On the Motion of Objects in Contact*

John Hopcroft
Gordon Wilfong

TR 84-602
May 1984

Department of Computer Science
Cornell University
Ithaca, New York 14853

*This work was supported in part by NSF grant ECS-8312096 and an NSERC graduate scholarship.

ON THE MOTION OF OBJECTS IN CONTACT

John Hopcroft and Gordon Wilfong
Department of Computer Science
Cornell University
Ithaca, N.Y. 14853

Abstract

There is an increasing use of computers in the design, manufacture and manipulation of physical objects. An important aspect of reasoning about such actions concerns the motion of objects in contact. The study of problems of this nature requires not only the ability to represent physical objects but the development of a framework or theory in which to reason about them. In this paper such a development is investigated and a fundamental theorem concerning the motion of objects in contact is proved. The simplest form of this theorem states that if two objects in contact can be moved to another configuration in which they are in contact, then there is a way to move them from the first configuration to the second configuration such that the objects remain in contact throughout the motion. This result is proved when translation and rotation of objects are allowed. The problem dealing with more generalized types of motion is also discussed. This study has obvious applications in compliant motion and in motion planning.

Keywords: compliant motion, motion planning, solid modelling, robotics, complexity theory

---

This work was supported in part by NSF grant ECS-8312096 and an NSERC graduate scholarship.

Introduction

The increasing use of computers in the design, manufacture and manipulation of physical objects underscores the need for a theory to provide a framework for reasoning about transformations of objects. In this paper we take a first step towards developing such a theory.

One may define an object such as a shaft or connecting rod without instantiating its position or orientation. More generally one can define a rectangular solid without specifying its dimensions. In fact one can define an object without instantiating its shape. For example, the shape of an ellipsoid parameterized by the ratio of its axes is determined only when the parameter is fixed. To obtain a specific instance of the ellipsoid one provides the ratio of major to minor axis, the length of the major axis and its orientation with respect to some coordinate system along with the position in 3-space of the center of the ellipsoid. In this framework an object is the image of an instance of a parameterized homeomorphism from a canonical region in $\mathbb{R}^3$ to $\mathbb{R}^3$. Thus all instances of an object are homeomorphic and therefore topologically equivalent. A sphere and an ellipsoid can be instances of the same object but a sphere and a torus cannot. A sphere might be given by the parameterized mapping

$$f(x,y,z)=(xr+a,\ yr+b,\ zr+c)$$

from a unit sphere centered at the origin in $\mathbb{R}^3$ where a,b and c give the coordinates of the center and r is the radius. Corresponding to a particular instantiation of the sphere is a point in the four dimensional parameter space.

In this generalized setting a motion is a continuous mapping from [0,1] into the appropriate parameter space. Thus a motion is a path in the parameter space. A motion can be a combination of translation, rotation, growth or more complicated continuous deformation of shape. It is our hope that this view will be useful in defining and manipulating generic objects as well as deformable or nonrigid objects. For our purposes we limit ourselves to motions where the transformation can be parameterized by a finite number of parameters.

Although it is traditional to think of objects in terms of their shape and dimension and then to deduce functionality from the shape, it is enticing to think of representing objects by functionality and then deducing shape. In designing for automatic assembly one is normally free to modify objects for ease of assembly. Thus designing for functionality and allowing the functionality and assembly process to determine shape and size is a desirable goal. Furthermore, parameterized design provides additional advantages. For example, instead of designing a drive shaft for a particular torque, it would be preferable to design the drive shaft with torque as a parameter. This allows changes in design specification without necessitating redesign of components. The study of problems of this nature will require substantial advances in the representation of physical objects and in our ability to reason about them. In this paper we begin with a modest step by establishing a fundamental theorem concerning the motion of objects in contact.

In the special case where motion is restricted to rotations and translations the theorem states that if there is a way to move a set of objects from an initial configuration where the objects form a connected component to a final configuration where the objects form a connected component then there is a way to move the objects from the initial to the final configuration such that at all times the objects form a connected component. To understand the theorem in a more general setting consider the motion of two objects A and B relative to one another. Normally one would consider A fixed and that B moves relative to A. For ordinary motions such as translations or rotations there is of course no loss in generality in fixing A. However, the fact that A may be changing shape makes it more desirable to view both objects as moving. A point in configuration space represents the values for the parameters of A and B. Certain points correspond to positions and orientations where B overlaps A. In the situation where configuration space is contractible to a point the theorem states that the existence of a path, P_1 in Figure 1, from initial configuration to final configuration where A and B always intersect and of a path, P_2 in Figure 1, where A and B do not overlap implies the existence of a path, P_3 in Figure 1, where A and B touch at all times but are not overlapping. One should observe that the point of contact need not be a continuous function even though motion is continuous.

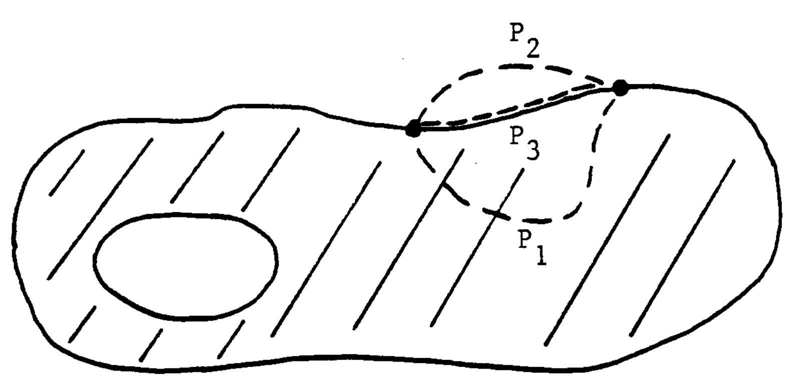

Figure 1. Configuration Space

Care must be exercised in applying the theorem. For example, in Figure 2 there are two objects A and B, A is fixed and B is permitted only to rotate about x. Rotating B $2\pi-\alpha$ in the clockwise direction results in the same apparent configuration as rotating B α radians in the counterclockwise direction. However, these two configurations are the same in configuration space only if we identify points that differ by a rotation of 2π. This results in a cylindrical shaped space that is not contractible to a point and hence the hypothesis of our theorem is not valid. Observe that in the above example the only motion from $\theta=0$ to $\theta=2\pi-\alpha$ is a motion where the objects do not overlap and thus there is no motion that keeps the objects intersecting.

