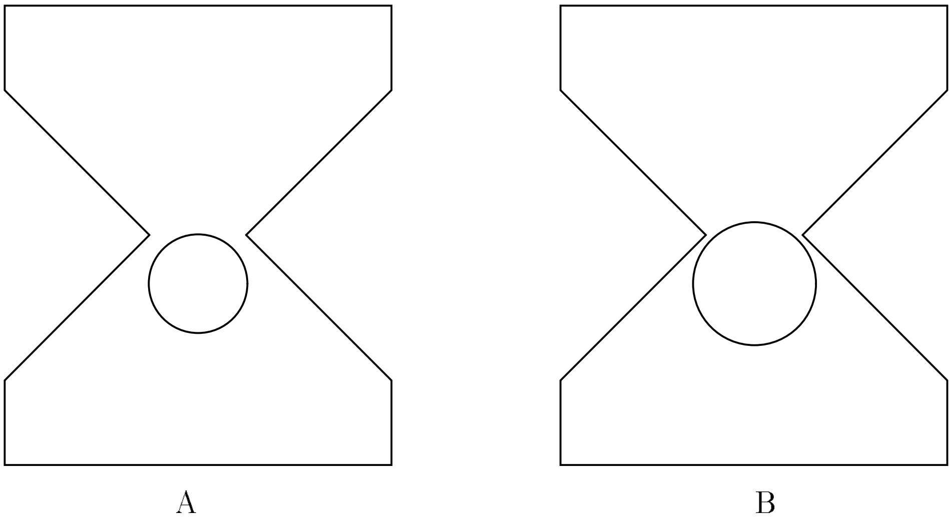
For most purpose, what is significant about a system of rigid objects is not so much the shape of the space of free configurations as the nature of the set of paths through that space, and systems are similar to the extent that the sets of free paths are similar. However, there is more than one possible measure of closeness in terms of the paths through a given space. For example, we might say that spaces S and T are close if

(1) The trace of every path through S is close, in terms of Hausdorff distance, to a path through T, and vice versa.

(2) Any path through S can be tracked closely by a path through T, and vice versa.

(3) Say that points $s \in S$ and $t \in T$ correspond closely if any path p starting at s can be tracked closely by a path q starting at t and vice versa. Say that S and T are close if every point $s \in S$ has a closely corresponding point $t \in T$ and vice versa.

Taking (0) to be the criterion that the Hausdorff distance from S to T is small, it is immediate that (3) implies (2) implies (1) implies (0). Figures 2 -4 illustrate that none of the reverse implications hold. In figure 2, the two free spaces are close in the Hausdorff distance, but paths in A that go from bottom to top are not close in Hausdorff distance to any path in B. In figure 3, the trace of any path



In A, the configurations where the ball is on top and those where it is on the bottom are in a single connected component of free space. In B, they are two separate connected components.

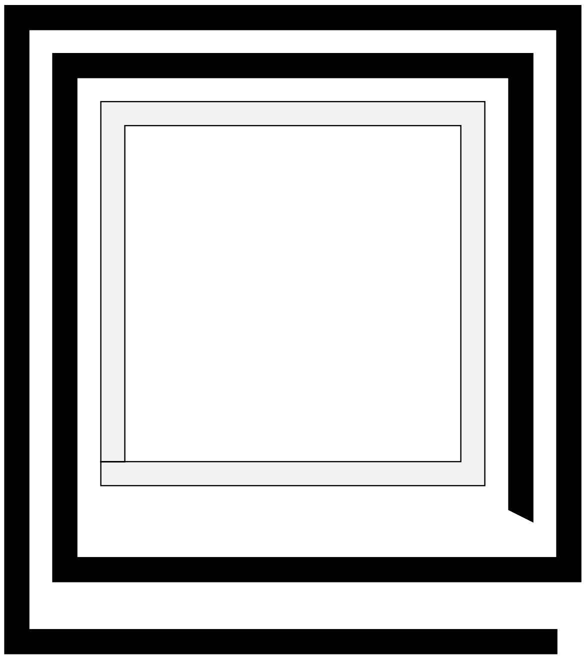
Figure 2: Free space vs. path traces

through A lies close to the trace of some path through B, and vice versa; but no detective in B can tail a criminal in A for more than two revolutions. In figure 4, any path in A can be traced by a path in B, but the "stuck" position in A has no corresponding configuration in B, since every position in B leads to arbitrarily long paths which cannot be tracked from the stuck position.

In this section, we will address criteria (1) and (2). We will show that the analogue of theorem 2.14 holds for criterion (1) and that, if certain anomalous shapes are excluded, the analogues of theorems 2.13 and 2.14 hold for both criteria (1) and (2). (The analogues of theorem 2.2 also hold for criteria (1) and (2), but these are both trivial and unimportant, as no one cares about the space of forbidden paths.)

Criterion (3) is more difficult; establishing it requires constraining the surface normal to the objects. The analysis needed is quite different in character. This will be discussed in section 4.

Figure 5 illustrates an anomalous system of a type that creates difficulties for many of these theorems. The system consists of two objects, a sphere of radius 1 and a frame. The frame nearly meets the sphere at its bottom point and close to its top point. The lower frame surface just is a horizontal circular disk of radius 1 about the z-axis in the plane $z = -1$. The upper frame surface consists of the paraboloid $z = 1 - (1-r)^2/50$ for $0 \leq r \leq 1$, where r is the distance from the z-axis, with a spiral track carved out of it. The center of this track follows the curve $r = 1 - 2^{-\theta/2\pi}$, where $\theta \in [0, \infty)$ is the angle in the $x - y$ plane around the z-axis. This curve is a spiral that goes through infinitely many revolutions, converging on the circle of radius 1 as a limit cycle. At the center of the track, the height of the surface is $1 + (1-r)^2/1000$. A radial cross section of one notch in the track, centered at angle θ_0 and radius $r_0 = 1 - 2^{\theta_0/2\pi}$, follows the circular arc of radius 1 and apex at $r = r_0, z = 1 + (1-r_0)^2/1000$, until meeting the parabolic cross section. These curves meet at roughly the points $r = 1 - .8(1-r_0)$ and $r = 1 - 1.2(1-r_0)$, so consecutive tracks do not overlap. Thus, the track is carved out so that the sphere fits with a little room to spare, but cannot slide from one groove to the next. The remainder of the frame is placed so as to hold these two surfaces in place, without interfering with the motion of the sphere.



A is the closed square (in grey); B is the spiral (in black)
Any path in A is close to a path in B,
but paths in A that go around more than twice cannot be closely tracked in B.

Figure 3: Path traces vs. tracking paths

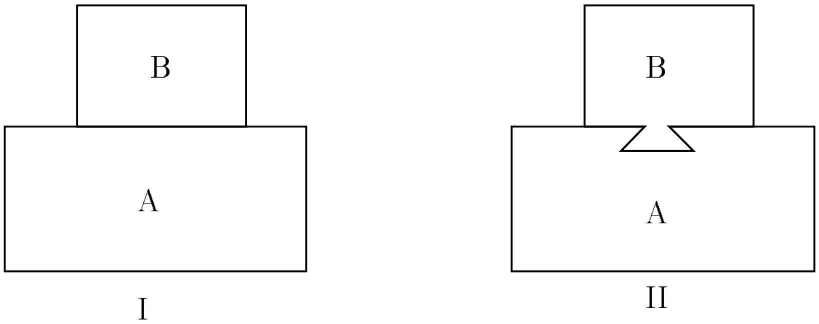
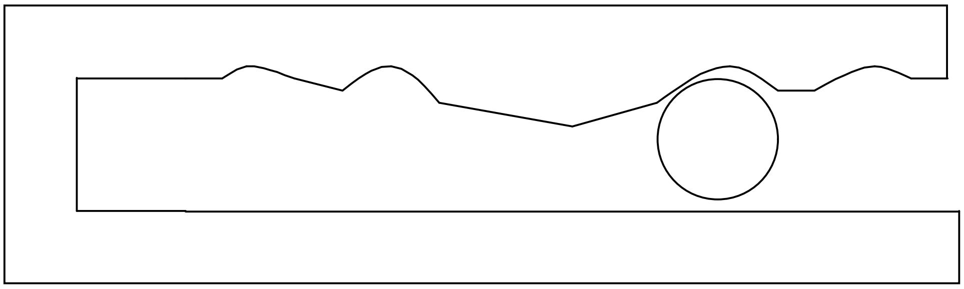
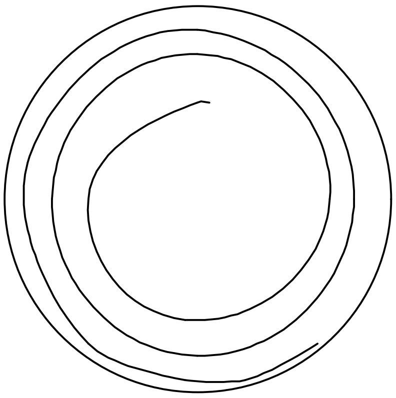


Figure 4: Path tracking vs. close correspondence of points



Cross-section of frame and sphere in the $x - z$ plane.
(Note: this diagram is inaccurate and not to scale; it is meant only to be suggestive.
Two notches of the track on either side are shown; in fact, there are infinitely many.)



The beginning of the spiral that forms the center of the track.
The spiral has infinitely many revolutions, converging to the outer circle.

Figure 5: Anomalous kinematic system

Therefore, this display permits a contact-free motion of the sphere around the track arbitrarily often. It does not contain any paths that allow the sphere to escape the track. However, if the frame is expanded by an arbitrary δ, then all the grooves in the channel of depth less than δ will be filled in, and the sphere can only execute finitely many revolutions in the track. If the frame is contracted by distance δ, then, where the height of the paraboloid is greater than $1-\delta$, the sides of the track no longer hold the sphere, which can then escape to the outside of the frame. Thus, the system is not continuous inward even in the Hausdorff distance between paths; arbitrarily small shrinkage of the frame allows the sphere to execute a path from the inside of the track to far away, a path that is not close to any path in the original system. It is not continuous outward in the path-distance between paths; an arbitrarily small swelling of the frame confines the sphere to N revolutions in the track, and a path of N revolutions cannot closely track a path of $N+1$ revolutions. (Incidentally, the construction here relies on a three-dimensional physical space. I do not know whether analogous constructions can be found in a two-dimensional space, or whether stronger theorems apply there.)

However, small expansion of the frame does allow paths that are close to the original paths in terms of the Hausdorff distance. This, it turns out, holds in general, over compact regions of configuration space. (Theorem 3.3 below).

Definition 3.1: Let U be a metric space with metric μ. A *path* in U is a continuous function from the real interval [0,1] into U. The *trace* of a path is its image in U.

$$\text{tr}(\phi) = \{ \ \phi(T) \mid T \in [0,1] \ \}$$

The distance measure between two paths ϕ_1 and ϕ_2 in criterion (1) above is then just the Hausdorff distance relative to μ between the respective traces, $\mu_H(\text{tr}(\phi_1),\text{tr}(\phi_2))$. The distance measure in criterion (2), which we will denote $\mu_t(\phi_1,\phi_2)$ is defined as the maximal distance between corresponding points on the paths.

$$\mu_t(\phi_1,\phi_2) = \max_{t\in[0,1]} \mu(\phi_1(t),\phi_2(t))$$

Definition 3.2: Let ϕ be a path in configuration space and let D be a display. We define overlap$^D(\phi)$ as the maximal value of overlap^D over ϕ and clearance$^D(\phi)$ as the minimal value of clearance^D over ϕ. Path ϕ is *free* over D if overlap$^D(\phi) = 0$; otherwise, it is *forbidden*. Path ϕ is *contact-free* over D if clearance$^D(\phi) > 0$; The function "path_free(D)" maps a display D into the set of free paths over D; the function "path_cfree(D)" maps a display D into the set of contact-free paths over D.

Lemma 3.1: Let U be a compact space with metric μ. Let Φ be the space of all closed bounded subsets of U. Then Φ is compact under the Hausdorff distance μ_H.

Proof: See [Munkres, 75, p. 279].

Lemma 3.2: Let U be a compact space with metric μ and let f be a continuous function on U. For any $\epsilon > 0$ there exists a $\delta > 0$ such that, for any path ϕ on which f is always greater than 0, there exists a path ϕ' on which f is always greater than δ such that $d_H(\text{tr}(\phi),\text{tr}(\phi')) < \epsilon$.

