

Figure 3.9 A planar mechanism with six links (A through F), seven revolute joints, and one degree of freedom.

This is known as *Grübler's formula* for closed chains, and it is only valid if the constraints due to the joints are independent. In the planar mechanism of figure 3.9, there are seven joints, each with one degree of freedom, and six links, yielding a mobility of $3(6 - 7 - 1) + 7 = 1$.

As an application of the ideas in this section, determine the number of degrees of freedom of your arm by adding the degrees of freedom at your shoulder, elbow, and wrist. To test your answer, place your palm flat down on a table with your elbow bent. Without moving your torso or your palm, you should find that it is still possible to move your arm. This internal freedom means that your arm is *redundant* with respect to the task of positioning your hand (or a rigid body grasped by your hand) in space; an infinity of arm configurations puts your hand in the same place.⁴ How many internal degrees of freedom do you have? How many holonomic constraints were placed on your arm's configuration when you fixed your hand's position and orientation? The sum of the number of constraints and internal degrees of freedom is the number of degrees of freedom of your (unconstrained) arm, and you should find that your arm has more than six degrees of freedom. A robot is said to be *hyper-redundant* with respect to a task if it has many more degrees of freedom than required for the task. (There is no strict definition of "many" here.)

3.4 The Topology of the Configuration Space

Now that we understand how to determine the dimension of a configuration space, we can begin to explore its topology and geometry, each of which plays a vital role in developing and analyzing motion-planning algorithms. Some basic concepts from topology are discussed in appendixes B and C.

⁴ Provided your arm is away from its joint limits.

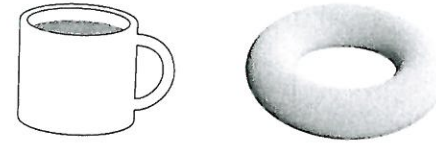


Figure 3.10 The surfaces of the coffee mug and the torus are topologically equivalent.

Topology is a branch of mathematics that considers properties of objects that do not change when the objects are subjected to arbitrary continuous transformations, such as stretching or bending. For this reason, topology is sometimes referred to as "rubber sheet geometry." Imagine a polygon drawn on a rubber sheet. As the sheet is stretched in various directions, the polygon's shape changes; however, certain properties of the polygon do not change. For example, points that are inside the polygon do not move to the outside of the polygon simply because the sheet is stretched.

Two spaces are topologically different if cutting or pasting is required to turn one into the other, as cutting and pasting are not continuous transformations. For example, the configuration spaces of the circular mobile robot (\mathbb{R}^2) and the two-joint planar arm (T^2) are topologically different. If we imagine T^2 as the surface of a rubber doughnut, we see that no matter how we stretch or deform the doughnut (without tearing it), the doughnut will always have a hole in it. Also, if we imagine \mathbb{R}^2 as an infinite rubber sheet, there is no way to stretch it (without tearing it) such that a hole will appear in the sheet. To a topologist, all rubber doughnuts are the same, regardless of how they are stretched or deformed (figure 3.10). Likewise, all rubber sheet versions of \mathbb{R}^2 are the same.

One reason that we care about the topology of configuration space is that it will affect our representation of the space. Another reason is that if we can derive a path-planning algorithm for one kind of topological space, then that algorithm may carry over to other spaces that are topologically equivalent (see, e.g., chapter 4, section 4.6).

Since topology is concerned with properties that are preserved under continuous transformations, we begin our study of the topology of configuration spaces by describing two types of continuous transformations: homeomorphisms and diffeomorphisms. Appendix C provides an introduction to differentiable transformations.

3.4.1 Homeomorphisms and Diffeomorphisms

A mapping $\phi: S \rightarrow T$ is a rule that places elements of S into correspondence with elements of T . We respectively define the *image* of S under ϕ and the *preimage*

CHOSSET

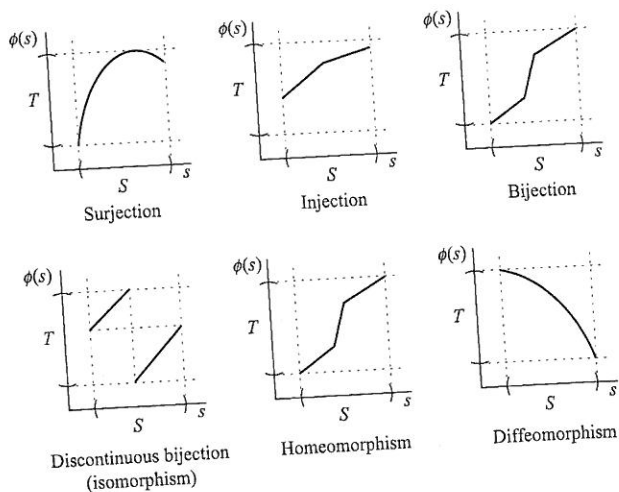


Figure 3.11 Representative ways of looking at surjective, injective, and bijective mappings. Bijections may become homeomorphisms or diffeomorphisms if they are sufficiently differentiable.