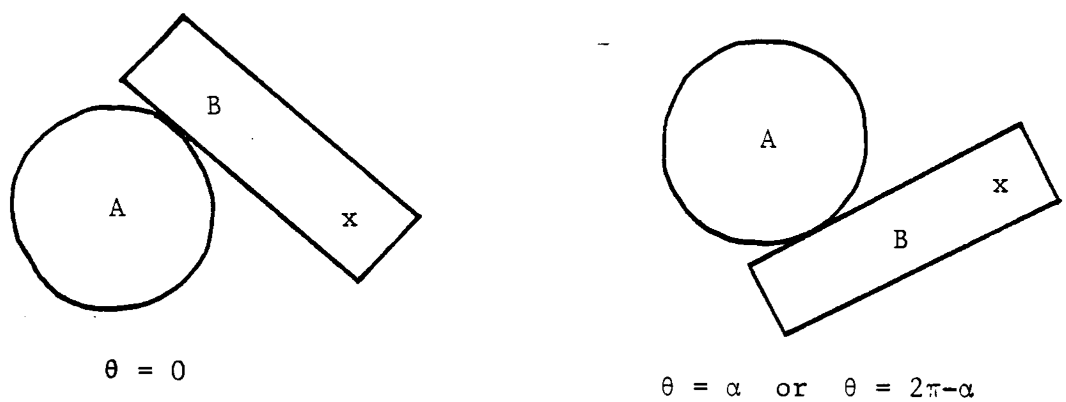

Figure 2. Restriction to rotational motion

In addition to the obvious applications in compliant motion, the theorem has potential applications in motion planning and in complexity theory. In planning of coordinated motion not only must trajectories be determined but also the relative timing of objects as they move on their individual trajectories. A path in configuration space contains this information about relative timing. Thus searching the paths in configuration space conceptually simplifies the problem. In general, configuration space is of very high dimension. The above theorem reduces the search from this high dimensional space to that of a lower dimensional surface in the space. The surface can be thought of as composed of faces that intersect in lower dimensional faces that in turn intersect in still lower dimensional faces. Under suitable restrictions, the surface of contact in configuration space will have vertices that are edge connected. In order to move a set of objects from one configuration to another we first push the objects together in both the initial and final configurations. Then we move the objects along faces until they reach lower dimensional faces. We continue this process until the initial and final configurations have been converted to vertices of the surface in configuration space at which the objects are in contact. This reduces the problem to a graph searching problem. In general the number of vertices of the graph will be astronomical. We need not construct the entire graph but only generate vertices and edges as the search progresses. With a suitable heuristic it may be possible, in practical situations, to find the desired path having only generated a tiny fraction of the graph. The knowledge of such a path could be used in constructing a path where the objects do not touch one another.

In complexity theory it is often important to show that if a certain motion exists, then a canonical motion exists. In the case of linkages [4], for example, it is important from a complexity point of view that various joints need not be moved to locations that are algebraically independent in order for a motion to take place. Our theorem establishes that if a motion exists, then one that follows features of the surface exists and hence a canonical motion exists.

The paper consists of four sections. In the first section, some general properties of the space of all configurations are developed. These properties are used in the second section to show that certain regions of the space of all configurations are path connected or contractible to a point. From this it is shown in the second section that if there is a motion between configurations in which two objects touch then there is a motion between them such that at all times two objects touch. In the third section an inductive argument is developed to show the main result. That is, it is shown that if there is a motion of rotations and translations between two configurations in which the objects form a connected component then there is a motion which keeps the objects in a connected component. In the fourth section we discuss the case where more general motions are allowed.

1. Basic Properties of Configuration Space

Let A be a set in $\mathbb{R}^n$. The *interior* of A, denoted int(A), is the union of all open sets of $\mathbb{R}^n$ contained in A. A point x is a *limit point* of A if there exists a sequence $\{x_i\}$ of points in A such that

$$\lim_{i\to\infty} x_i = x.$$

The *closure* of A is the set of all limit points of A.

An *object* is a convex, compact region of $\mathbb{R}^n$ that is the closure of its interior and is bounded by algebraic surfaces. (The limitation of convexity will be removed later by introducing composite objects.) Each object contains a designated point, called the *origin*, at which the origin of a coordinate system, affixed to the object, is located. The position and orientation of an object are specified by the location of the origin of the object in $\mathbb{R}^n$ and orientation of the affixed coordinate system relative to the coordinate system of $\mathbb{R}^n$. Given a set of objects, a *configuration* is a vector whose components specify the position and orientation of each object. The space of all such vectors is called *configuration space*. Given a set of objects and a point x in the corresponding configuration space we let B(x) denote the region in $\mathbb{R}^n$ occupied by object B in the given configuration. If b is a point on object B then let b(x) be the point in $\mathbb{R}^n$ occupied by b when B is in the position and orientation specified by x.

Objects B_i and B_j intersect in configuration x if $B_i(x) \cap B_j(x) \neq \emptyset$. The objects overlap if their interiors intersect. If the objects intersect but do not overlap then we say that they touch in configuration x.

It is convenient to partition the set of objects into subsets called composite objects. A composite object is intended to be a single object made up of smaller objects. With each composite object associate a graph whose vertices are the objects and whose edges are pairs of objects that intersect. A composite object is connected if the associated graph is connected. A configuration is proper if each composite object is connected. A configuration is valid if in addition to being proper the interiors of each pair of objects in a composite object do not intersect. Let VALID denote the set of valid configurations.

Composite objects are used in an inductive argument in Section 3. By considering two or more objects to be a single composite object we are able to establish a motion for n objects from a motion for n-1 objects, one of which is a composite object. We introduce the notion of valid so that individual objects in a composite object will touch but not overlap throughout the motion.

Let OVERLAP denote the set of valid configurations in which two or more composite objects overlap. Let TOUCH denote the set of valid configurations in which two or more composite objects touch and no two composite objects overlap. Let NONOVERLAP be the complement of OVERLAP with respect to the set of valid configurations.

Figure 3 shows two objects, B_1 and B_2. Object B_2 is stationary and object B_1 is allowed only translational motion. Given two configurations in which the two objects are touching and a motion between the configurations in which the objects do not overlap, we wish to show that there is a motion where the objects are always in contact. The graphic representation of configuration space of Figure 1 suggests that the boundary of the region where the objects intersect corresponds to the configurations where the objects touch. Figure 3 shows that this is not exactly the case. The configurations where B_1 is in the opening of B_2 are not in the boundary of the space of configurations where the objects intersect. The first goal of this section is to show that the configurations where at least two composite objects touch is exactly

$$\text{cl(OVERLAP)} - \text{OVERLAP} .$$

That is, TOUCH = cl(OVERLAP) $\cap$ NONOVERLAP.