Proof: Suppose not. Then we can construct the following sequence of paths and quantities.

ϕ_0 is any path on which f is always greater than 0. δ_0 is the minimal value of f on ϕ_0.

ϕ_1 is a path on which f is always greater than 0 such that no path within ϵ of ϕ_1 is always greater than δ_0. δ_1 is the minimal value of f on ϕ_1.

...

ϕ_i is a path on which f is always greater than 0 such that no path within ϵ of ϕ_i is always greater than δ_{i-1}. δ_i is the minimal value of f on ϕ_i.

...

Then, by construction, the Hausdorff distance between ϕ_i and ϕ_j is always greater than ϵ, for all $i \neq j$. But this is impossible, since by lemma 3.1, these are elements in a compact space.

Theorem 3.3: Let D be a display and let $\epsilon > 0$. Let U be a compact region of configuration space. Then there exists a $\delta > 0$ such that, for any display D', if $\mathrm{d}_H(D, D') < \delta$ then any path ϕ through U that is contact-free over D is within ϵ of some path ϕ' through U that is contact-free over D'.

Proof: Applying lemmas 3.1 and 3.2 to the function clearance^D, infer that there is a $\delta > 0$ such that, for every ϕ through U such that clearance$^D(\phi) > 0$, there exists a path ϕ' through U within ϵ of ϕ such that clearance$^D(\phi') > \delta$. By lemma 2.1, ϕ' is contact-free in D'. ▮

Somewhat surprisingly, theorem 3.3 does not generalize to the case of paths through unbounded configuration space. Imagine taking the frame with the infinite spiral in figure 5, and moving it the whole system through space in the z direction at unit speed, while spinning the ball along the track at one revolution per unit time. Now, if the track is slightly filled in, so that the ball can only execute finitely many revolution, the Hausdorff distance between the path on the finite track and path on the infinite track will become 1 within half a revolution of the time that the ball reaches the end of the finite track. The Hausdorff distance between traces, since it compares position at one time to position at another, is not invariant under uniform change in velocity, a substantial defect in this measure.

To generalize theorem 3.3 to unbounded space, or to establish that path_free is inward continuous, we must rule out systems such as those of figure 5, which are undoubtedly not things you can pick up in most hardware stores. One condition that suffices is to require that the configuration space be *ordinarily connected.*

Definition 3.3: A region O in a metric space is *ordinarily connected* if the following holds: Let S be a solid sphere, and let $P = S \cap O$. Then (i) P has finitely many connected components, and (ii) any connected component of P is path-connected.

Theorems 3.7 and 3.8 will apply to configuration spaces that are ordinarily connected. It is known [Mishra, 94] that every semi-algebraic set is ordinarily connected. Since a semi-algebraic display gives rise to a semi-algebraic configuration space, it suffices for the application of the theorem that the shapes in the display be all semi-algebraic. (I also conjecture that theorems 3.7 and 3.8 hold for all two-dimensional displays of regular, connected objects, though not all such displays are ordinarily connected.)

Lemma 3.4: Let O be ordinarily connected, and let C be a connected component of O. Then there is an open set $U \supset C$ which is disjoint from any other connected component of O.

Proof: By definition of connectivity, given any two connected components C and C', there are disjoint open regions $V \supset C$ and $V' \supset C'$. By the definition of ordinary connectivity, there are only finitely many such C'. Therefore, the intersection of all these regions V is the desired open set. ▮

Lemma 3.5: Let O be a compact space with metric μ which is ordinarily connected. Then for any $\epsilon > 0$ there exists a $\delta > 0$ such that for any $p, q \in O$, if $\mu(p, q) < \delta$ then there is a path ϕ from p to q through O such that ϕ is always within ϵ of p.

Proof: Suppose not. Then there exists an ϵ such that for any N we may choose points p_N, q_N for which $\mu(p_N, q_N) < 1/N$ but there is no path from p_N to q_N which remains within ϵ of p_N. Consider the sequence of pairs $\langle p_1, q_1\rangle, \langle p_2, q_2\rangle, \langle p_3, q_3\rangle, \ldots$ Since O is compact, there is a subsequence of the points p_k that converges to a limit point $p \in O$. Clearly, p is the limit of the corresponding q_k as well.

Let S be the subset of O within distance $\epsilon/2$ of the point p. Let R be the connected component of S containing p. By lemma 3.4, there is an open set U containing R and disjoint from $S - R$. Since U is open, it must contain all the points p_N and q_N in the subsequence converging to p for all sufficiently large N. Therefore, these points must be in R. But R is path-connected, and it lies in a sphere of radius $\epsilon/2$, so there must be a path from $\mathbf{p}_N$ to $\mathbf{q}_N$ that stays within ϵ of $\mathbf{p}_N$. This complete the contradiction. ∎

Lemma 3.6: Let O be a compact space with metric μ which is ordinarily connected. Then for any $\epsilon > 0$ there exists a $\delta > 0$ for which the following holds: Let Q be any space within Hausdorff distance δ of O: $\mu_H(O, Q) < \delta$. Let ψ be any path through Q. Then there exists a path ϕ through O such that the path distance $\mu_t(\phi, \psi) < \epsilon$.

Proof: By lemma 3.5, we can choose a value δ_1 such that, if the distance between two points $q, r \in O$ is less than δ_1, then there is a path from q to r that lies within $\epsilon/3$ of q. I claim that if δ is chosen less than $\min(\epsilon/3, \delta_1/3)$ then the theorem holds.

Let Q and ψ be chosen as above. We can choose a finite sequence of points $\langle p_0, p_1 \ldots p_k \rangle$ on ψ such that p_0 is the starting point of ψ, p_k is the ending point of ψ, and the section of the path ψ between p_i and p_{i+1} is always within $\delta/3$ of p_i. Since the Hausdorff distance from O to Q is less than δ, we may choose points $q_0, q_1 \ldots q_k$ in O such that each q_i is within δ of p_i. Therefore, the distance from q_i to q_{i+1},

$$\mu(q_i, q_{i+1}) < \mu(q_i, p_i) + \mu(p_i, p_{i+1}) + \mu(p_{i+1}, q_{i+1}) < \delta_1$$

By definition of δ_1, therefore, there is a path from q_i to q_{i+1} that remains within $\epsilon/3$ of q_i. Therefore, any point on the path ψ between p_i to p_{i+1} lies within ϵ of the corresponding point between q_i and q_{i+1}. We now string together the paths between q_i and q_{i+1}, and the proof is complete. ∎

Let $\mathrm{p}_t(\phi, \phi')$ be the path distance between paths ϕ and ϕ'. Applying the Hausdorff construction over the space of paths, we can define a distance measure p_{tH} between one set of paths and another. This, in turn, defines a topology over the space of sets of paths.

Theorem 3.7: Let $\mathcal{D}$ be the space of displays D such that free(D) is ordinarily connected. Topologize $\mathcal{D}$ using the complement-Hausdorff metric d_{Hc} and topologize the space of sets of paths using the metric p_{tH}, the Hausdorff construction applied to the path-distance. Then the function "path_free(D)", mapping a display D to the set of paths free over D, is inward continuous over $\mathcal{D}$.

Proof: What this states is, that for any D and ϵ, δ can be chosen such that, if D' is a contraction of D and $\mathrm{d}_{Hc}(D, D') < \delta$ then (a) every path in path_free(D) is near a path in path_free(D') and (b) every path in path_free(D') is near a path in path_free(D). Part (a) is trivial, since path_free(D') is a superset of path_free(D).

We establish part (b) in two steps. First, consider only the space of paths that remain in the basic configuration region. Since this region is compact, (b) follows immediately from lemmas 3.6 and theorem 2.10. Thus, for any $\epsilon > 0$ and display D, there exists a δ such that, if $\mathrm{d}_{Hc}(D, D') < \delta$ and ϕ' is a path in path_free(D') that stays in the basic configuration region, then there exists a path ϕ in path_free(D) such that $\mathrm{p}_t(\phi, \phi') < \epsilon$.

Second, we use lemma 2.8 to reduce configuration space as a whole to the basic configuration region, as follows: Assume w.l.o.g. that $\epsilon < 1$. Using the first step above, choose a suitable δ for ϵ and D. Using lemma 2.8, construct a function $\Gamma^D(C)$ from configuration space into the basic configuration region over D and let $\Theta^D(C)$ be the transformation from C to $\Gamma^D(C)$. Let D' be any display such that $\mathrm{d}_{Hc}(D, D') < \delta$ and let ϕ' be any path in free_space(D'). Let $\phi_1(T) = \Gamma^D(\phi'(T))$, and let $\theta(T) = \Theta^D(\phi'(T))$. By lemma 2.8, ϕ_1 stays in the compact basic configuration region. Also by lemma 2.8, since Γ^D preserves the function overlap$^{D'}$, overlap$^{D'}(\phi_1)$ = overlap$^{D'}(\phi) = 0$, so ϕ_1 is in path_free(D'). By lemma 3.6, there is a path ϕ_2 in path_free(D) such that $\mathrm{p}_t(\phi_2, \phi_1) < \epsilon$. Now, let $\phi(T) = (\theta(T))^{-1}(\phi(T))$. It follows immediately from lemma 2.8 that $\mathrm{p}(\phi, \phi') < \epsilon$ and that $\phi \in$path_free(D).∎

Theorem 3.8: Let $\mathcal{D}$ be the space of displays D such that free(D) is ordinarily connected. Topologize the space of displays using the Hausdorff metric d_H and topologize the space of sets of paths using the metric p_{tH}. Then the function "path_cfree(D)", mapping a display D to the set of paths contact-free over D, is outward continuous over $\mathcal{D}$.

Proof: The proof is essentially identical to that of theorem 3.7, using theorem 2.14 instead of theorem 2.10.

As in theorem 2.15, in the case of displays whose free space is the closure of their contact-free space, we can allow change in either direction:

Theorem 3.9: Let $\mathcal{D}$ be the space of displays D such that free(D) is ordinarily connected. Topologize the space of displays using the dual-Hausdorff metric d_{Hd} and topologize the space of sets of paths using the metric p_{tH}. Let D be a display such that free(D) is the closure of cfree(D). Then the functions path_free($\cdot$) and path_cfree($\cdot$) are continuous over a neighborhood of $D \in \mathcal{D}$.

Proof: Analogous to the proofs of theorems 3.7 and 3.8. ∎

4 Lifting from contact and the approximation of tangents

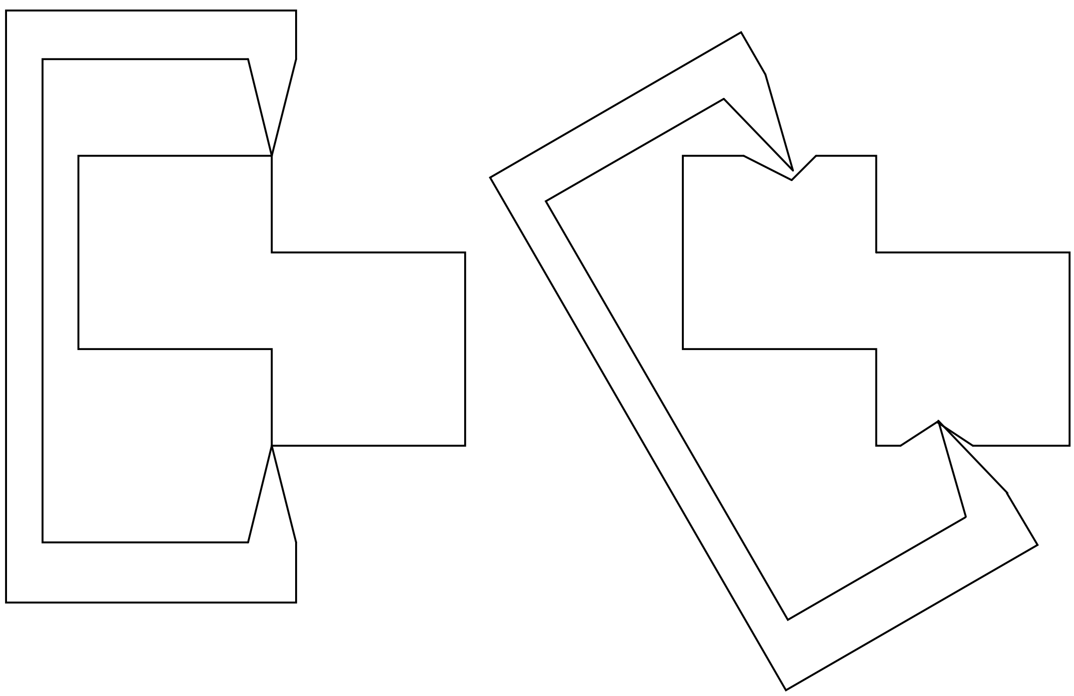
As figure 4 illustrates, it is possible to change a system where all feasible configurations are connected to one that contains configurations that are "stuck" using changes that can be arbitrarily small in both the Hausdorff and complement-Hausdorff distance, and that can be either expansions or contractions. The kinds of criteria that we have used in the previous sections thus do not suffice to prevent this kind of change in the free space.