of T by

$$\phi(S) = \{\phi(s) \mid s \in S\} \text{ and } \phi^{-1}(T) = \{s \mid \phi(s) \in T\}.$$

If $\phi(S) = T$ (i.e., every element of T is covered by the mapping) then we say that ϕ is *surjective* or *onto*. If ϕ puts each element of T into correspondence with at most one element of S , i.e., for any $t \in T$, $\phi^{-1}(t)$ consists of at most one element in S , then we say that ϕ is *injective* (*one-to-one*). If ϕ is injective, then when $s_1 \neq s_2$ we have $\phi(s_1) \neq \phi(s_2)$ for $s_1, s_2 \in S$. Maps that are both surjective and injective are said to be *bijective*. Figure 3.11 illustrates these definitions. As another example, the map $\sin : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow (-1, 1)$ is bijective, whereas $\sin : \mathbb{R} \rightarrow [-1, 1]$ is only surjective. Bijective maps have the property that their inverse exists at all points in the range T , and thus they allow us to move easily back and forth between the two spaces S and T . In our case, we will use bijective maps to move back and forth between configuration spaces (whose geometry can be quite complicated) and Euclidean spaces.

We will consider two important classes of bijective mappings.

DEFINITION 3.4.1 If $\phi : S \rightarrow T$ is a bijection, and both ϕ and ϕ^{-1} are continuous, then ϕ is a homeomorphism. When such a ϕ exists, S and T are said to be homeomorphic.

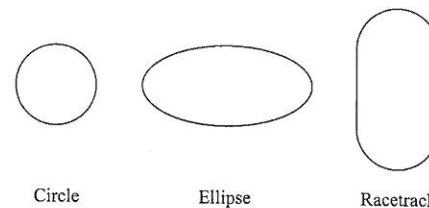


Figure 3.12 A circle, an ellipse, and a racetrack.

A mapping $\phi : U \rightarrow V$ is said to be *smooth* if all partial derivatives of ϕ , of all orders, are well defined (i.e., ϕ is of class C^∞). With the notion of smoothness, we define a second type of bijection.

DEFINITION 3.4.2 A smooth map $\phi : U \rightarrow V$ is a diffeomorphism if ϕ is bijective and ϕ^{-1} is smooth. When such a ϕ exists, U and V are said to be diffeomorphic.

The condition for diffeomorphisms (smoothness) is stronger than the condition for homeomorphisms (continuity), and thus all diffeomorphisms are homeomorphisms.

To illustrate these ideas, consider three one-dimensional surfaces: a circle, denoted by \mathcal{M}_c ; an ellipse, denoted by \mathcal{M}_e ; and a "racetrack," denoted by \mathcal{M}_r . The racetrack consists of two half-circles connected by straight lines (figure 3.12). We define these shapes mathematically as

$$(3.2) \quad \mathcal{M}_c = \{(x, y) \mid f_c(x, y) = x^2 + y^2 - 1 = 0\}$$

$$(3.3) \quad \mathcal{M}_e = \{(x, y) \mid f_e(x, y) = \frac{x^2}{4} + y^2 - 1 = 0\}$$

$$(3.4) \quad \mathcal{M}_r = \{(x, y) \mid f_r(x, y) = 0\}$$

with

$$(3.5) \quad f_r(x, y) = \begin{cases} x - 1 & : -1 \leq y \leq 1, x > 0 \\ (y + 1)^2 + x^2 - 1 & : y < -1 \\ (y - 1)^2 + x^2 - 1 & : y > 1 \\ x + 1 & : -1 \leq y \leq 1, x < 0 \end{cases}$$

Note that these surfaces are implicitly defined as being the set of points that satisfy some equation $f(x, y) = 0$.

In some ways, these three surfaces are similar. For example, they are all simple, closed curves in the plane; all of $f_c(x, y)$, $f_e(x, y)$, and $f_r(x, y)$ are continuous. In other ways, they seem quite different. For example, both $f_c(x, y)$ and $f_e(x, y)$ are

continuously differentiable, while $f_r(x, y)$ is not. We can more precisely state the similarities and differences between these surfaces using the concepts of homeomorphism and diffeomorphism. In particular, it can be shown that \mathcal{M}_c , \mathcal{M}_e , and \mathcal{M}_r are all homeomorphic to each other. For example, the map $\phi: \mathcal{M}_e \rightarrow \mathcal{M}_c$ given by

$$\phi(x, y) = \left[\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}} \right]^T$$

is a homeomorphism.