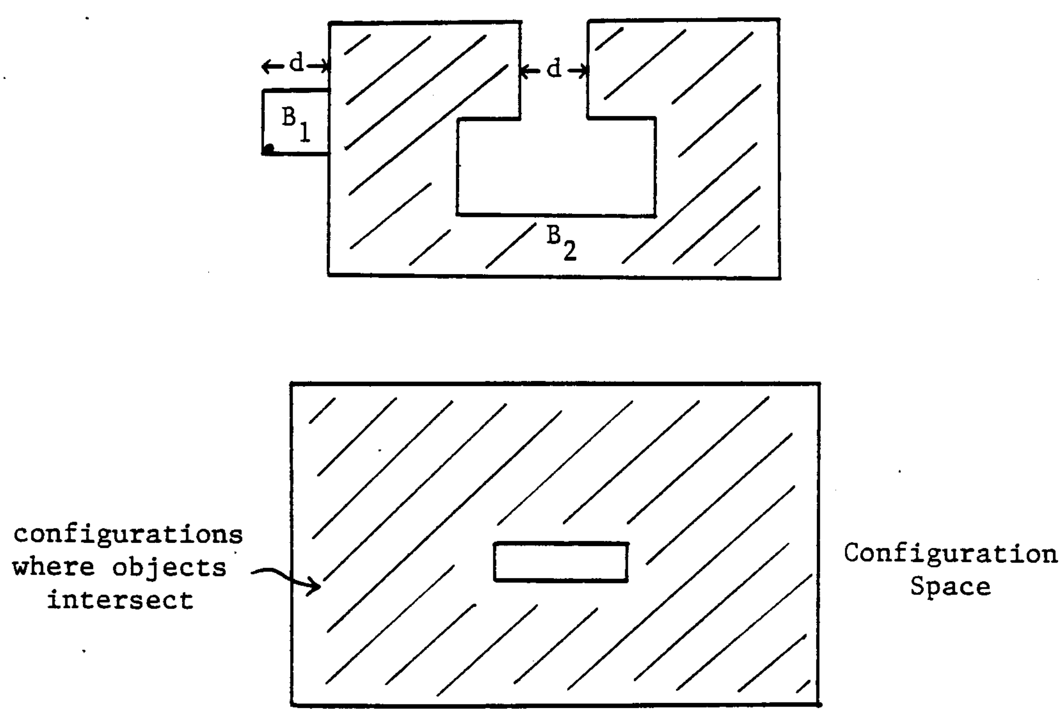

Figure 3. Two objects and corresponding configuration space

The second aim of this section is to prove a lemma concerning TOUCH that will aid us in the next section in proving that certain motions in TOUCH exist. Towards this end, we designate one object in each composite object as the base object. We will call the origin of the base object the origin of the composite object. Let BASE be the set of proper configurations in which the base object of some composite object intersects the base object of the n^{th} composite object. Let

$$\text{FILL} = \text{cl(OVERLAP)} \cup \text{BASE} .$$

In Section 2 we will need the fact that FILL ∪ NONOVERLAP is contractible to a point. Suppose we had not included BASE in FILL. Consider the example shown in Figure 4. Here we have three circles that are allowed to move along a line. Object B is fixed and objects A and B form one composite object in which B is the base object. In this case NONOVERLAP is four rays and cl(OVERLAP) ∪ NONOVERLAP is two parallel lines. Thus not only is cl(OVERLAP) ∪ NONOVERLAP not contractible to a point it is not path connected. However when we include BASE, the set FILL ∪ NONOVERLAP becomes contractible to a point. Thus the points of FILL fill in the holes of NONOVERLAP so that the union of FILL and NONOVERLAP is contractible to a point. We will show that

$$\text{TOUCH} = \text{FILL} \cap \text{NONOVERLAP} .$$

This will be used in the next section to show that if there is a path in NON-OVERLAP between two configurations in TOUCH then there is a path in TOUCH between them. Throughout this section B_i will denote an object and A_i will denote a composite object.

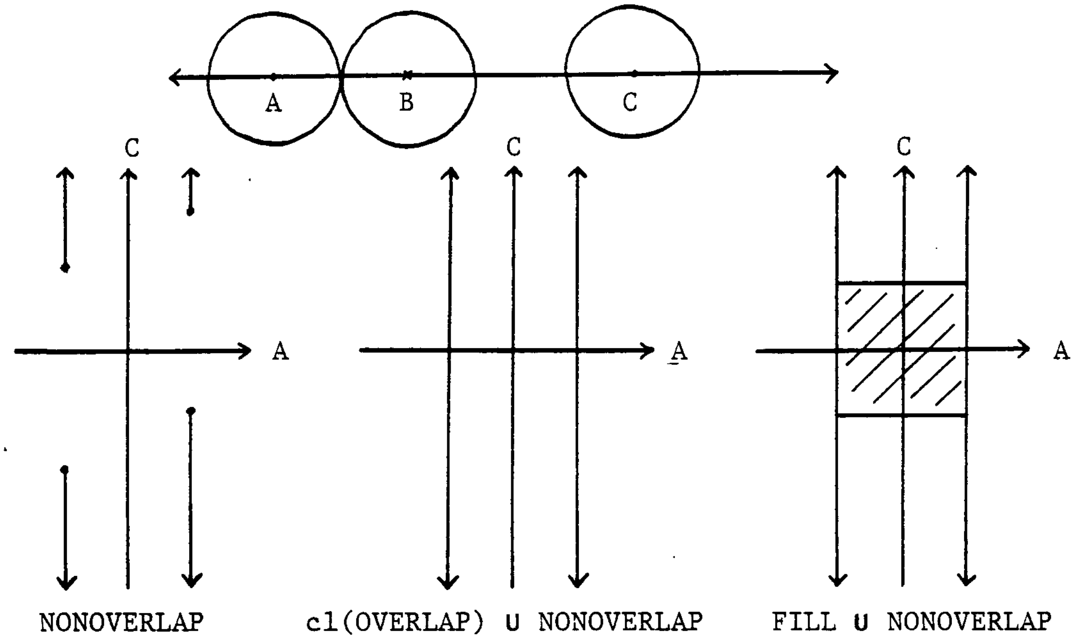

Figure 4. Three circles constrained to horizontal motion, B is fixed. A and B form a composite object with B as base.

We now proceed with a series of lemmas. Lemmas 1.1 through 1.4 are used to establish in Theorem 1.5 that TOUCH = cl(OVERLAP) ∩ NONOVERLAP and in Theorem 1.6 that TOUCH = FILL ∩ NONOVERLAP. This latter theorem is used in the next section to prove that certain motions in TOUCH exist.

First we show that for any configuration in which two objects intersect there is an arbitrarily close configuration in which the two objects overlap.

__Lemma 1.1__: Let $S = \{x | int(B_i(x)) \cap int(B_j(x)) \neq \emptyset\}$. Then $\{x | B_i(x) \cap B_j(x) \neq \emptyset\} \subseteq cl(S)$.

__Proof__: Let x be a configuration such that $B_i(x) \cap B_j(x) \neq \emptyset$. Let y be a point on $B_i(x)$ and z be a point on $B_j(x)$ such that y and z occupy the same point b in

$\mathbb{R}^n$. Since an object is the closure of its interior there are sequences of points $\{y_\alpha\}$ of $int(B_i(x))$ and $\{z_\alpha\}$ of $int(B_j(x))$ such that $\lim_{\alpha\to\infty} y_\alpha=y$ and $\lim_{\alpha\to\infty} z_\alpha=z$. Let $\{x_\alpha\}$ be the sequence of configurations in S such that all objects except B_i and B_j have the same position and orientation as in configuration x and $B_i(x_\alpha)$ and $B_j(x_\alpha)$ have the same orientation as in configuration x but y_α and z_α are at position b in $\mathbb{R}^n$. Thus $\lim_{\alpha\to\infty} x_\alpha=x$ and so x is in cl(S). □

Next we show that the closure of the set of configurations in which two objects overlap is contained within the set of configurations where the two objects intersect. This result combined with the previous lemma establishes that these two sets are equal.

<u>Lemma 1.2</u>: $cl(\{x|int(B_i(x))\cap int(B_j(x))\neq\emptyset\}) \subseteq \{x|B_i(x)\cap B_j(x)\neq\emptyset\}$.