Fortunately, figure 4 also suggests where a solution might lie. The construction of the "hook and eye" that so radically changes the behavior of the object requires only a small variation in terms of the points occupied by the object, but requires a large change in the local surface normal. This suggests that if the approximation is required to be close both in distance and in the direction of the local surface normal, then it may be possible to infer that a small modification does not generate new connected components in configuration space.

This is indeed possible, at least two dimensions. In this section, we will prove the following result: Given a display consisting of two-dimensional objects such that in every configuration it is possible to separate the objects by lifting or twisting one off the other, then, with certain exceptions, if the objects are modified by a sufficiently small amount with with sufficiently small change in boundary direction, then the connected components of the free space remain essentially unchanged. Connected components do not appear, disappear, merge, or split. Figures 6 and 7 illustrates two of the "exceptions" to this general rule; our theorem below imposes conditions that rule out all of these.

The definitions that we will need for this result are rather complex and the proof rather long, so let me begin by giving a rough sketch of the structure of the proof. We wish to show that, in situations like figure 4, if the changes to the shape and surface normal are kept small, then alterations in the configuration space such as appear in figure 4 can be avoided. Specifically, what happens in side II of figure 4 is that there is a modification of the shapes in I and there are two configurations both close to the configuration in I — one hooked and the other (not shown) unhooked — which are not connected by any feasible path. This is related to the fact that though in scenario I object B can move away from object A over a range of upward directions, in scenario II it cannot move away from A cannot move upward. We say that the upward motion *strongly separates* B from A in scenario I. The main steps of our proof are then as follows:

- If motion M strongly separates object B from object A, then B can be moved along M a finite distance, and escapes contact with A immediately after the start of motion. (Lemma 4.17)
- Suppose that there is some motion M strongly separates B from A in scenario $\langle D, C\rangle$. We now

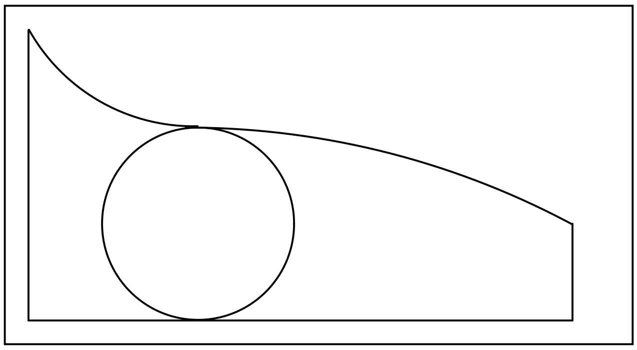


A

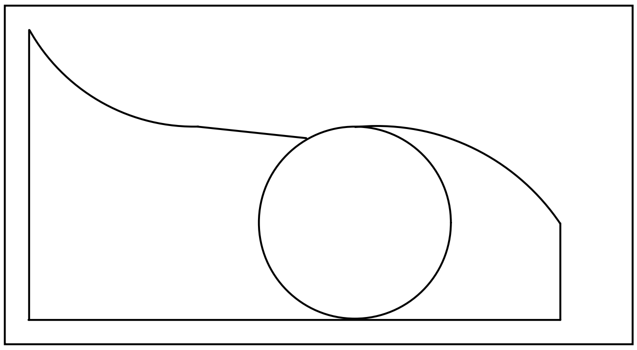
B

The pincer is “stuck” in (B) if the sides of the notch are steeper than the circle through one tooth centered at the other tooth. With a sufficiently small rotation, this can be accomplished with an arbitrarily small and shallow notch.

Figure 6: Change in connectedness under small change in tangent: I



A



B

In (A) the ball is free to move to the left.
However, a “trapped” position can be created by carving out a “notch” as in (B),
of arbitrarily small depth and change in angle.

Figure 7: Change in connectedness under small change in tangent: II

slightly change the shape of D to D_0, enforcing the rule that the surface normal may change only slightly, and we find two configurations C_1 and C_2 that are close to C and are feasible over D_0. Then there is a path from C_1 to C_2 that is feasible over D_0. (Lemma 4.22)

The final theorem states that uniform bounds for maximal change in shape can be found such that the connected components of free configuration space remain essentially unchanged.

Down to business. For two points $\mathbf{p} \neq \mathbf{q}$, we write "dir($\mathbf{p}, \mathbf{q}$)" to mean the the direction from $\mathbf{p}$ to $\mathbf{q}$: $\mathrm{dir}(\mathbf{p}, \mathbf{q}) = (\mathbf{q} - \mathbf{p})/ \mid \mathbf{q} - \mathbf{p} \mid$. For a unit vector $\hat{u}$ and angle θ, we will write $\hat{u} + \theta$ to mean the unit vector $\hat{v}$ at positive angle θ from $\hat{u}$, slightly abusing notation. We will write $\angle \vec{u}\vec{v}$ to mean the unsigned angle between vectors $\vec{u}$ and $\vec{v}$ ($= \cos^{-1}(\vec{u} \cdot \vec{v}/ \mid \vec{u} \mid\mid \vec{v} \mid)$). We will write $(\hat{u}, \hat{v})$ to mean the interval of all unit vectors $\hat{w}$ strictly between $\hat{u}$ and $\hat{v}$ in the positive direction. Thus, for $\theta \in (0, \pi)$, $(\hat{u} - \theta, \hat{u} + \theta)$ is the interval of all unit vectors $\hat{w}$ such that $\angle \hat{w}\hat{u} < \theta$. Also, we write $[\hat{u}, \hat{v}]$ for the closed interval from $\hat{u}$ to $\hat{v} = (\hat{u}, \hat{v}) \cup \{\hat{u}, \hat{v}\}$. We denote the boundary of region $\mathbf{R}$ as "Bd($\mathbf{R}$)".

We will assume that the shapes of our objects are *solid polycurves*.

Definition 4.1: A *polycurve* is a simple closed curve $\mathbf{C}$ such that

- $\mathbf{C}$ is smooth at all but finitely many points; and
- For every point $\mathbf{p} \in \mathbf{C}$, the tangent to $\mathbf{C}$ at $\mathbf{q}$ converges to a limit as $\mathbf{q}$ approaches $\mathbf{p}$ from each direction.

A *solid* polycurve is an object-like region whose boundary is the union of finitely many disjoint polycurves.

We now define what it means to approximate a shape both in distance and in the surface normal.

Definition 4.2: Let $\mathbf{A}$ and $\mathbf{B}$ be polycurves, let $\epsilon > 0$ be a length and let $\phi > 0$ be an angle. $\mathbf{B}$ *approximates* $\mathbf{A}$ *in tangent*, with parameters ϵ, α, if there exists a homeomorphism Γ from $\mathbf{A}$ to $\mathbf{B}$ such that,

i. For any $\mathbf{a} \in \mathbf{A}$, $\mathrm{d}(\mathbf{a}, \Gamma(\mathbf{a})) < \epsilon$; and

ii. For any $\mathbf{a} \in \mathrm{Bd}(\mathbf{A})$, let $\hat{e}, \hat{f}$ be the tangents to $\mathbf{A}$ at $\mathbf{a}$ in the positive and negative directions, and let $\hat{g}, \hat{h}$ be the tangents to $\mathbf{B}$ at $\Gamma(\mathbf{a})$ in the positive and negative directions. Then $\angle \hat{e}\hat{g} < \alpha$ and $\angle \hat{f}\hat{h} < \alpha$.

Lemma 4.1: If $\mathbf{B}$ approximates $\mathbf{A}$ in tangent (α, β) and $\mathbf{C}$ approximates $\mathbf{B}$ in tangent (μ, θ) then $\mathbf{C}$ approximates $\mathbf{A}$ in tangent $(\mu + \alpha, \theta + \beta)$.

Proof: Immediate from the definition by composing the two homeomorphisms.

Lemma 4.2: If region $\mathbf{S}$ approximates region $\mathbf{R}$ in tangent (δ, α), then the dual-Hausdorff distance from $\mathbf{S}$ to $\mathbf{R}$ is less than 2δ.

Proof: It is immediate that the Hausdorff distance from $\mathbf{S}$ to $\mathbf{R}$ is less than δ. To show the result for the complement-Hausdorff distance, let Γ be a homeomorphism between $\mathbf{R}$ and $\mathbf{S}$ that moves every point by less than δ. Let $\mathbf{p}$ be any point in the complement of $\mathbf{S}$. If $\mathbf{p}$ is also outside $\mathbf{R}$, then the distance from $\mathbf{p}$ to $\mathbf{R}^c$ is 0. If $\mathbf{p} \in \mathbf{R}$, then, by definition, $\Gamma(\mathbf{p}) \in \mathbf{S}$ and $\mathrm{d}(\mathbf{p}, \Gamma(\mathbf{p})) < \delta$. Thus, the line segment from $\mathbf{p}$ to $\Gamma(\mathbf{p})$ must intersect the boundary of $\mathbf{S}$, so there is a point $\mathbf{b}$ on the boundary of $\mathbf{S}$ within δ of $\mathbf{p}$. Hence, the distance from $\Gamma^{-1}(\mathbf{b})$ to $\mathbf{p}$ is less than 2δ. Since $\Gamma^{-1}(\mathbf{b})$ is a boundary point of $\mathbf{R}$, $\mathbf{p}$ is within 2δ of a point in the complement of $\mathbf{R}$. Thus every point in the complement of $\mathbf{S}$ is within 2δ of a point in the complement of $\mathbf{R}$. Switching the roles of Γ and Γ^{-1} gives the same result in the opposite direction. ∎

We begin by recalling a useful lemma on compact sets:

Lemma 4.3: Let U be a compact space. For $u \in U$ and $x_1 \dots x_k \in \Re$, let $\Phi(u, x_1 \dots x_k)$ be a property of u such that

a. If $0 < y_i < x_i$ for $i = 1 \dots k$, and $\Phi(u, x_1 \dots x_k)$ then $\Phi(u, y_1 \dots y_k)$.

b. For any $u \in U$ there exists a neighborhood V of u and values $z_1 > 0 \dots z_k > 0$ such that, for all $v \in V$, $\Phi(v, z_1 \dots z_k)$.

Then there exist $x_1 > 0 \dots x_k > 0$ such that for all $u \in U$, $\Phi(u, x_1 \dots x_k)$.

Proof: Let $U_n = \{u \mid \Phi(u, 1/n \dots 1/n)\}$. Then the collection of $U_n, n = 1 \dots$ is a covering of U by open sets. Since U is compact, this has a finite subcover. But, for $n > m$, $U_n \supset U_m$. Thus the last U_n in the finite subcover contains all the rest and is so equal to U. Thus $\Phi(u, 1/n \dots 1/n)$ for all $u \in U$. ▮

Definition 4.3: Let $\mathbf{A}$ and $\mathbf{B}$ be two non-overlapping regions, and let $\mathbf{p}$ be a point in $\mathbf{A} \cap \mathbf{B}$. Let U be an open interval in the unit circle. U *strongly separates* $\mathbf{B}$ from $\mathbf{A}$ at $\mathbf{p}$ if, for every vector $\hat{u} \in U$,

1. for every neighborhood $\mathbf{V}$ of $\mathbf{p}$ there exist points $\mathbf{a} \in \mathbf{A} \cap \mathbf{V}$ and $\mathbf{b} \in \mathbf{B} \cap \mathbf{V}$, $\mathbf{a} \neq \mathbf{b}$, such that $\hat{u}$ is the direction from $\mathbf{a}$ to $\mathbf{b}$; and
2. there exists a neighborhood $\mathbf{V}$ of $\mathbf{p}$ in which there do not exist points $\mathbf{a} \in \mathbf{A} \cap \mathbf{V}$ and $\mathbf{b} \in \mathbf{B} \cap \mathbf{V}$, $\mathbf{a} \neq \mathbf{b}$, such that $\hat{u}$ is the direction from $\mathbf{b}$ to $\mathbf{a}$.