For this choice of ϕ , both ϕ and ϕ^{-1} are smooth, and therefore, \mathcal{M}_c is diffeomorphic to \mathcal{M}_e . Neither \mathcal{M}_c nor \mathcal{M}_e is diffeomorphic to \mathcal{M}_r , however. This is because $f_r(x, y)$ is not continuously differentiable, while both $f_c(x, y)$ and $f_e(x, y)$ are. This can be seen by examining the curvatures of the circle, ellipse, and racetrack. For the circle, the curvature is constant (and thus continuous), and for the ellipse, curvature is continuous. For the racetrack, there are discontinuities in curvature (at the points $(-1, 1)$, $(-1, -1)$, $(1, 1)$, $(1, -1)$), and therefore there is no smooth mapping from either the circle or the ellipse to the racetrack.

We are often concerned only with the local properties of configuration spaces. Local properties are defined on *neighborhoods*. For metric spaces⁵, neighborhoods are most easily defined in terms of open balls. For a point p of some manifold \mathcal{M} , we define an open ball of radius ϵ by

$$B_\epsilon(p) = \{p' \in \mathcal{M} \mid d(p, p') < \epsilon\},$$

where d is a metric on \mathcal{M} .⁶ A neighborhood of a point $p \in \mathcal{M}$ is any subset $\mathcal{U} \subseteq \mathcal{M}$ with $p \in \mathcal{U}$ such that for every $p' \in \mathcal{U}$, there exists an open ball $B_\epsilon(p') \subset \mathcal{U}$. Any open ball is itself a neighborhood. The open disk in the plane is an example. For the point (x_0, y_0) in the plane, an open ball defined by the Euclidean metric is

$$B_\epsilon(x_0, y_0) = \{(x, y) \mid (x - x_0)^2 + (y - y_0)^2 < \epsilon^2\}.$$

We say that S is *locally diffeomorphic* (resp. *locally homeomorphic*) to T if for each $p \in S$ there exists a diffeomorphism (resp. homeomorphism) f from S to T on some neighborhood \mathcal{U} with $p \in \mathcal{U}$.

The sphere presents a familiar example of these concepts. At any point on the sphere, there exists a neighborhood of that point that is diffeomorphic to the plane. It is not surprising that people once believed the world was flat — they were only looking at their neighborhoods!

5. A metric space is a space equipped with a distance metric. See appendix C.

6. One can define all topological properties, including neighborhoods, without resorting to the use of metrics, but for our purposes, it will be easier to assume a metric on the configuration space and exploit

Let us now reflect on the two examples from the beginning of this chapter. For the circular mobile robot, the workspace and the configuration space are diffeomorphic. This is easy to see, since both are copies of \mathbb{R}^2 . In this case, the identity map $\phi(x) = x$ is a perfectly fine global diffeomorphism between the workspace and configuration space. In contrast, the two-joint manipulator has a configuration space that is T^2 , the torus. The torus T^2 is not diffeomorphic to \mathbb{R}^2 , but it is locally diffeomorphic. If the revolute joints in the two-joint manipulator have lower and upper limits, $\theta_i^l < \theta_i < \theta_i^u$, so that they cannot perform a complete revolution, however, then the configuration space of the two-joint manipulator becomes an open subset of the torus, which is diffeomorphic to \mathbb{R}^2 (globally). This follows from the fact that each joint angle lies in an *open* interval of \mathbb{R}^1 , and we can “stretch” that open interval to cover the line. An example of such a stretching function is $\tan: (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}$.

3.4.2 Differentiable Manifolds

For all of the configuration spaces that we have seen so far, we have been able to uniquely specify a configuration by n parameters, where n is the dimension of the configuration space (two for the planar two-joint arm, three for a polygon in the plane, etc.). The reason that we could do so was that these configuration spaces were all “locally like” n -dimensional Euclidean spaces. Such spaces, called *manifolds*, are a central topic of topology.

DEFINITION 3.4.3 (Manifold) A set S is a k -dimensional manifold if it is locally homeomorphic to \mathbb{R}^k , meaning that each point in S possesses a neighborhood that is homeomorphic to an open set in \mathbb{R}^k .

While a general k -dimensional manifold is locally homeomorphic to \mathbb{R}^k , the configuration spaces that we will consider are locally diffeomorphic to \mathbb{R}^k , an even stronger relationship. In fact, when we parameterized configurations in section 3.1, we were merely constructing diffeomorphisms from configuration spaces to \mathbb{R}^2 . If we construct enough of these diffeomorphisms (so that every configuration in \mathcal{Q} is in the domain of at least one of them), and if these diffeomorphisms are compatible with one another (an idea that we will formalize shortly), then this set of diffeomorphisms together with the configuration space define a *differentiable manifold*. We now make these ideas more precise.