<u>Proof</u>: Let $S = \{x|int(B_i(x))\cap int(B_j(x))\neq\emptyset\}$. Let x be in cl(S). Then there exists a sequence of configurations $\{x_\alpha\}$ in S such that $\lim_{\alpha\to\infty} x_\alpha=x$. Corresponding to $\{x_\alpha\}$ is a sequence $\{\langle y_\alpha,z_\alpha\rangle\}$ where y_α and z_α are points of $B_i(x_\alpha)$ and $B_j(x_\alpha)$ that occupy the same location in $\mathbb{R}^n$. Since the objects are compact the cross product space is compact and so there is a subsequence $\{\langle\hat{y}_\alpha,\hat{z}_\alpha\rangle\}$ that converges to some pair of points $\langle y,z\rangle$. Since objects are closed, y and z are in B_i and B_j. Define the usual distance metric d. Since $d(y_\alpha,z_\alpha)=0$ for all α, clearly $\lim_{\alpha\to\infty} d(\hat{y}_\alpha,\hat{z}_\alpha)=0$. Since d is continuous we can move the limit inside and get $d(y,z)=0$. Thus $B_i(x) \cap B_j(x)\neq\emptyset$. □

We now conclude from Lemmas 1.1 and 1.2 that the set of configurations in which two given objects intersect is equal to the closure of the set of configurations in which the interiors of the two objects intersect. Since a composite object is some union of objects we get that two composite objects intersect in a configuration if and only if there is an object in each composite object that intersect in the configuration. Thus, the set of configurations where the two composite objects intersect is some union of sets of configurations where two objects intersect. Similarly the set of configurations where the interiors of two given composite objects intersect is some union of sets of configurations in which the interiors of two objects intersect. Since the closure of the union is the union of the closure, the closure of the set of configurations in which the interiors of two given composite objects intersect is some union of the closure of sets of configurations in which the interiors of two objects intersect. Thus we can conclude that the set of configurations in which two given composite objects intersect is equal to the closure of the set of configurations in which the interiors of the two composite objects intersect. By the same argument we can show that the set of configurations in which there are at least two composite objects which intersect is equal to the closure of the set of configurations where the interiors of at least two composite objects intersect.

The next step is to show that a composite object also has the property that it is the closure of its interior. Note that interior points of a composite object may not be interior points of any object.

Lemma 1.3: For composite object A, A=cl(int(A)).

Proof: Let $A=\bigcup_i B_i$. Suppose $y \in A$. Then $y \in B_i$ for some i. Since B_i=cl(int(B_i)) there is a sequence $\{y_\alpha\}$ in int(B_i) such that $\lim_\alpha y_\alpha = y$. But each $y_\alpha \in$ int(B_i) implies each $y_\alpha \in$ int(A). Thus $y \in$ cl(int(A)) and so $A \subseteq$ cl(int(A)).

Suppose $y \in cl(int(A))$. Then y is the limit point of a sequence in int(A) and hence the limit point of a sequence in A. Since A is closed y must be in A. Thus $cl(int(A)) \subseteq A$. □

Since VALID is a closed set we can compute the closure of OVERLAP by taking the closure of all configurations where two composite objects overlap and then intersecting with VALID.

<u>Lemma</u> <u>1.4</u>: $cl(OVERLAP)=cl(\{x \mid \exists i,j \quad i\neq j \quad int(A_i(x)) \cap int(A_j(x)) \neq \emptyset\} \cap VALID)$ $=\{x \mid \exists i,j \; i\neq j \; A_i(x) \cap A_j(x) \neq \emptyset\} \cap VALID$.

<u>Proof</u>: Let $F=\{x \mid \exists i,j \quad i\neq j \quad int(A_i(x)) \cap int(A_j(x)) \neq \emptyset\}$. Then cl(OVERLAP) = $cl(F \cap VALID) \subseteq cl(F) \cap cl(VALID) = cl(F) \cap VALID$ because VALID is closed.

Let $x \in cl(F) \cap VALID$. We want to show $x \in cl(OVERLAP)$. Since $x \in cl(F)$ we know by the remark after Lemma 1.2 that for some i and j, $i \neq j$, $A_i(x) \cap A_j(x) \neq \emptyset$. By a construction similar to that in Lemma 1.3 we create a sequence $\{x_\alpha\}$ with limit point x such that $int(A_i(x_\alpha)) \cap int(A_j(x_\alpha)) \neq \emptyset$ and $x_\alpha \in VALID$. Thus $x \in cl(F \cap VALID)$ = cl(OVERLAP). □

We can now establish the result that TOUCH = cl(OVERLAP) ∩ NONOVERLAP.

<u>Theorem</u> <u>1.5</u>: TOUCH = cl(OVERLAP) ∩ NONOVERLAP

<u>Proof</u>: By definition $TOUCH=\{x \mid \exists i,j \quad i\neq j \quad A_i(x) \cap A_j(x) \neq \emptyset\} \cap \{x \mid \forall i,j \quad i\neq j \quad int(A_i(x)) \cap int(A_j(x)) = \emptyset\} \cap VALID$. Therefore, $TOUCH = NONOVERLAP \cap \{x \mid \exists i,j \; i\neq j \; A_i(x) \cap A_j(x) \neq \emptyset\} = NONOVERLAP \cap cl(OVERLAP)$ by Lemma 1.4. □

<u>Theorem</u> <u>1.6</u>: TOUCH = FILL ∩ NONOVERLAP

Proof: By definition FILL = cl(OVERLAP) $\cup$ BASE. Since TOUCH = cl(OVERLAP)$\cap$NONOVERLAP by Theorem 1.5 we get that TOUCH $\subseteq$ cl(OVERLAP) $\subseteq$ FILL $\subseteq$ FILL$\cap$NONOVERLAP.

Let $x\epsilon$FILL$\cap$NONOVERLAP. If $x\epsilon$cl(OVERLAP) then $x\epsilon$TOUCH. Suppose $x\epsilon$BASE. Since $x\epsilon$NONOVERLAP we have $x\epsilon$VALID. Also $x\epsilon$BASE implies two composite objects intersect and so $x\epsilon$cl(OVERLAP). Thus $x\epsilon$TOUCH. Therefore, FILL$\cap$NONOVERLAP $\subseteq$ TOUCH. □

2. Requiring Two Objects To Touch Throughout A Motion

In this section we show the following intermediate result. Given two configurations x and y with n objects, at least two of which are touching in each configuration, if it is possible to move the objects from configuration x to configuration y, then it is possible to do so by a motion such that two objects are always touching. To do this, we make use of the Mayer-Vietoris theorem from algebraic topology to show that the path connected components of TOUCH are in one to one correspondence with the path connected components of NONOVERLAP.

Notice that if TOUCH was a retract of NONOVERLAP then if there was a motion in NONOVERLAP between two configurations in TOUCH then we could conclude that there was a motion in TOUCH between these configurations. This is because there must be a continuous function f:NONOVERLAP→TOUCH such that $f(t)=t$ $\forall$ $t\epsilon$TOUCH and so if m is a motion in NONOVERLAP between $t_1,t_2\epsilon$TOUCH then $f(m(t))$ is a continuous path in TOUCH between t_1 and t_2.

However we cannot guarantee that TOUCH will be a retract of NONOVERLAP. Consider Figure 5 where there is one stationary object A and another object B which is free to move about. Configuration space is just $\mathbb{R}^2$ where a configuration consists of the position of the center of the disk. NONOVERLAP is the unshaded region in the figure. Thus in this case NONOVERLAP is not retractible to TOUCH, the boundaries of NONOVERLAP in configuration space.