A vector $\hat{u}$ strongly separates $\mathbf{B}$ from $\mathbf{A}$ if some neighborhood U of $\hat{u}$ strongly separates $\mathbf{B}$ from $\mathbf{A}$. We denote the set of vectors that strongly separate $\mathbf{B}$ from $\mathbf{A}$ at $\mathbf{p}$ as "sep($\mathbf{B}, \mathbf{A}, \mathbf{p}$)".

Lemma 4.4: For any regions $\mathbf{B}, \mathbf{A}$ and point $\mathbf{p} \in \mathbf{A} \cap \mathbf{B}$, sep($\mathbf{B}, \mathbf{A}, \mathbf{p}$) is an open set in the unit circle.

Proof: Immediate from the definition. ▮.

It is straightforward to characterize the space of directions that strongly separate polycurves $\mathbf{B}$ from $\mathbf{A}$ at any contact point $\mathbf{p}$ in terms of the tangents to $\mathbf{A}$ and $\mathbf{B}$ at $\mathbf{p}$: (Figure 8) Let $\mathbf{A}$ and $\mathbf{B}$ be two non-overlapping polycurves, and let $\mathbf{p}$ be a point in $\mathbf{A} \cap \mathbf{B}$. Let $\hat{e}$ be the unit vector tangent to $\mathbf{A}$ at $\mathbf{p}$ in the positive (counter-clockwise) direction; let $\hat{f}$ be the unit vector tangent to $\mathbf{A}$ at $\mathbf{p}$ in the negative direction; let $\hat{g}$ be the unit vector tangent to $\mathbf{B}$ at $\mathbf{p}$ in the positive direction; and let $\hat{h}$ be the unit vector tangent to $\mathbf{B}$ at $\mathbf{p}$ in the negative direction. Thus $\hat{e} = -\hat{f}$ if $\mathbf{A}$ is smooth at $\mathbf{p}$. A vector $\hat{u}$ sited at $\mathbf{p}$ points into $\mathbf{A}$ if $\hat{u} \in (\hat{e}, \hat{f})$ and points out of $\mathbf{A}$ if $\hat{u} \in (\hat{f}, \hat{e})$; and correspondingly for $\mathbf{B}$. Note that $\hat{e}, \hat{f}, \hat{g}, \hat{h}$ are in non-strict positive cyclic order.

Definition 4.4: (Figure 8). The function sep1$(\hat{e}, \hat{f}, \hat{g}, \hat{h})$ is defined as follows:

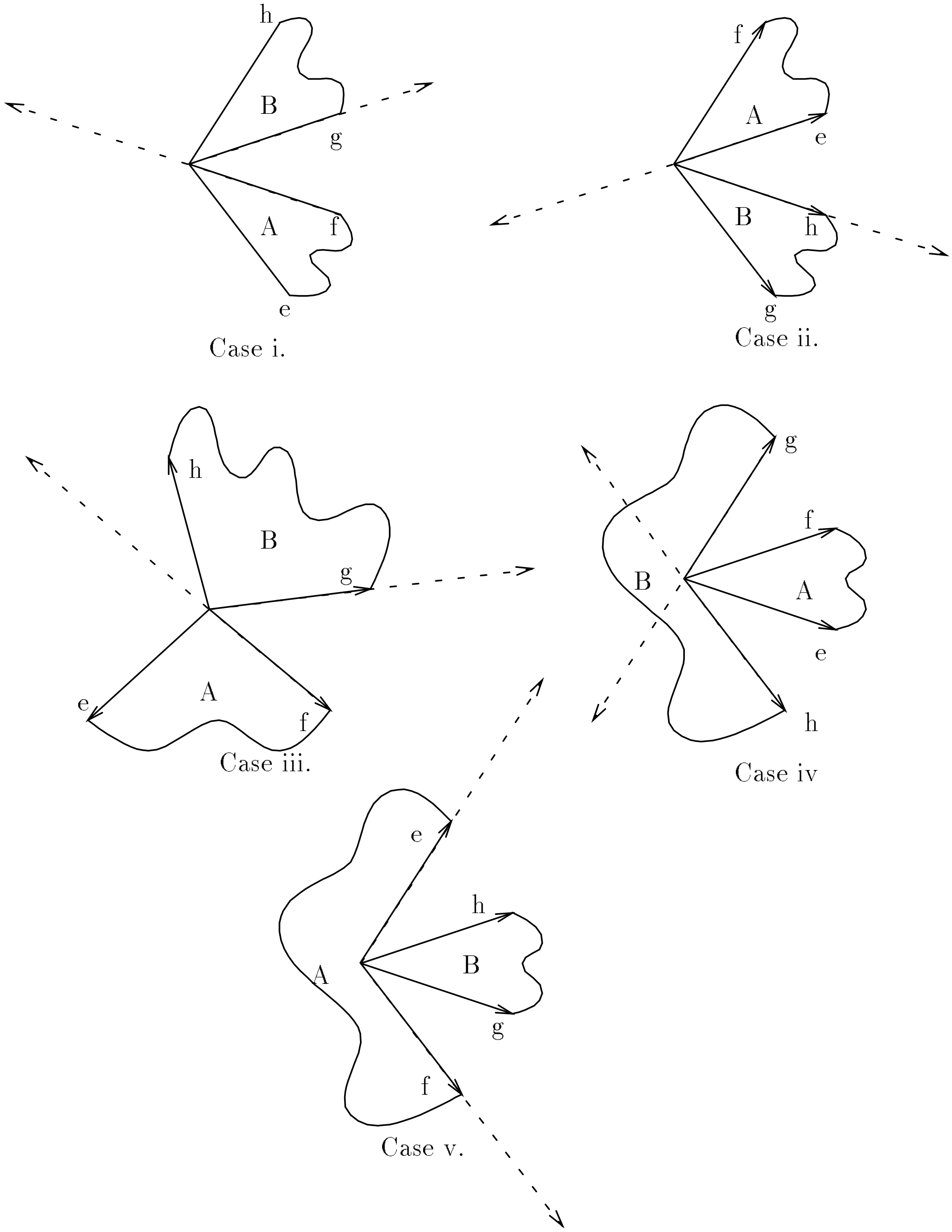
i. If the angle from $\hat{e}$ to $\hat{h}$ is less than or equal to π, then sep1$(\hat{e}, \hat{f}, \hat{g}, \hat{h}) = (\hat{g}, -\hat{f})$.

ii. If the angle from $\hat{g}$ to $\hat{f}$ is less than or equal to π, then sep1$(\hat{e}, \hat{f}, \hat{g}, \hat{h}) = (-\hat{e}, \hat{h})$.

iii. If the angles from $\hat{e}$ to $\hat{f}$, from $\hat{f}$ to $\hat{g}$, from $\hat{g}$ to $\hat{h}$ and from $\hat{h}$ to $\hat{e}$ are all less than or equal to π, then sep1$(\hat{e}, \hat{f}, \hat{g}, \hat{h}) = (\min(-\hat{e}, \hat{g}), \max(-\hat{f}, \hat{h}))$, where min and max are in the sense of the smaller positive rotation between the two.

iv. If the angle from $\hat{h}$ to $\hat{g}$ is less than or equal to π, then sep1$(\hat{e}, \hat{f}, \hat{g}, \hat{h}) = (-\hat{h}, -\hat{g})$.

v. If the angle from $\hat{f}$ to $\hat{e}$ is less than or equal to π, then sep1$(\hat{e}, \hat{f}, \hat{g}, \hat{h}) = (\hat{f}, \hat{e})$.



Dashed lines indicate the range of directions separating B from A.

Figure 8: Separating directions

It is easily verified that in figures where two or more of these vectors are angle π apart, so that more than one rule applies, the different rules give the same results.

Lemma 4.5: For any polycurves $\mathbf{A}, \mathbf{B}$ and point $\mathbf{p} \in \mathbf{A} \cap \mathbf{B}$, $\text{sep}(\mathbf{B}, \mathbf{A}, \mathbf{p}) = \text{sep1}(\hat{e}, \hat{f}, \hat{g}, \hat{h})$ where $\hat{e}$, $\hat{f}$, $\hat{g}$, $\hat{h}$ are the tangents to $\mathbf{A}$ and $\mathbf{B}$ at $\mathbf{p}$, as defined above.

Proof: Straightforward from the geometry. ∎.

Lemma 4.6:

A. $\text{sep1}(\hat{e}, \hat{f}, \hat{g}, \hat{h}) = -\text{sep1}(\hat{g}, \hat{h}, \hat{e}, \hat{f})$

B. If $\hat{g} \in [-\hat{e}, \hat{e}]$ and $\hat{h} \in [-\hat{e}, \hat{e}]$, then $\text{sep}(\hat{e}, -\hat{e}, \hat{g}, \hat{h}) = (-\hat{e}, \hat{e})$.

C. If $(\hat{e}, \hat{f}) \subset [\hat{q}, -\hat{q}]$ then $\text{sep1}(\hat{e}, \hat{f}, \hat{g}, \hat{h}) \subset \text{sep1}(\hat{q}, -\hat{q}, \hat{r}, \hat{s})$, where $\hat{r}$ and $\hat{s}$ are any two vectors in $(-\hat{q}, \hat{q})$.

D. If $(\hat{e}, \hat{f}) \supset [\hat{q}, -\hat{q}]$ then $\text{sep1}(\hat{e}, \hat{f}, \hat{g}, \hat{h}) \subset \text{sep1}(\hat{q}, -\hat{q}, \hat{r}, \hat{s})$, where $\hat{r}$ and $\hat{s}$ are any two vectors in $(-\hat{q}, \hat{q})$.

Proof: Immediate from definition 4.4. ∎

Lemma 4.7: Let $\hat{e}$, $\hat{f}$, $\hat{g}$, $\hat{h}$ and $\hat{e}'$, $\hat{f}'$, $\hat{g}'$, $\hat{h}'$, be two quadruples of vectors, both in positive order, such that $\angle \hat{e}'\hat{e} < \theta$, $\angle \hat{f}'\hat{f} < \theta$, $\angle \hat{g}'\hat{g} < \theta$, and $\angle \hat{h}'\hat{h} < \theta$. Let $\hat{u}$ be a vector such that the interval $(\hat{u} - \theta, \hat{u} + \theta) \subset \text{sep1}(\hat{e}, \hat{f}, \hat{g}, \hat{h})$. Then $\hat{u} \in \text{sep1}(\hat{e}', \hat{f}', \hat{g}', \hat{h}')$.

Proof: Immediate from lemma 4.6. ∎

Definition 4.5: A *motion* is either (a) a unit vector $\hat{u}$, representing translation in the $\hat{u}$ direction; or (b) a pair $\langle \mathbf{o}, S \rangle$ of a point $\mathbf{o}$ and a sign $S = \pm 1$ representing counter-clockwise (if $S = 1$) or clockwise (if $S = -1$) rotation about the point $\mathbf{o}$.

Definition 4.6: The *bounded distance* from point $\mathbf{p}$ to motion M, denoted $\bar{d}(\mathbf{p}, M)$ is equal to $\min(\text{d}(\mathbf{p}, \mathbf{o}), 1)$ if $M = \langle \mathbf{o}, S \rangle$ and equal to 1 if $M = \hat{v}$. If $\mathbf{R}$ is a region then $\bar{d}(\mathbf{R}, M) = \inf_{\mathbf{p} \in \mathbf{R}} \bar{d}(\mathbf{p}, M)$.

Definition 4.7: The *flow* from point $\mathbf{p}$ to motion M, denoted "$\text{flow}(\mathbf{p}, M)$" is the direction of motion of point $\mathbf{p}$ under motion M. It is defined as follows:

- If $M = \hat{v}$ then $\text{flow}(\mathbf{p}, M) = \hat{v}$.
- If $M = \langle \mathbf{o}, S \rangle$ and $\mathbf{p} = \mathbf{o}$, then $\text{flow}(\mathbf{p}, M) = \vec{0}$.
- If $M = \langle \mathbf{o}, S \rangle$ and $\mathbf{p} \neq \mathbf{o}$, then $\text{flow}(\mathbf{p}, M) = \text{dir}(\mathbf{o}, \mathbf{p}) + S\pi/2$

Lemma 4.8: For any point $\mathbf{p}, \mathbf{q}$ and motion M, if neither $\text{flow}(\mathbf{p}, M)$ nor $\text{flow}(\mathbf{q}, M)$ is equal to $\vec{0}$, let θ be the angle between $\text{flow}(\mathbf{p}, M)$ and $\text{flow}(\mathbf{q}, M)$. Then $\sin(\theta) \leq \text{d}(\mathbf{p}, \mathbf{q}) \ / \ \bar{d}(\mathbf{p}, M)$.