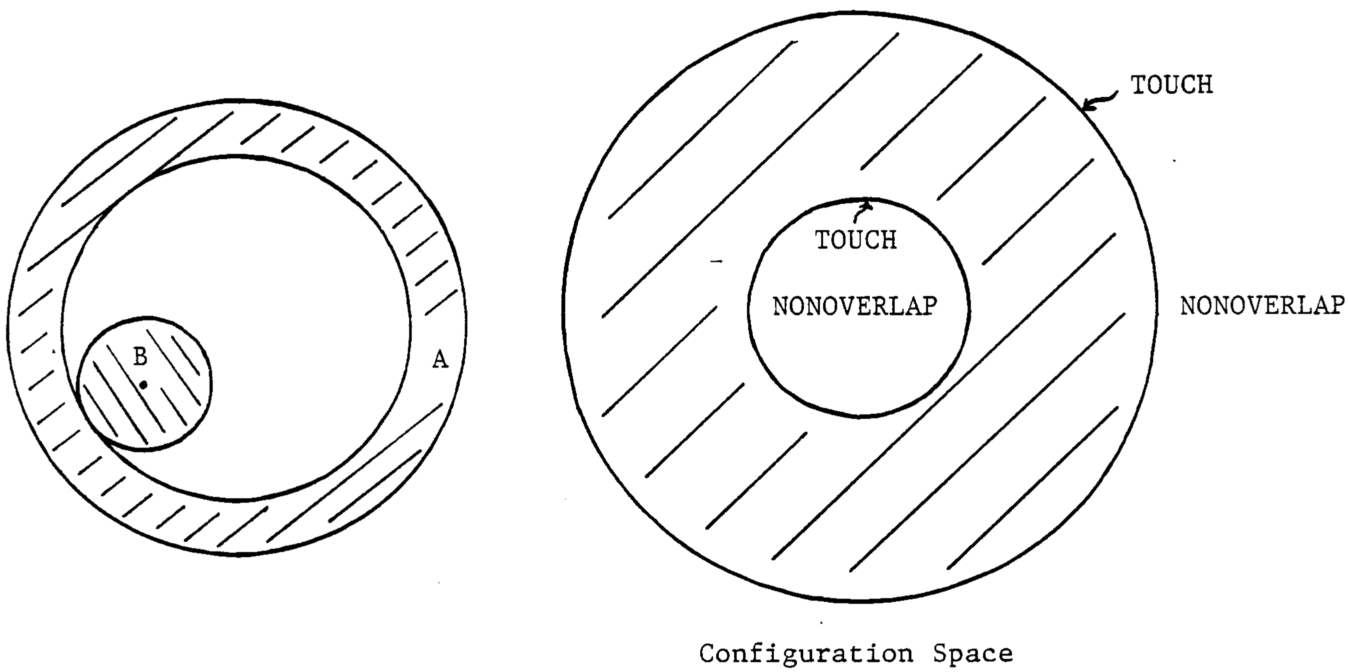

Figure 5. TOUCH not a retract of NONOVERLAP

We begin by defining a motion. A movement of the objects corresponds in an obvious manner to a path in configuration space. Thus a motion is a continuous function from [0,1] to configuration space. If m is a motion then the reversal of the motion m^r is defined as $m^r(t) = m(1-t)$. If m_1 and m_2 are motions where $m_1(1) = m_2(0)$ then the composition $m = m_1 || m_2$ is a motion defined by

$$m(t) = \begin{cases} m_1(2t) & 0 \le t \le 1/2 \\ m_2(2t-1) & 1/2 \le t \le 1. \end{cases}$$

At certain times we shall be concerned with motions where the orientation of each object is maintained while each object is moving along a straight line at a constant rate. Thus in configuration x(t) the location of point b of an object is given by

$$b(x_t) = b(x_0) + [b(x_1) - b(x_0)]t.$$

When we talk about a motion in which objects move in a straight line we are referring to a motion of the above type. When we talk about moving a composite object in a straight line, the objects making up the composite object maintain their relative spacing.

In many of the following results a straight line motion is used. In Lemma 2.1 we show that a straight line motion of two objects keeps the objects intersecting if they intersect at the beginning and at the end of the straight line motion.

<u>Lemma 2.1</u>: Let x and y be configurations and B_1 and B_2 be objects. Suppose that b_1 and c_1 are points of B_1 and b_2 and c_2 are points of B_2 such that $b_1(x)=b_2(x)$ and $c_1(y)=c_2(y)$. Then moving B_1 and B_2 along straight lines so that c_1 and c_2 are positioned at $c_1(y)$ keeps the objects intersecting. See Figure 6.

<u>Proof</u>: At time t during the motion, the point of B_1 on the line between b_1 and c_1 given by $b_t=b_1+(c_1-b_1)t$ occupies the same point as the point of B_2 on the line between b_2 and c_2 given by $b_t=b_2+(c_2-b_2)t$. □

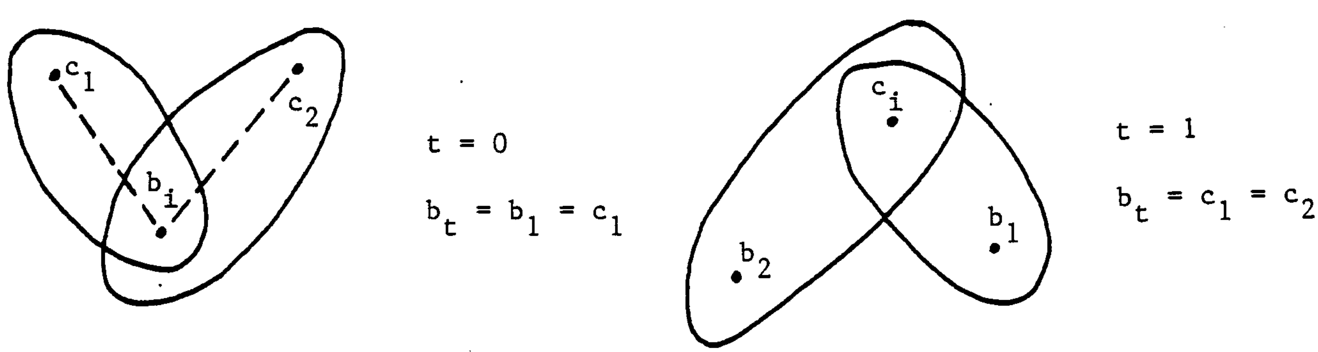

Figure 6. Straight line motion

Now it is shown that there is a motion between any two configurations in PROPER in which the base object of some composite object intersects the base object of the n^{th} composite object such that during the motion all configurations have that property. This is done by showing that there is some fixed configuration with the property such that there is a motion, which keeps the property true, between any configuration with the property and the fixed configuration.

__Lemma 2.2__: BASE is path connected.

__Proof__: Fix some configuration y in BASE such that in y, the origin of every object has the same location. Let xϵBASE. Move the objects in a straight line from configuration x to the configuration z that has the same orientations as x but the location in z of the origins of the objects are as in y. By Lemma 2.1 any objects that intersect in x will intersect throughout the motion. Now rotate the objects about their origins to the orientations given by y. Thus there is a path in BASE from any xϵBASE to y. Since motions are reversible there is a path in BASE from y to any xϵBASE and so BASE is path connected. □

The following theorem shows that FILL is path connected by constructing a motion from any configuration in FILL to some configuration in BASE. BASE is path connected, by the previous result, and since BASE is contained in FILL we conclude that FILL is path connected.

__Theorem 2.3__: FILL is path connected.

Proof: Let $x \in$ FILL. We will show that there is a path in FILL from x to some configuration in BASE and since BASE is path connected and motions are reversible we will conclude that FILL is path connected.

By definition FILL = cl(OVERLAP) $\cup$ BASE. If $x \in$ BASE then we are done. Suppose $x \in$ cl(OVERLAP). Then $A_i(x) \cap A_j(x) \neq \emptyset$ for some i and j, $i \neq j$ and $x \in$ VALID. Let b be the origin of the base object of A_i and b' the origin of the base object of the n^{th} composite object. Move A_i and A_j in a straight line (considering A_i and A_j as one composite object) to the configuration where the location of b is b'(x). All other objects remain stationary. Thus the motion is in cl(OVERLAP) and hence in FILL and the resulting configuration is in BASE. Therefore FILL is path connected. □

The next lemma will be used when we show that FILL $\cup$ NONOVERLAP is contractible to a point. We show that there is a path in VALID from any configuration in VALID to a configuration in BASE in which the origins of all the composite objects have the same location. The same construction can be used to construct a path in BASE from any configuration in BASE to some configuration in BASE in which the origins of all the composite objects have the same location.

Lemma 2.4: From every configuration in VALID (BASE) there is a path in VALID (BASE) to some configuration in BASE in which the origins of the composite objects coincide.

Proof: Let $x \in$ VALID (BASE). Move the composite objects in a straight line from x to the configuration where all the origins have location equal to the location in x of the origin of the n^{th} composite object. The resulting configuration is in BASE and the motion described is as desired by Lemma 2.1. □

We now use the motions constructed in Lemmas 2.2 and 2.4 to show that FILL ∪ NONOVERLAP is contractible to a point. That is, we show that there is a configuration $y \in S = \text{FILL} \cup \text{NONOVERLAP}$ and a continuous function $f: S \times [0,1] \to S$ such that

$$\left.\begin{aligned} f(x,0) &= x \\ f(x,1) &= y \end{aligned}\right\} \forall x \in S$$
$$f(y,t) = y \qquad \forall t \in [0,1] \ .$$

In order that FILL∪NONOVERLAP be contractible to a point we cannot identify a rotation of 2π with no rotation at all as is done in Schwartz and Sharir [7]. Thus in configuration space a dimension corresponding to a rotation is infinite even though every 2π radians the object returns to its apparent initial position.

<u>Theorem 2.5</u>: FILL ∪ NONOVERLAP is contractible to a point.

<u>Proof</u>: Let S=FILL ∪ NONOVERLAP. Then S=BASE ∪ VALID. Let y be the fixed configuration in BASE as in Lemma 2.2. Define $f: S \times [0,1] \to S$

$$f(x,t) = \begin{cases} m_1(2t) & 0 \le t \le 1/2 \text{ where } m_1 \text{ is the motion described in Lemma 2.4 and } m_1(0) = x \\ m_2(2t-1) & 1/2 \le t \le 1 \text{ where } m_2 \text{ is the motion described in Lemma 2.2 and } m_2(0) = m_1(1) \end{cases}$$

Then
$$\left.\begin{aligned} f(x,0) &= x \\ f(x,1) &= y \end{aligned}\right\} x \in S$$
$$f(y,t) = y \qquad t \in [0,1]$$

By the construction of m_1 and m_2, f is continuous. Thus f is a homotopy between the retraction $r:S\to\{y\}$ and the identity $i:S\to S$. That is $\{y\}$ is a deformation retract of S and so S is contractible to a point. □

The above construction gives a motion in S=FILL∪NONOVERLAP from any configuration in S to the fixed configuration $y\in S$. Thus if x_1 and x_2 are two configurations in S and m_i is the motion constructed from x_i to y in S then $m_1 || m_2^r$ is a motion in S from x_1 to x_2 and so S is path connected.

<u>Corollary 2.6</u>: FILL∪NONOVERLAP is path connected.

If configuration space is restricted so that each parameter that corresponds to an orientation is only allowed to range within some closed and bounded interval then the portion of FILL in this restricted space is clearly still path connected. Also FILL∪NONOVERLAP is still contractible to a point in this restricted configuration space. For the rest of this section we will be considering such a restricted configuration space. Thus when we speak of some set such as NONOVERLAP then we will mean the part of the set which is in the restricted configuration space.

<u>Theorem 2.7</u>: NONOVERLAP consists of a finite number of path connected components.

<u>Proof</u>: As in Schwartz and Sharir [7] we divided configuration space into finitely many cells such that the set of polynomials that describe the relative positions of the objects are sign invariant within each cell. Then NONOVERLAP is the finite union of some of these cells and for any two points x,y in a cell there is a path within the cell between them. Thus there must be finitely many path connected components in NONOVERLAP. □

Theorem 2.8 (Mayer-Vietoris): The sequence

$$H_1(A \cup B) \to H_0(A \cap B) \to H_0(A) \oplus H_0(B) \to H_0(A \cup B) \to \{0\}$$

is an exact sequence.

Proof: See Massey [5]. □

In the following we will use the Mayer-Vietoris sequence with A = FILL and B = NONOVERLAP to show that in our restricted configuration space TOUCH and NONOVERLAP have the same number of path connected components.

Theorem 2.9: H_0(NONOVERLAP) $\simeq$ H_0(FILL$\cap$NONOVERLAP) = H_0(TOUCH).

Proof: By Theorem 2.5, S=FILL$\cup$NONOVERLAP is contractible to a point and so $\Pi_1(S)=\{0\}$ (see Massey [6]). Since $\Pi_1(S)$ is abelian $H_1(S)=\Pi_1(S)=\{0\}$. Also we have that $H_0(S)=\mathbf{Z}$ because by Corollary 2.6, S is path connected. By Theorem 2.3 FILL is path connected so $H_0(\text{FILL})=\mathbf{Z}$. Therefore the sequence in Theorem 2.8 is as follows:

$$\{0\} \xrightarrow{h_1} H_0(\text{TOUCH}) \xrightarrow{h_2} \mathbf{Z} \oplus H_0(\text{NONOVERLAP}) \xrightarrow{h_3} \mathbf{Z} \xrightarrow{h_4} \{0\}.$$

Since the sequence is exact, $Im(h_3)=ker(h_4)=\mathbf{Z}$ and so h_3 is onto. Also $ker(h_2)=Im(h_1)=\{0\}$ and so h_2 is one-to-one. Thus we have the situation shown in Figure 7.