Proof: Immediate. ∎

Definition 4.8: Let $\mathbf{A}$ and $\mathbf{B}$ be two non-overlapping polycurves. Motion M strongly separates $\mathbf{B}$ from $\mathbf{A}$ if, for every point $\mathbf{p} \in \mathbf{A} \cap \mathbf{B}$, $\text{flow}(\mathbf{p}, M) \in \text{sep}(\mathbf{B}, \mathbf{A}, \mathbf{p})$. Note that, if M strongly separates $\mathbf{B}$ from $\mathbf{A}$ then, $\bar{d}(\mathbf{A} \cap \mathbf{B}, M) > 0$.

For the next several lemmas, we will assume that $\mathbf{A}$ and $\mathbf{B}$ are non-overlapping solid polycurves; and that M is a motion strongly separating $\mathbf{B}$ from $\mathbf{A}$. All of lemmas 4.9 - 4.16 below are stated relative to an arbitrary fixed choice of $\mathbf{A}$, $\mathbf{B}$, and M.

Lemma 4.9: For any point $\mathbf{p} \in \mathbf{B}$, there exist $\epsilon > 0, \alpha > 0$ such that, for any $\mathbf{a} \in \mathbf{A}$, $\mathbf{b} \in \mathbf{B}$ if both $\mathbf{a}$ and $\mathbf{b}$ are within ϵ of $\mathbf{p}$, then the angle between flow$(\mathbf{p}, M)$ and dir$(\mathbf{b}, \mathbf{a})$ is at least α.

Proof: If $\mathbf{p}$ is not in $\mathbf{A}$, then let $\epsilon <$ d$(\mathbf{p}, \mathbf{A})$. The condition is then satisfied vacuously, as there is no such $\mathbf{a}$. If $\mathbf{p} \in \mathbf{A} \cap \mathbf{B}$, then the condition is immediate from definitions 4.3 and 4.7.

Lemma 4.10: There exist $\epsilon > 0, \alpha > 0$ such that, for any points $\mathbf{p}, \mathbf{b} \in \mathbf{B}$ and $\mathbf{a} \in \mathbf{A}$, if both $\mathbf{a}$ and $\mathbf{b}$ are within ϵ of $\mathbf{p}$, then the angle between flow$(\mathbf{p}, M)$ and dir$(\mathbf{b}, \mathbf{a})$ is at least α. (This is the same as lemma 4.9, except here α and ϵ are quantified with larger scope than $\mathbf{p}$.)

Proof: We apply lemma 4.3, taking the domain U to be $\mathbf{B}$ and the property $\Phi(\mathbf{p}, \epsilon, \alpha)$ to be the relation, "For any point $\mathbf{b} \in \mathbf{B}$ and $\mathbf{a} \in \mathbf{A}$, if both $\mathbf{a}$ and $\mathbf{b}$ are within ϵ of $\mathbf{p}$, then the angle between flow$(\mathbf{p}, M)$ and dir$(\mathbf{b}, \mathbf{a})$ is at least α." Property (a) of lemma 4.3 is immediate. Property (b) requires establishing that, for any point $\mathbf{p} \in \mathbf{B}$ there exist $\delta > 0, \mu > 0, \theta > 0$, such that, if $\mathbf{u} \in \mathbf{B}$ and d$(\mathbf{u}, \mathbf{p}) < \delta$ then $\Phi(\mathbf{u}, \mu, \theta)$. We do this as follows: For any point $\mathbf{p} \in \mathbf{B}$, use lemma 4.9 to find $\epsilon > 0, \alpha > 0$ such that $\Phi(\mathbf{p}, \epsilon, \alpha)$. Let $\theta = \alpha/2$ and let $\delta = \mu = \min(\epsilon/2, \bar{d}(\mathbf{p}, M)\sin(\theta))$. Now choose $\mathbf{u}$ within δ of $\mathbf{p}$ and $\mathbf{a} \in \mathbf{A}$, $\mathbf{b} \in \mathbf{B}$ such that both $\mathbf{a}$ and $\mathbf{b}$ are within μ of $\mathbf{u}$. Now since $\mathbf{a}$ and $\mathbf{b}$ are within ϵ of $\mathbf{p}$, dir$(\mathbf{b}, \mathbf{a})$ be at least α from flow$(\mathbf{p}, M)$. By lemma 4.8, flow$(\mathbf{u}, M)$ is within $\alpha/2$ of flow$(\mathbf{p}, M)$, so the angle between dir$(\mathbf{b}, \mathbf{a})$ and flow$(\mathbf{u}, M)$ must be at least $\alpha/2$. Therefore we can apply lemma 4.3, and conclude that α and ϵ can be chosen uniformly over all $\mathbf{p} \in \mathbf{B}$. ∎

Lemma 4.11: Let $\mathbf{p} \in$Bd$(\mathbf{A})$, let $\hat{e}$ be the tangent to $\mathbf{A}$ at $\mathbf{p}$ in the counterclockwise direction, and let $\hat{f}$ be the tangent to $\mathbf{A}$ at $\mathbf{p}$ in the clockwise direction. Then for any $\epsilon > 0$ there exists a $\delta > 0$ such that for any boundary point $\mathbf{q}$ within ϵ of $\mathbf{p}$, if $\mathbf{q}$ is counterclockwise from $\mathbf{p}$ then the positive tangent to $\mathbf{A}$ at $\mathbf{p}$ is within ϵ of $\hat{e}$ and the negative tangent to $\mathbf{A}$ at $\mathbf{p}$ is within ϵ of $-\hat{e}$; and if $\mathbf{q}$ is clockwise from $\mathbf{p}$ then the positive tangent to $\mathbf{A}$ at $\mathbf{p}$ is within ϵ of $-\hat{f}$ and the negative tangent to $\mathbf{A}$ at $\mathbf{p}$ is within ϵ of $\hat{f}$.

Proof: Immediate from definition 4.1 of a polycurve. ∎

Definition 4.9: Let $\mathbf{C}, \mathbf{D}$ be two solid polycurves. Let $\mathbf{p} \in$Bd$(\mathbf{C})$ and $\mathbf{q} \in$Bd$(\mathbf{D})$. Let $\hat{e}$ and $\hat{f}$ be the tangents to $\mathbf{C}$ at $\mathbf{p}$ in the positive and negative directions; and let $\hat{g}$ and $\hat{h}$ be the tangents to $\mathbf{D}$ at $\mathbf{q}$ in the positive and negative directions. Let $\epsilon > 0$. We say that "the tangents to $\mathbf{D}$ at $\mathbf{q}$ are within ϵ of the tangents to $\mathbf{C}$ at $\mathbf{p}$," if one of the following holds:

i. $\hat{g}$ is within ϵ of $\hat{e}$ and $\hat{h}$ is within ϵ of $\hat{f}$; or

ii. $\hat{g}$ is within ϵ of $\hat{e}$ and $\hat{h}$ is within ϵ of $-\hat{e}$; or

iii. $\hat{g}$ is within ϵ of $-\hat{f}$ and $\hat{h}$ is within ϵ of $\hat{f}$.

(Note this is not a symmetric relation between $\mathbf{C}, \mathbf{p}$ and $\mathbf{D}, \mathbf{q}$.)

Lemma 4.12: Let $\mathbf{A}$ be a solid polycurve and let $\mathbf{p} \in$Bd$(\mathbf{A})$. Then for any $\epsilon > 0$ there exists a $\delta > 0$ such that, if $\mathbf{q} \in$Bd$(\mathbf{A})$ and d$(\mathbf{p}, \mathbf{q}) < \delta$ then the tangents to $\mathbf{A}$ at $\mathbf{q}$ are within ϵ of the tangents to $\mathbf{A}$ at $\mathbf{p}$.

Proof: This is just a restatement of lemma 4.11, adding the case that $\mathbf{q} = \mathbf{p}$. ∎

Definition 4.10: The *angle from blockage* of unit vectors $\hat{e}, \hat{f}, \hat{g}, \hat{h}, \hat{u}$ is the minimum ϵ for which there exist $\hat{e}', \hat{f}', \hat{g}', \hat{h}', \hat{u}'$ such that $\angle\hat{e}\hat{e}' \leq \epsilon$, $\angle\hat{f}\hat{f}' \leq \epsilon$, $\angle\hat{g}\hat{g}' \leq \epsilon$, $\angle\hat{h}\hat{h}' \leq \epsilon$, $\angle\hat{u}\hat{u}' \leq \epsilon$; $\hat{e}', \hat{f}', \hat{g}', \hat{h}'$ are in (non-strict) positive cyclic order; and $\hat{u}' \not\in$sep1$(\hat{e}', \hat{f}', \hat{g}', \hat{h}')$.

That is, the angle from blockage is the minimal change in all these vectors so that the tangent vectors can correspond to those of two non-overlapping objects in contact, and so that $\hat{u}$ does not strongly separate the two objects.

Definition 4.11: Let $\mathbf{p} \in$Bd$(\mathbf{A})$ and $\mathbf{q} \in$Bd$(\mathbf{B})$. Let $\hat{e}, \hat{f}, \hat{g}, \hat{h}$ be the positive and negative tangents to $\mathbf{A}$ at $\mathbf{p}$ and to $\mathbf{B}$ at $\mathbf{q}$. Let $\hat{u}$ be a vector. The function "free_ang$(\mathbf{A}, \mathbf{p}, \mathbf{B}, \mathbf{q}, \hat{u})$" is the angle from blockage of $\hat{e}, \hat{f}, \hat{g}, \hat{h}, \hat{u}$.

Lemma 4.13: Let $\mathbf{p} \in \mathbf{A} \cap \mathbf{B}$ and let $\hat{u} \in$sep$(\mathbf{B}, \mathbf{A}, \mathbf{p})$. Then there exist $\delta > 0$ and $\theta > 0$ such that, for all $\mathbf{q} \in$Bd$(\mathbf{A})$ and $\mathbf{r} \in$Bd$(\mathbf{B})$, if d$(\mathbf{p}, \mathbf{q}) < \delta$ and d$(\mathbf{p}, \mathbf{r}) < \delta$, then free_ang$(\mathbf{A}, \mathbf{q}, \mathbf{B}, \mathbf{r}, \hat{u}) \geq \theta$.

Proof: Let $\hat{e}, \hat{f}, \hat{g}, \hat{h}$ be the positive and negative tangents to $\mathbf{A}$ and $\mathbf{B}$ at $\mathbf{p}$. If $\hat{e}$ is between $\hat{g}$ and $-\hat{g}$, let $\gamma_e = \pi - \angle \hat{e}\hat{g}$ else $\gamma_e = \pi$. If $\hat{f}$ is between $-\hat{h}$ and $\hat{h}$, let $\gamma_f = \pi - \angle \hat{f}\hat{h}$ else $\gamma_f = \pi$. If $\hat{g}$ is between $\hat{e}$ and $-\hat{e}$, let $\gamma_g = \pi - \angle \hat{g}\hat{e}$ else $\gamma_g = \pi$. If $\hat{h}$ is between $-\hat{f}$ and $\hat{f}$, let $\gamma_h = \pi - \angle \hat{h}\hat{f}$ else $\gamma_h = \pi$. Using lemma 4.4, choose γ_u so that $(\hat{u} - \gamma_u, \hat{u} + \gamma_u) \in$sep$(\mathbf{B}, \mathbf{A}, \mathbf{p})$. Now, let $\theta = (1/3)\min(\gamma_e, \gamma_f, \gamma_g, \gamma_h, \gamma_u)$.