Therefore $H_0(\text{TOUCH}) \simeq IM(h_2)$. The α_i's are the cosets of $Im(h_2)$ and so there is a one-to-one correspondence between any α_i and $Im(h_2)$. Hence $\mathbf{Z}\oplus H_0(\text{NONOVERLAP}) \simeq \mathbf{Z}\oplus Im(h_2) \simeq \mathbf{Z}\oplus H_0(\text{TOUCH})$. Then since $H_0(\text{NONOVERLAP}) \simeq \mathbf{Z}\oplus\ldots\oplus\mathbf{Z}$ (n copies of $\mathbf{Z}$) and $H_0(\text{TOUCH}) \simeq \mathbf{Z}\oplus\ldots\oplus\mathbf{Z}$ (m copies of $\mathbf{Z}$) by Theorem 2.7 it must be that m=n and so $H_0(\text{NONOVERLAP}) \simeq H_0(\text{TOUCH})$. □

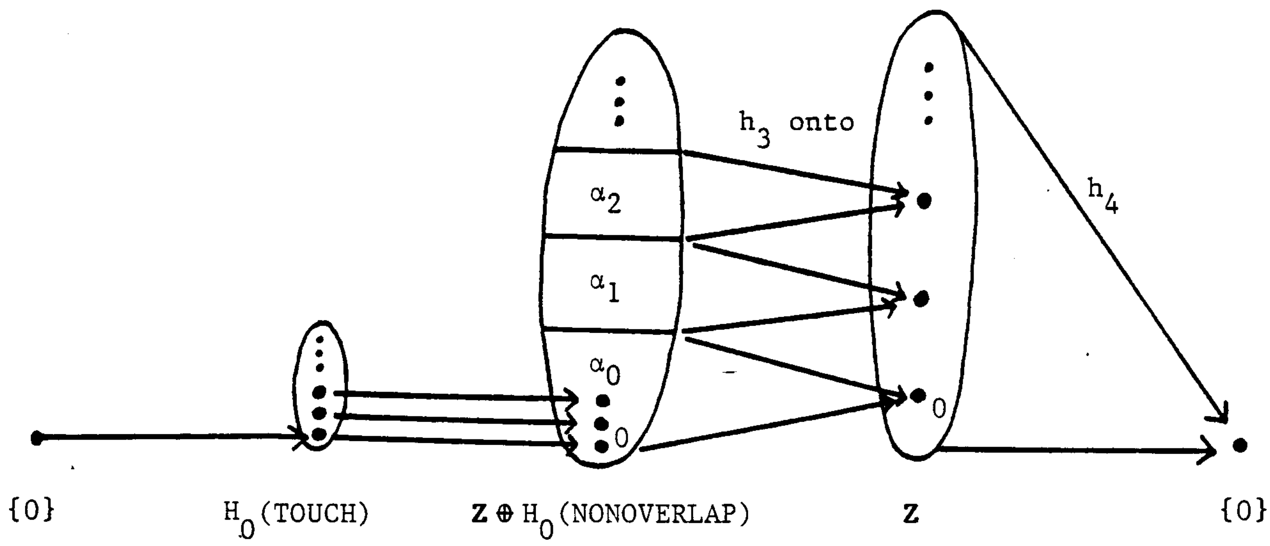

Figure 7. Mayer-Vietoris sequence

Thus it has been shown that the number of path connected components of NONOVERLAP, denoted by #NONOVERLAP, equals the number of path connected components of TOUCH, denoted by #TOUCH. Let m=#TOUCH=#NONOVERLAP and let $t_1,\ldots,t_m$ be the path connected components of TOUCH and $n_1,\ldots,n_m$ be the path connected components of NONOVERLAP.

Lemma 2.10: Each t_i intersects at most one n_k.

Proof: Suppose there is a t_i such that $t_i \cap n_k \neq \phi$ and $t_i \cap n_j \neq \phi$. Let $x_1 \in n_j$, $x_2 \in n_k$, $x_3 \in t_i \cap n_j$ and $x_4 \in t_i \cap n_k$. Then since n_j is path connected there is a path P_1 in n_j from x_1 to x_3. Similarly there is a path P_2 in n_k from x_4 to x_2 and a path P_3 from x_3 to x_4 in t_i. Since TOUCH=FILL∩NONOVERLAP by Theorem 1.6 we get TOUCH⊆NONOVERLAP and so a path in t_i is a path in NONOVERLAP. Thus $P_1 || P_3 || P_2$ is a path from x_1 in n_j to x_2 in n_k and the path is in NONOVERLAP. Hence n_j and n_k must be the same path connected component of NONOVERLAP. □

Thus for each i, $t_i \cap n_k \neq \phi$ for at most one k and since TOUCH$\subseteq$NONOVERLAP we know that for every t_i there is one n_k such that $t_i \subseteq n_k$. Now it will be shown that each n_k contains at most one t_i and so each n_k contains exactly one t_i.

Lemma 2.11: For each n_k there is a t_j such that $t_j \subseteq n_k$.

Proof: We must show that for any $x \in$NONOVERLAP there is motion in NONOVERLAP from x to some configuration $y \in$TOUCH. Let $x \in$NONOVERLAP and suppose $x \in$TOUCH. Let a_1 and a_2 be the origins of A_1 and A_2 respectively and let $s(t)=ta_2(x)+(1-t)a_1(x)$. Let m(t) be the motion such that $a_1(m(t))=s(t)$ and everything else stays constant. Let $t_0=\min_t\{m(t)\notin \text{NONOVERLAP-TOUCH}\}$. Thus $\exists i\, j$ such that $A_i(m(t_0)) \cap A_j(m(t_0)) \neq \phi$. If $a \in A_i(m(t_0)) \cap A_j(m(t_0))$ then $a \in A_i(m(t_0))-\text{int}(A_i(m(t_0)))$ and $a \in A_j(m(t_0))-\text{int}(A_j(m(t_0)))$ otherwise we would contradict the definition of t_0. Therefore $m(t_0) \in$TOUCH as required. □

Theorem 2.12: For each n_k there is exactly one t_j such that $t_j \subseteq n_k$.

Proof: By Lemma 2.11 we know there is at least one $t_j \subseteq n_k$. By Lemma 2.10 each t_i is contained in at most one n_k. Since #NONOVERLAP = #TOUCH we conclude that there is exactly one t_j contained in each n_k. □

We call a motion m of n objects a k-component motion if for all t in the closed interval [0,1], m(t) has at most k connected components. A configuration x is said to be a k-component configuration if x has at most k connected components.

Thus Theorem 2.12 can be restated: If there is an n-component motion from x to y (i.e. $x,y \in n_k$) and x and y are (n-1)-component configurations, then there is an (n-1)-component motion from x to y (i.e. $x,y \in t_j$).

3. The Existence of a Motion in Contact

In this section we establish our main result. Suppose we have n objects and that x and y are two configurations in which the n objects form a connected composite object. Suppose further that there is a motion from configuration x to configuration y such that no two objects overlap. Then there is a motion such that all configurations throughout the motion are also connected.

For the remainder of the section, all configurations will be in NONOVERLAP. Thus if it is said that a configuration x has k components then we mean that x is in NONOVERLAP and x has k components.