Using lemma 4.12, find $\mu > 0$ such that for all $\mathbf{q} \in$Bd$(\mathbf{A})$ within μ of $\mathbf{p}$, the tangents to $\mathbf{A}$ at $\mathbf{q}$ are within θ of the tangents to $\mathbf{A}$ at $\mathbf{p}$, and such that for all $\mathbf{r} \in$Bd$(\mathbf{B})$ within μ of $\mathbf{p}$, the tangents to $\mathbf{B}$ at $\mathbf{r}$ are within θ of the tangents to $\mathbf{B}$ at $\mathbf{p}$. Let $\delta = \min(\mu, \bar{d}(\mathbf{p}, M)\sin(\theta))$. Now, choose $\mathbf{q} \in \mathbf{A}$, $\mathbf{r} \in \mathbf{B}$ such that $\mathbf{q}$ and $\mathbf{r}$ are both within δ of $\mathbf{p}$. Let $\hat{e}_1, \hat{f}_1, \hat{g}_1, \hat{h}_1$ be the positive and negative tangents to $\mathbf{A}$ at $\mathbf{q}$ and to $\mathbf{B}$ at $\mathbf{r}$. It is easily verified that, if $\hat{e}_1, \hat{f}_1, \hat{g}_1, \hat{h}_1$ are not in positive order then they are at least θ from being in positive order, while if they are in positive order, then $\hat{u}$ is at least θ from the boundaries of sep1$(\hat{e}, \hat{f}, \hat{g}, \hat{h})$. ∎

Lemma 4.14: There exist $\delta > 0$ and $\theta > 0$ such that, for all $\mathbf{p} \in \mathbf{A} \cap \mathbf{B}$, $\mathbf{q} \in$Bd$(\mathbf{A})$, $\mathbf{r} \in$Bd$(\mathbf{B})$, if $\hat{u}$ =flow$(\mathbf{p}, M)$, d$(\mathbf{p}, \mathbf{q}) < \delta$, and d$(\mathbf{p}, \mathbf{r}) < \delta$, then free_ang$(\mathbf{A}, \mathbf{q}, \mathbf{B}, \mathbf{r}, \hat{u}) \geq \theta$.

Proof: We apply lemma 4.3, taking the domain U to be $\mathbf{A} \cap \mathbf{B}$ and the property $\Phi(\mathbf{p}, \delta, \theta)$ to be the property, "For any $\mathbf{q} \in$Bd$(\mathbf{A})$, $\mathbf{r} \in$Bd$(\mathbf{B})$, if $\hat{u}$ =flow$(\mathbf{p}, M)$, d$(\mathbf{p}, \mathbf{q}) < \delta$, and d$(\mathbf{p}, \mathbf{r}) < \delta$, then free_ang$(\mathbf{A}, \mathbf{q}, \mathbf{B}, \mathbf{r}, \hat{u}) \geq \theta$." Property (a) of lemma 4.3 is immediate. Property (b) requires establishing that, for any point $\mathbf{p} \in \mathbf{A} \cap \mathbf{B}$ there exist a $\mu > 0, \beta > 0, \alpha > 0$, such that, for $\mathbf{x} \in \mathbf{A} \cap \mathbf{B}$, if d$(\mathbf{x}, \mathbf{p}) < \mu$, then $\Phi(\mathbf{x}, \beta, \alpha)$. We proceed as follows: For any such $\mathbf{p}$, find δ and θ satisfying lemma 4.13. Let $\alpha = \theta/2$ and let $\beta = \delta/2$ and let $\mu = \min(\delta/2, \bar{d}(\mathbf{p}, M)\sin(\alpha))$. Let $\mathbf{x}$ be any point in $\mathbf{A} \cap \mathbf{B}$ within μ of $\mathbf{p}$ and let $\mathbf{q}, \mathbf{r}$ be points in $\mathbf{A}$ and $\mathbf{B}$ within β of $\mathbf{x}$. Then certainly d$(\mathbf{p}, \mathbf{r}) < \delta$, so free_ang$(\mathbf{A}, \mathbf{q}, \mathbf{B}, \mathbf{r}$,flow$(\mathbf{p},M)) \geq \theta$. Moreover, flow$(\mathbf{x},M)$ is within α of flow$(\mathbf{p},M)$, It then follows immediately, from definitions 4.10 and 4.11, that free_ang$(\mathbf{A}, \mathbf{q}, \mathbf{B}, \mathbf{r}$,flow$(\mathbf{x},M)) \geq \theta - \alpha = \alpha$. ∎

Lemma 4.15: For any $\epsilon > 0$ there exists a $\gamma > 0$ such that if $\mathbf{a} \in \mathbf{A}$, $\mathbf{b} \in \mathbf{B}$ and d$(\mathbf{a}, \mathbf{b}) < \gamma$ then there exists a point $\mathbf{p} \in \mathbf{A} \cap \mathbf{B}$ such that d$(\mathbf{p}, \mathbf{a}) < \epsilon$ and d$(\mathbf{p}, \mathbf{b}) < \epsilon$.

Proof: Immediate from lemma 2.5, with the compact domain $\mathbf{U}$ being the cross product $\mathbf{A} \times \mathbf{B}$, the function $f(\langle \mathbf{a}, \mathbf{b} \rangle) = $ d$(\mathbf{a}, \mathbf{b})$, and the metric $\mu(\langle \mathbf{a}, \mathbf{b} \rangle, \langle \mathbf{c}, \mathbf{d} \rangle) = \max(d(\mathbf{a}, \mathbf{c})$, d$(\mathbf{b}, \mathbf{d}))$.

Lemma 4.16: There exist $\beta > 0, \psi > 0$ such that, if $\mathbf{X}, \mathbf{Y}$ are non-overlapping polycurves that respectively approximate $\mathbf{A}, \mathbf{B}$ in tangent (β, ψ), then $\mathbf{M}$ strongly separates $\mathbf{Y}$ from $\mathbf{X}$.

Proof: Find δ and θ to satisfy lemma 4.14. Let $\psi < \theta$, and let $D = (1/2)\bar{d}(\mathbf{A} \cap \mathbf{B}, M)\sin(\psi)$. Let $\epsilon = \min(\delta/2, D)$. Find a value of γ to satisfy lemma 4.15 for this value of ϵ, and let $\beta = \min(\gamma/2, D)$.

Now let $\mathbf{X}, \mathbf{Y}$ be polycurves that respectively approximate $\mathbf{A}, \mathbf{B}$ in tangent (β, ψ), with approximating homeomorphisms Γ and Δ. Let $\mathbf{p} \in \mathbf{X} \cap \mathbf{Y}$; let $\mathbf{a} = \Gamma^{-1}(\mathbf{p})$; and let $\mathbf{b} = \Delta^{-1}(\mathbf{p})$. Then d$(\mathbf{a}, \mathbf{b}) < 2\beta$, so there exists a point $\mathbf{q} \in \mathbf{A} \cap \mathbf{B}$ such that d$(\mathbf{q}, \mathbf{a}) < \epsilon$, d$(\mathbf{q}, \mathbf{b}) < \epsilon$. Therefore, free_ang$(\mathbf{A}, \mathbf{a}, \mathbf{B}, \mathbf{b}$,flow$(\mathbf{q},M)) \geq \theta$. Now, the tangents to $\mathbf{A}$ at $\mathbf{a}$ are within ψ of the tangents to $\mathbf{X}$ at $\mathbf{p}$ and the tangents to $\mathbf{B}$ are within ψ of the tangents to $\mathbf{Y}$ at $\mathbf{p}$. Also, since d$(\mathbf{p}, \mathbf{q}) < 2D$, flow$(\mathbf{p}, M)$ is within ψ of flow$(\mathbf{q}, M)$. Therefore, free_ang$(\mathbf{X}, \mathbf{p}, \mathbf{Y}, \mathbf{p}$,flow$(\mathbf{p}, M)) \geq \theta - \psi$. Since $\mathbf{X}$ and $\mathbf{Y}$ do not overlap, the tangents to $\mathbf{X}$ and $\mathbf{Y}$ at $\mathbf{p}$ are necessarily in proper cyclic order, so flow$(\mathbf{p}, M) \in$ sep$(\mathbf{Y}, \mathbf{X}, \mathbf{p})$. ∎

We now discharge the assumptions on $\mathbf{A}, \mathbf{B}$ and M, and return to the language of displays, configurations, and paths.

Definition 4.12: Let C be a configuration over the two objects A and B; let M be a motion; and let $\Delta > 0$. Then "arc(C, M, Δ)" is the path ϕ that starts in C, leaves A fixed and moves B along M until the distance moved is equal to Δ. Formally,

i. $\phi(0) = C$.

ii. For all $T \in [0,1]$, $\phi(T)[A] = C[A]$.

iii. For all $T \in (0,1)$ and for all $\mathbf{p}$, there exists $k > 0$ such that $d/dT(\phi(T)[B](\mathbf{p})) = k \cdot \text{flow}(\phi(T)[B](\mathbf{p}), M)$,

iv. Let $\mathbf{B} = CD[B]$. Then $\mathrm{p}^D(\phi(0), \phi(1)) = \Delta$.

v. For $T \in (0,1)$, $\mathrm{p}^D(\phi(0), \phi(T)) < \Delta$. (This, to prevent ϕ from traversing a full rotation before ending up at $\phi(1)$.)

Definition 4.13: We say that path ϕ *escapes contact* over D if, for all $T \in (0,1]$ $\langle D, \phi(T)\rangle$ is contact-free.

Lemma 4.17: Let $\langle D, C\rangle$ be a feasible scenario over objects A and B, where CD[A] and CD[B] are solid polycurves. Let M be a motion that strongly separates CD[B] from CD[A]. Then there exists a $\Delta > 0$ such that arc(C, M, Δ) escapes contact.

Proof: Find ϵ, α to satisfy lemma 4.10 for $\mathbf{A} = CD$[A], $\mathbf{B} = CD$[B]. If M is a translation, then choose $\Delta = \epsilon$. Otherwise, if $M = \langle \mathbf{o}, S\rangle$, let R be the maximal value of $\mathrm{d}(\mathbf{o}, \mathbf{p})$ for $\mathbf{p} \in \mathbf{B}$, and let $\Delta = \min(\epsilon, R\sin(\alpha))$. Let $\phi = \text{arc}(C, M, \Delta)$. Then for any $\mathbf{p} \in \mathbf{B}$, let $\mathbf{b} = C^{-1}(\mathbf{p}) \in D$[B]; and let $\mathbf{W}(T)$ be the curve $\phi(T)[B](\mathbf{b})$. We wish to show that $\mathbf{W}(T)$ does not intersect $\mathbf{A}$ for any $T \in (0,1]$. If M is a rotation with center $\mathbf{p}$, then $\mathbf{p} \notin \mathbf{A}$, and $\mathbf{W}(T) = \mathbf{p}$ for all T, so the result is immediate. If not, for any $T \in (0,1]$, $\mathrm{d}(\mathbf{W}(T), \mathbf{p}) \leq \Delta$. Let $\dot{\mathbf{W}}(T)$ be the tangent to $\mathbf{W}$ at time T; then, for all T, $\dot{\mathbf{W}}(T) = \text{flow}(\mathbf{W}(T), M)$ which is within α of flow$(\mathbf{p}, M)$. Then, for any $T \in (0,1]$, dir$(\mathbf{p}, \mathbf{W}(T))$ is within α of flow$(\mathbf{p}, M)$. By lemma 4.10, therefore, $\mathbf{W}(T)$ is not in $\mathbf{A}$. ▮

Lemma 4.17 establishes that, if M strongly separates CD[B] from CD[A], as defined in definition 4.8, then there actually is a path along M that separates B from A.

Lemma 4.18: Let D be a display and let C and C' be configurations over objects A and B. Let Δ be the minimum of the diameters of D[A] and D[B], and let $\lambda = \mathrm{p}^D(C, C')$. If $2\lambda < \Delta$, then $C'D$ approximates CD in tangent with parameters $(\lambda, 2\sin^{-1}(\lambda/\Delta))$. Moreover $C'C^{-1}$ is the approximating homeomorphism.

Proof: Let $\mathbf{p}$ and $\mathbf{q}$ be two points either both in D[A] or both in D[B] such that $\mathrm{d}(\mathbf{p}, \mathbf{q}) = \Delta$. If there is a rotation of angle ψ between C and C' then one of $\mathbf{p}$ and $\mathbf{q}$ must be at least $\Delta/2$ from the center of rotation, and therefore must move a distance at least $\Delta\sin(\psi/2)$. Since neither $\mathbf{p}$ nor $\mathbf{q}$ moves more than λ, ψ must be no more than $2\sin^{-1}(\lambda/\Delta)$. The result is then immediate. ▮

Corollary 4.19: Let $\langle D, C\rangle$ be a scenario over objects A,B. For any $\epsilon > 0, \psi > 0$, there exist $\mu > 0, \theta > 0, \lambda > 0$ such that, for any scenario $\langle D', C'\rangle$, if $\mathrm{p}^D(C, C') < \lambda$ and D' approximates D in tangent (μ, θ), then $C'D'$ approximates CD in tangent (ϵ, ψ).

Proof: Let Δ be the minimum of the diameters of D[A], D[B]; let $\theta = \psi/2$; let $\mu = \lambda = \min(\epsilon/2, \Delta\sin(\theta/2))$; and apply lemmas 4.2 and 4.18. ▮.

Lemma 4.20: Let $\langle D, C\rangle$ be a feasible scenario over objects A,B, and let M be a motion that strongly separates CD[B] from CD[A]. Then there exist $\epsilon > 0, \gamma > 0, \theta > 0$ satisfying the following: Let D_0 be a display that approximates D in tangent (γ, θ) and let C_1 and C_2 be feasible configurations

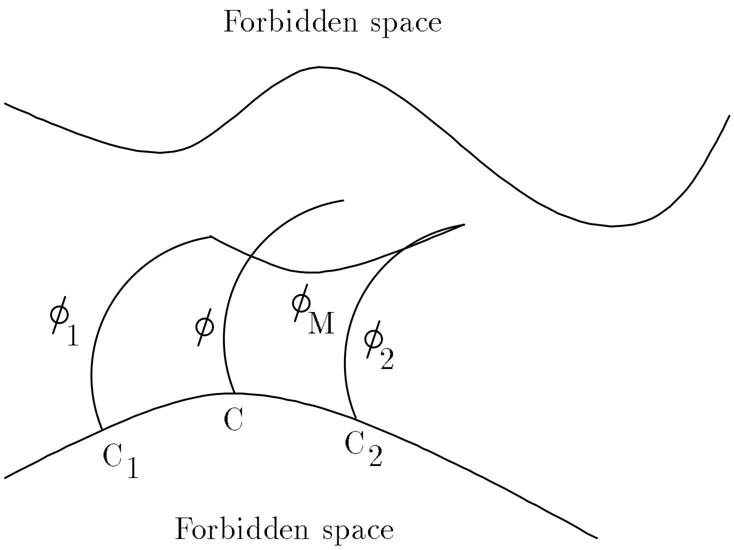


Figure 9: Illustration for lemma 4.20

over D_0 such that C_1[A] = C_2[A]=C[A], $\mathrm{p}^D(C, C_1) < \epsilon$ and $\mathrm{p}^D(C, C_2) < \epsilon$. Then there is a path connecting C_1 and C_2 that is feasible over D_0.

Proof: (Figure 9) Find β and ψ to satisfy lemma 4.16 for $\mathbf{A} = CD$[A], $\mathbf{B} = CD$[B]. Using corollary 4.19, find $\epsilon_0, \gamma_0, \theta_0$ such that, if D' approximates D (γ_0, θ_0) and $\mathrm{p}^D(C, C') < \epsilon_0$ then $C'D'$ approximates CD (β, ψ). Using lemma 4.17, there exists $\Delta > 0$ such that, for all $T \in (0, 1]$, arc$(C, M, \Delta)(T)$ $\in$cfree(D). Let $\phi =$arc$(C, M, \min(\epsilon_0/2, \Delta))$ and let $\alpha = (1/2)$clearance$(\langle D, \phi(1)\rangle))$ Using corollary 4.19 again, find $\epsilon_1, \gamma, \theta$ such that, if D' approximates D (γ, θ) and $\mathrm{p}^D(C, C') < \epsilon_1$ then $C'D'$ approximates CD $(\min(\beta, \alpha), \psi)$. Finally, let $\epsilon = \min(\epsilon_1, \epsilon_0/2)$.

Now, let D_0 be a display that approximates D in tangent (γ, θ) and let C_1 and C_2 be feasible configurations over D_0 within ϵ of C such that C_1[A] = C_2[A]=C[A], Let ϕ_1, ϕ_2 be paths parallel to ϕ starting at C_1, C_2; that is, $\phi_i(T) = \phi(T)C^{-1}C_i$ for $i = 1, 2$. Let ϕ_M be the uniform translation or rotation taking $\phi_1(1)$ into $\phi_2(1)$ through an angle less than π. We now claim that the path $\phi_0 = \langle \phi_1, \phi_M, \phi_2^{-1}\rangle$ is feasible.

We first note that, by construction C_1D_0 approximates CD in tangent (α, ψ) and therefore $d_H(C_1D_0, CD) < \alpha$. Let $\mathbf{p}$ be any point in D_0[B]. Let Γ[B] be the approximating homeomorphism on B and let $\mathbf{q} = \Gamma[\mathrm{B}]^{-1}(\mathbf{p})$. Then $\mathrm{d}(C_1(\mathbf{p}), C(\mathbf{q})) < \alpha$ and $\mathrm{d}(C_2(\mathbf{p}), C(\mathbf{q})) < \alpha$. Since ϕ, ϕ_1, ϕ_2 all move in parallel along the motion M, it follows that, for every $T \in [0, 1]$, $\mathrm{d}(\phi_1(T)(\mathbf{p}), \phi(T)(\mathbf{q}))$ $= \mathrm{d}(\phi_1(0)(\mathbf{p}), \phi(0)(\mathbf{q})) = \mathrm{d}(C_1(0)(\mathbf{p}), C(\mathbf{q})) < \alpha$; and by the same token $\mathrm{d}(\phi_2(T)(\mathbf{p}), \phi(T)(\mathbf{q})) < \alpha$. Let $\mathbf{A} = CD$[A] and $\mathbf{A}_0 = CD_0$[A]. Since the position of object A is constant, $\mathbf{A}_0$ is the place of A along the entire path ϕ_0. Since $\mathrm{d}_H(\mathbf{A}, \mathbf{A}_0) < \alpha$, and since clearance$(D, \phi(1)) > 2\alpha$, it follows that $\mathrm{d}(\phi(1)(\mathbf{q}), \mathbf{A}_0) > \alpha$, and so $\phi_1(1)(\mathbf{p}) \notin \mathbf{A}_0$. Thus $\phi_1(1)$ is feasible, and, by the same token, so is $\phi_2(1)$.

Moreover, the trace of point $\mathbf{p}$ along path ϕ_M is either a circle of arc at most π or a line segment. Hence the maximal value over C_M on ϕ_M of $\mathrm{d}(C_M(\mathbf{p}), \phi(1)(\mathbf{q}))$ is attained either at $C_M = \phi_1(1)$ or at $C_M = \phi_2(1)$ and so is always less than α. Hence $C_M(\mathbf{p})$ is never in $\mathbf{A}_0$, so the path ϕ_M is feasible over D_0.

To show that ϕ_1 is feasible over D_0, we use proof by contradiction. Suppose ϕ_1 is not feasible.

Let T be the maximal value such that, for all $T' \in [0,T]$, $\phi_1(T')$ is feasible. Since free(D_0) is closed and since $\phi_1(0)$ is feasible, $\phi_1(T)$ is feasible. Now $\mathrm{p}^D(C,\phi_1(T)) \leq \mathrm{p}^D(C,C_1) + \mathrm{p}^D(C_1,\phi_1(T)) \leq \epsilon + \epsilon_0/2 < \epsilon_0$. Hence $\phi_1(T)(D_0)$ approximates CD in tangent (β,ψ). Let $\mathbf{X} = \phi_1(T)D_0[\mathrm{A}]$ and $\mathbf{Y} = \phi_1(T)D_0[\mathrm{B}]$. By lemma 4.16, M strongly separates $\mathbf{Y}$ from $\mathbf{X}$. By lemma 4.10, there is a $\Gamma > 0$ such that arc($\phi_1(T), M, \Gamma$) escapes contact over D_0. But this path is just a continuation of the path ϕ_1 past time T, which is a contradiction.

The proof for ϕ_2 is the same as for ϕ_1. ∎.

Definition 4.14: A display D over two objects A and B is *always strongly separable* if D[A] and D[B] are both solid polycurves, and, for every any feasible configuration C over D, there exists a motion M that strongly separates CD[B] from CD[A].

Lemma 4.21: Let D be a display over two objects A and B that is always strongly separable. Then there exist $\epsilon > 0, \gamma > 0, \theta > 0$ satisfying the following: Let D_0 be a display that approximates D in tangent (γ,θ); let C be a feasible configuration over D; and let C_1 and C_2 be feasible configurations over D_0 such that $C_1[\mathrm{A}] = C_2[\mathrm{A}] = C[\mathrm{A}]$, $\mathrm{p}^D(C,C_1) < \epsilon$ and $\mathrm{p}^D(C,C_2) < \epsilon$. Then there is a path connecting C_1 and C_2 that is feasible over D_0.

Proof: Let Δ = diameter(D[A]) + diameter(D[B]), and let us constrain ϵ and γ to be less than Δ. We can then w.l.o.g. restrict attention to the compact space U of configurations in which the configuration of object A is the same as in C and and the distance from A to B is less than or equal to 2Δ.

We apply lemma 4.3 yet again, U being this compact space of configurations, and $\Phi(C,\epsilon,\gamma,\theta)$ being the property "For any display D_0 that approximates D in tangent (γ,θ); and configurations C_1 and C_2 that are feasible over D_0, if $C_1[\mathrm{A}] = C_2[\mathrm{A}]{=}C[\mathrm{A}]$, $\mathrm{p}^D(C,C_1) < \epsilon$ and $\mathrm{p}^D(C,C_2) < \epsilon$, then there is a path connecting C_1 and C_2 that is feasible over D_0." We must show that, if $C \in U$, there exists a $\delta > 0$, $\alpha > 0$, $\beta > 0$, $\psi > 0$ such that, for all C' within δ of C, $\Phi(C',\alpha,\beta,\psi)$. By lemma 4.20, for any C there exists ϵ,γ,θ such that $\Phi(C,\epsilon,\gamma,\theta)$. The result is then immediate if we choose $\alpha = \delta = \epsilon/2$; $\beta = \gamma$, and $\psi = \theta$. ∎

We now have to generalize this to the case where object A is not fixed in C, C_1 and C_2. Intuitively, this is clear enough, since we can just view everything from A's reference frame. However, the point is subtle enough that it's worth working through explicitly.

Lemma 4.22: Let D be a display over two objects A and B that is always strongly separable. Then there exist $\epsilon > 0, \gamma > 0, \theta > 0$ satisfying the following: Let D_0 be a display that approximates D in tangent (γ,θ); let C be a feasible configuration over D; and let C_1 and C_2 be feasible configurations over D_0 such that $\mathrm{p}^D(C,C_1) < \epsilon$ and $\mathrm{p}^D(C,C_2) < \epsilon$. Then there is a path connecting C_1 and C_2 that is feasible over D_0. (This is the same as lemma 4.21, but dropping the condition $C_1[\mathrm{A}] = C_2[\mathrm{A}] = C[\mathrm{A}]$.)

Proof: Let $\Delta = 2$ (diameter(D[A]) + diameter(D[B])). As in the proof of lemma 4.21, we can confine attention to configurations C in which the diameter of $CD[\mathrm{A}] \cup CD[\mathrm{B}]$ is at most Δ, the lemma being trivial if it is more than Δ. Let $E{=}\Delta$/diameter(D[A]). Find ϵ_0, γ_0, θ_0 to satisfy lemma 4.21. Let $\epsilon_1 = \epsilon_0/4E$. Using lemma 4.18 find ϵ, γ, θ such that, for any C_0, D_0 if D_0 approximates D in tangent (γ,θ) and C_0 is within ϵ of C then $C_0 D_0$ approximates CD in tangent $(\min(\gamma_0,\epsilon_1/2),\theta_0)$.

Now let D_0 be a display approximating D (γ,θ) and let C_1 and C_2 be feasible configurations over D_0 within ϵ of C. Let $\Lambda_1 = C^{-1}C_1$ (again, composition is performed placewise, so $C^{-1}C_1$ denotes the pair of transformations $\langle C^{-1}[\mathrm{A}]C_1[\mathrm{A}], C^{-1}[\mathrm{B}]C_1[\mathrm{B}]\rangle$.) Let $D_1 = \Lambda_1 D_0$. Then $CD_1 = C_1 D_0$. Since $\langle D_0, C_1\rangle$ is feasible, so is $\langle D_1, C\rangle$. Also by construction, D_1 approximates D $(\min(\gamma_0,\epsilon_1/2),\theta_0)$, the approximating homeomorphism being Λ_1. Thus, Λ_1 in effect eliminates the shift of configuration from C to C_1 by shifting the shapes from D_0 to D_1.