Let $P_1, P_2, \ldots P_p$ be the partitionings of the n objects into k+1 or fewer connected components. We say that a configuration x satisfies P_i if the connected component of P_i are contained in the connected components of x. Let x_1 and x_2 be configurations of n composite objects both of which have k, $1 \le k \le n$, or less connected components and let m be a (k+1)-component motion from x_1 to x_2. Partition NONOVERLAP into regions so that all x satisfying a given P_i are in one region. Further partition NONOVERLAP into path connected components. Without loss of generality we assume that m never returns to a path connected component satisfying P_i once it has left it. Let $T_i = \{t \mid m(t) \text{ satisfies } P_i\}$, $1 \le i \le p$. Since m(t) is a continuous function and T_i is the set of all t such that m(t) is a closed region of configuration space, each T_i is a closed set. Furthermore, each T_i is a finite union of closed intervals. Partition the interval [0,1] into a finite number of closed subintervals $J_i = [a_i, a_{i+1}]$, $1 \le i \le t$, such that for each J_i there are k+1 sets of composite objects where each set remains as a connected composite object during the motion m on interval J_i. We can assume that the J_i's are maximal with respect to the above conditions. Thus $m(a_i)$, $1 \le i \le t+1$, is a k-component configuration and the motion m_i, which is m on J_i, is a (k+1)-component motion. Notice that during m_i the composite objects in each of the k+1 sets that define J_i remain as a connected composite object and so we can think of m_i as a motion of k+1 composite objects rather than a motion of n composite objects.

We will use such a partitioning of the interval [0,1] in the following result where we show that it is possible to reduce the number of components during a motion to the number of components in the initial and final configurations if there is a motion during which there is one more component.

Lemma 3.1: Let n be the number of composite objects and let k be such that $1 \le k < n$. If there is a (k+1)-component motion between two k-component configurations then there is a k-component motion between them.

Proof: The proof is by induction n, the number of composite objects.

Base Step: If n=2 then k=n-1 and the result follows from Theorem 2.12.

Induction Step: Assume the result holds when there are less than n composite objects. Suppose we have n composite objects. For k=n-1 the lemma is true by Theorem 2.12.

Let k be such that $1 \le k < n-1$. Partition [0,1] into $J_1,\ldots,J_t$ as above. Then m_i is a (k+1)-component motion of k+1 composite objects between k-component configurations a_i and a_{i+1}. Since $k+1<n$ the induction hypothesis holds for each m_i and so there is a k-component motion m'_i between a_i and a_{i+1}. Thus $m' = m'_1 || m'_2 || \ldots || m'_t$ is a k-component motion between the given k-component configurations. □

Now an immediate corollary to Lemma 3.1 for the case of connected configurations is stated.

Corollary 3.2: If there is a (k+1)-component motion between two connected configurations of composite objects ($1 \le k < n$) then there is a k-component motion between them.

It is now possible to prove the major aim of this paper. Thus we now show that if there is any motion between two connected configurations such that during the motion no two composite objects overlap then there is a motion between the configurations such that all configurations during the motion are connected and no two composite objects overlap.

Theorem 3.3: If there is any motion between two connected configurations x and y then there is a motion between them such that throughout the motion the configurations are connected.

Proof: Suppose there is a (k+1)-component motion between x and y for $k \ge 1$. Then by Corollary 3.2 there is a k-component motion between x and y. Thus by induction there is a 1-component motion between x and y. □

4. Generalizations

In the previous sections we restricted motion to be translation and rotation. There it was shown that if there was a motion between two connected configurations such that throughout the motion no two objects overlapped, then there was a motion between the configurations such that throughout the motion the objects formed a connected configuration. This result depended on the fact that certain subsets of configuration space were path connected and that one subset, namely FILL ∪ NONOVERLAP was contractible to a point. These facts used translational motions in their proofs. Notice that the results hold if the only motions allowed are translations. However as noted earlier if only rotations are permitted then the result does not hold as stated.

For more general motion that allows continuous deformation of the objects such as stretching or radial growth about some point of an object, we must make sure that FILL and FILL ∪ NONOVERLAP are again path connected and FILL ∪ NONOVERLAP is contractible to a point. If so, we can again conclude that if there is a generalized motion, that keeps the objects from overlapping one another, between two connected configurations then there is a generalized motion between these configurations that keeps the configurations connected throughout.

Suppose motions consist of translations and any kind of continuous deformation of objects such that the objects remain convex and the deformation has a fixed point. The motions described in Theorem 2.3 and Theorem 2.5 can be extended in the obvious way to include the type of motion described above. Thus FILL ∪ NONOVERLAP is contractible to a point and FILL is path connected for these motions and hence we can conclude that if there is a motion in NONOVERLAP between two connected configurations then there is a motion that keeps the objects connected throughout.

For some types of generalized motions FILL will not be path connected. Suppose FILL is not path connected but it is the case that the set consisting of all the configurations of one path connected component of FILL and all the configurations in a path connected component of NONOVERLAP which intersects the path connected component of FILL is contractible to a point. Then by taking A to be the path connected component of FILL and B to be the path connected component of NONOVERLAP in the Mayer-Vietoris sequence of section 2 we can conclude that for two configurations which are in both the path connected component of FILL and the path connected component of NONOVERLAP and hence in TOUCH, there is a motion in TOUCH between these configurations.

If we strengthen the definition of VALID such that VALID remains a closed set in configuration space then the results of the previous sections still hold. For example instead of just requiring that a composite object be connected we could insist that the objects of the composite object touch each other in a specific manner. In this way we could have nonconvex objects by dividing them into convex pieces and then defining VALID so that a configuration is in VALID only if the convex pieces form the nonconvex object that is required. An example of this is shown in Figure 8 where the four rectangular objects form a nonconvex composite object and a configuration must have these objects touching in this way for it to be in VALID.

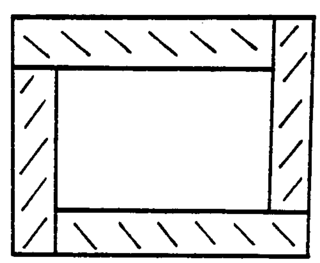

Figure 8. A nonconvex object

Acknowledgements

The authors would like to thank Peter Kahn for pointing out the use of the Mayer-Vietoris sequence.

References

[1] Croom, F. H., "Basic Concepts of Algebraic Topology", Springer-Verlag, New York, 1978.

[2] Goldberg, R. R., "Methods of Real Analysis", John Wiley & Sons, Inc., New York, 1976.

[3] Hopcroft, J. E., J. T. Schwartz and M. Sharir, "On the Complexity of Motion Planning for Multiple Independent Objects; PSpace Hardness of the 'Warehouseman's Problem'", Robotics Research Technical Report No. 14, Computer Science Division, New York University, February 1984.

[4] Hopcroft, J. E., D. Joseph and S. Whitesides, "Movements Problems for 2-Dimensional Linkages", Dept. of Computer Science Technical Report No. TR82-515, Cornell University, 1982. (To appear in the SIAM J. on Computing.)

[5] Massey, W. S., "Homology and Cohomology Theory", Marcel Dekker, Inc., New York, 1978.

[6] Massey, W. S., "Algebraic Topology: An Introduction", Springer-Verlag, New York, 1967.

[7] Schwartz, J. T., and M. Sharir, "On The Piano Mover's Problem II: General Techniques for Computing Topological Properties of Real Algebraic Manifolds", New York University Technical Report No. 41, Computer Science Department, 1982.

[8] Willard, S., "General Topology", Addison-Wesley, Reading, Massachusetts, 1970.