Let $C_3 = C_2\Lambda_1^{-1}$; thus $C_3 D_1 = C_2 D_0$. Thus, C_3 is the replacement needed for C_2 since we have shifted our shapes from D_0 to D_1. Again since $\langle D_0, C_2 \rangle$ is feasible, so is $\langle D_1, C_3 \rangle$. Now, for any point $\mathbf{p} \in D$[A], $\mathrm{d}(C_2[A](\mathbf{p}), C_3[A](\mathbf{p})) = \mathrm{d}(C_3[A]\Lambda_1[A](\mathbf{p}), C_3[A](\mathbf{p})) =$ (since C_3[A] is a rigid transformation that preserves distances) $\mathrm{d}(\mathbf{p}, \Lambda_1[A](\mathbf{p})) \leq \epsilon_1/2$. Likewise, for any point $\mathbf{q} \in D$[B], $\mathrm{d}(C_2[B](\mathbf{q}), C_3[B](\mathbf{q})) < \epsilon_1/2$. Hence $\mathrm{p}^D(C_3, C_2) \leq \epsilon_1/2$ and $\mathrm{p}^D(C_3, C) \leq \epsilon_1 < \epsilon_0/2$.

There may also be a net motion of A between C_1 and C_2, which will appear as a congruent motion between C and C_3. We therefore split the change from C to C_3 into two parts: the relative motion of B with respect to A, which will take us from C to a new configuration C_4, and the uniform rotation of both object together taking us from C_4 to C_3. To do this, let Λ_2 be the single transformation $C[\mathrm{A}]C_3^{-1}[\mathrm{A}]$, and let $C_4 = \langle \Lambda_2, \Lambda_2 \rangle C_3 = \langle C[\mathrm{A}], \Lambda_2 C_3[\mathrm{B}] \rangle$. Since C_4 differs from C_3 only by the same rigid mapping Λ_2 on both objects, C_4 is likewise feasible. Also, the maximum distance moved by any point in A from C_4 to C_3 is $\mathrm{p}^A(C_3, C_4)$, and the ratio of the maximum distance moved by any point in B from C_4 to C_3 divided by the maximum distance moved by any point in A from C_4 to C_3 is at most $2E$. Hence $\mathrm{p}^D(C_3, C_4) \leq 2E\mathrm{p}^A(C_3, C_4) = 2E\mathrm{p}^A(C_3, C) \leq 2E\epsilon_1 = \epsilon_0/2$. So $\mathrm{p}^D(C_4, C) < \epsilon_0$.

Thus the conditions of lemma 4.21 are satisfied for the pair C,C_4 over the display D_1, so there is a path ϕ_4 through free(D_1) connecting C to C_4. Let ϕ_3 be the path consisting of ϕ_4 followed by the continuous rigid movement of the whole scenario from C_4 to C_3. Then ϕ_3 connects C with C_3 and is feasible over D_1. Finally, let $\phi = \phi_3\Lambda_1^{-1}$. Then for any T, $\phi(T)(D_0) = \phi_3(T)(D_1)$. Hence ϕ is feasible over D_0 and connects C_1 with C_2. ∎

Definition 4.15: Let D and D' be two displays. The connected components of free(D) and free(D') *correspond* if every connected component of free(D) intersects exactly one connected component of free(D') and vice versa.

Theorem 4.23: Let D be a display over two objects A and B that is always strongly separable, such that free(D) is ordinarily connected. Then there exist $\alpha > 0$ and $\psi > 0$ such that, if D' is a display that approximates D in tangent (α, ψ), then the connected components of free(D') correspond to those of free(D).

Proof: Find ϵ, γ, ψ to satisfy lemma 4.22. Let Δ be the minimal distance between any two path-connected components of free(D); since free(D) is ordinarily connected, $\Delta > 0$. Let $\delta = (1/2)\min(\epsilon, \Delta)$

It is immediate from lemma 4.17 that free(D) is equal to the closure of cfree(D). Therefore, by theorem 2.15 there exists a $\lambda_1 > 0$ such that, if $\mathrm{d}_{Hd}(D, D') < \lambda_1$ for display D', then $\mathrm{p}_H^D(\mathrm{free}(D),\mathrm{free}(D')) < \delta$. Let λ_2 be the minimum, over all path-connected components $\mathcal{P}$ of free(D), of the maximum of clearance(D, C) for $C \in \mathcal{P}$. Let $\alpha = \min(\lambda_1/2, \lambda_2/2, \delta)$.

Now let D' be any display that approximates D in tangent (α, ψ). Note that $\mathrm{d}_{Hd}(D, D') < 2\alpha$. We observe the following:

1. For any configuration C, clearance(D', C) > clearance(D, C) $- \lambda_2$. Therefore, in every path-connected component of free(D) there is a configuration $C \in$free(D').
2. Any configuration $C' \in$free(D') is within δ of some configuration C in free(D), and therefore within δ of the path-connected component of free(D) containing C.
3. Since δ is less than the distance between any two distinct path-connected components of free(D), no configuration $C' \in$free(D') is within δ of two different path-connected components of free(D). Therefore, if C'_1 and C'_2 are not within δ of the same path-connected component of free(D), then there cannot be a path from C'_1 to C'_2 through free(D').
4. Let C'_S and C'_E be any two configurations in free(D') that are within δ of a single path-connected component $\mathcal{P}$ of free(D). Let C_S, C_E be configurations in $\mathcal{P}$ that are within δ of

C'_S, C'_E respectively. Let ϕ be the path connecting C_S to C_E. Find a sequence of times $T_0 = 0, T_1 \dots T_n = 1$ such that, for $i = 1 \dots n$, $\mathrm{p}^D(\phi(T_{i-1}), \phi(T_i)) < \delta$. Choose a sequence C'_i such that $C'_0 = C'_S$, $C'_n = C'_E$, and for $i = 1 \dots n-1$, $\mathrm{p}^D(C'_i, \phi(T_i)) < \delta$. Then for each $i = 1 \dots n$, $\mathrm{p}^D(C'_{i-1}, \phi(T_i)) < \mathrm{p}^D(C'_{i-1}, \phi(T_{i-1})) + \mathrm{p}^D(\phi(T_{i-1}), \phi(T_i)) < 2\delta < \epsilon$. Therefore, by lemma 4.22, there exists a path from C'_{i-1} to C'_i through free(D'). Stringing these together gives a path from C'_S to C'_E through free(D').

Thus, for any path-connected component $\mathcal{P}$ of free(D), the set $\mathcal{P}'$ of all configurations within δ of $\mathcal{P}$ is a path-connected component of free(D') and intersects $\mathcal{P}$. Moreover, every configuration in free(D') is an element of one such $\mathcal{P}'$. Thus, our theorem is complete. ▮.

The same argument allows us to achieve criterion (3) mentioned at the start of section 3:

Definition 4.16: For any metric spaces S and T, we say that points $s \in S$ and $t \in T$ correspond within ϵ if any path p starting at s can be tracked within ϵ by a path q starting at t and vice versa. Say that S and T are close within ϵ if, for every point $s \in S$ there is a point $t \in T$ corresponding within ϵ and vice versa.

Theorem 4.24: Let D be a display over two objects A and B that is always strongly separable, and such that free(D) is ordinarily connected. Then for any $\epsilon > 0$ there exist $\alpha > 0$ and $\psi > 0$ such that, if D' is a display that approximates D in tangent (α, ψ), then free(D') and free(D) are close within ϵ.

Proof: Same construction as the proof of theorem 4.23. Change the construction in lemma 4.20 to use a path $\phi = \mathrm{arc}(C, M, \min(\Delta, \lambda/2, \epsilon))$. ▮

It is possible to extend theorem 4.23 to systems of several objects along the following lines: Let ϕ be a continuously differentiable path through configuration space. For point $\mathbf{p}$, define the *relative motion* of $\mathbf{p}$ in object J relative to object I under path ϕ as the velocity of $\mathbf{p}$ under the motion of J relative to its velocity under the motion of I: $(d/dt\phi(T)[J](\mathbf{p})) - (d/dt\phi(T)[I](\mathbf{p}))$ A motion ϕ is strongly separating in scenario $\langle D, C\rangle$ where $C = \phi(0)$ if, for every pair of objects $I \neq J$, and for every point $\mathbf{p} \in CD[I] \cap CD[J]$, the motion of $\mathbf{p}$ in J relative to I is non-zero and its direction strongly separates $CD[J]$ from $CD[I]$ at $\mathbf{p}$. Display D is always strongly separable if the objects in D are all solid polycurves and for every $C \in$free(D) there exists a motion ϕ that is strongly separating in scenario $\langle D, C\rangle$.

Theorem 4.25: Let D be a display over n objects that is always strongly separable, and such that free(D) is ordinarily connected. Then there exist $\alpha > 0$ and $\psi > 0$ such that, if D' is a display that approximates D in tangent (α, ψ), then the connected components of free(D') correspond to those of free(D).

Proof: The proof is exactly analogous to that of theorem 4.23. We have replaced circular motion by general continuous differentiable motion (even if we required each object to move in a circle, their relative motion would be a cycloid). However, the flow field of any differentiable motion at any instant is equal to the flow field of a uniform motion, and, by continuity, stays close to that uniform motion over some time. This is sufficient to carry out the above proof. ▮

It seems plausible to conjecture that the analogous theorem holds in $\Re^k$ for $k > 2$ as well, and, indeed the same proof applies directly in the case of smooth shapes. However, I have not been able to prove the analogue of lemma 4.14 for a suitable class of piecewise smooth shapes in dimensions higher than 2. It is also plausible to conjecture that, if D' approximates D sufficiently well in tangent, then free(D') is homeomorphic to free(D) and, further, that free(D') can be made to approximate free(D) arbitrarily well in tangent. This last result is likely to be necessary, or at least useful, in proving that small changes to a mechanism do not change its qualitative behavior, since corners in configuration space correspond to jammed positions of a mechanism, and it is therefore desirable to give conditions under which no such corners can emerge. However, I have not proven either of these results.

Two further limitations of the above result should be noted. First, since the proof is non-constructive (and very conservative), it gives very little help in answering the practical questions, "Given a situation, how accurate an approximation is needed to get reliable results?" or "Given an approximation of known accuracy, do the results apply to the actual situation?" An algorithm that addresses these problems is given in [Joskowicz, Sacks, and Srinivasan, 95]. Second, theorems 4.23 - 4.25 apply only to motions that lift one object off of the other. By contrast, in physical reasoning about kinematic systems, the most important motions are those that maintain contact, and the most important approximations are the "gap-filling" approximations ([Joskowicz, 89], [Davis, in prep]) in which a system with a narrow contact-free configuration space is approximated by one in which there is no contact-free space. Such motions and approximations are not addressed by the above theorems at all.

References

Ernest Davis, "Approximation and Abstraction in Solid Object Kinematics," in preparation.

Rudolf Fleischer, Kurt Melhorn, Gunter Rote, Emo Welzl, and Chee Yap, "Simultaneous Inner and Outer Approximation of Shape," *Algorithmica,* vol. 8, 1992, pp. 365-389.

Leo Joskowicz, "Simplification and Abstraction of Kinematic Behaviors," *Proc. IJCAI-89,*, pp. 1337-1342.

Leo Joskowicz and Elisha Sacks, "Computational Kinematics," *Artificial Intelligence,* vol. 51, 1991, pp. 381-416.

Leo Joskowicz, Elisha Sacks, and Vijay Srinivasan, "Functional Kinematic Tolerancing," *Computer-Aided Design,* 1996, to appear.

Leo Joskowicz and Russell Taylor, "Interference-Free Insertion of a Solid Body into a Cavity: Algorithms and a Medical Application," submitted to *International Journal of Robotics Research,* MIT Press, 1994.

Bhubaneswar Mishra, *Algorithmic Algebra,* Springer-Verlag, New York, 1993.

James R. Munkres, *Topology: A First Course,* Prentice-Hall, 1975.

Aristides A.G. Requicha, "Towards a Theory of Geometric Tolerancing," *The International Journal of Robotics Research,* vol 2, no. 4, 1983, pp. 45-60.