DEFINITION 3.4.4 (Chart) A pair (U, ϕ) , such that U is an open set in a k -dimensional manifold and ϕ is a diffeomorphism from U to some open set in \mathbb{R}^k , is called a chart.

The use of the term *chart* is analogous to its use in cartography, since the subset U is “charted” onto \mathbb{R}^k in much the same way that cartographers chart subsets of the globe onto a plane when creating maps. Charts are sometimes referred to as *coordinate systems* because each point in the set U is assigned a set of coordinates in a Euclidean space [410]. The inverse diffeomorphism, $\phi^{-1} : \mathbb{R}^k \rightarrow U$, is referred to as a *parameterization* of the manifold.

As an example, consider the one-dimensional manifold $S^1 = \{x = (x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$. For any point $x \in S^1$ we can define a neighborhood that is diffeomorphic to \mathbb{R} . For example, consider the upper portion of the circle, $U_1 = \{x \in S^1 \mid x_2 > 0\}$. The chart $\phi_1 : U_1 \rightarrow (0, 1)$ is given by $\phi_1(x) = x_1$, and thus x_1 can be used to define a local coordinate system for the upper semicircle. In the other direction, the upper portion of the circle can be parameterized by $z \in (0, 1) \subset \mathbb{R}$, with $\phi_1^{-1}(z) = (z, (1 - z^2)^{1/2})$, which maps the open unit interval to the upper semicircle. But S^1 is not globally diffeomorphic to \mathbb{R}^1 ; we cannot find a single chart whose domain includes all of S^1 .

We have already used this terminology in section 3.1, when we referred to θ_1, θ_2 as parameters that represent a configuration of the two-joint arm. Recall that $(\theta_1, \theta_2) \in \mathbb{R}^2$, and that when considered as a representation of the configuration, they define a point in T^2 , the configuration space, which is a manifold. We now see that in section 3.1, when we represented a configuration of the planar arm by the pair (θ_1, θ_2) , we were in fact creating a chart from a subset of the configuration space to a subset of \mathbb{R}^2 .

A single mapping from T^2 to \mathbb{R}^2 shown in figure 3.2 encounters continuity problems at $\theta_i = \{0, 2\pi\}$. For many interesting configuration spaces, it will be the case that we cannot construct a single chart whose domain contains the entire configuration space. In these cases, we construct a collection of charts that cover the configuration space. We are not free to choose these charts arbitrarily; any two charts in this collection must be compatible for parts of the manifold on which their domains overlap. Two charts with such compatibility are said to be C^∞ -related (figure 3.13).

DEFINITION 3.4.5 (C^∞ -related) *Let (U, ϕ) and (V, ψ) be two charts on a k -dimensional manifold. Let X be the image of $U \cap V$ under ϕ , and Y be the image of $U \cap V$ under ψ , i.e.,*

$$X = \{\phi(x) \in \mathbb{R}^k \mid x \in U \cap V\}$$

$$Y = \{\psi(y) \in \mathbb{R}^k \mid y \in U \cap V\}.$$

These two charts are said to be C^∞ -related if both of the composite functions

$$\psi \circ \phi^{-1} : X \rightarrow Y,$$

$$\phi \circ \psi^{-1} : Y \rightarrow X$$

are C^∞ .

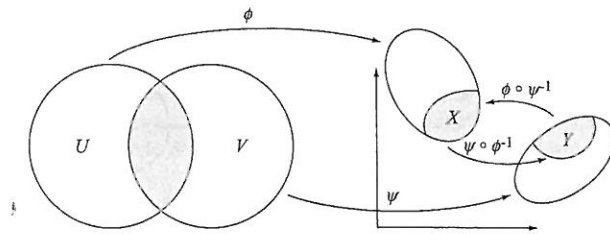


Figure 3.13 The charts (U, ϕ) and (V, ψ) map open sets on the k -dimensional manifold to open sets in \mathbb{R}^k .

If two charts are C^∞ -related, we can switch back and forth between them in a smooth way when their domains overlap. This idea will be made more concrete in the example of S^1 below.

A set of charts that are C^∞ -related, and whose domains cover the entire configuration space \mathcal{Q} , form an *atlas* for \mathcal{Q} . An atlas is sometimes referred to as the differentiable structure for \mathcal{Q} . Together, the atlas and \mathcal{Q} comprise a *differentiable manifold*. There are other ways to define differentiable manifolds, as we will see in section 3.5.

As an example, consider again the one-dimensional manifold S^1 . Above, we defined a single chart, (U_1, ϕ_1) . If we define three more charts, we can construct an atlas for S^1 . These four charts are given by

$$U_1 = \{x \in S^1 \mid x_2 > 0\}, \quad \phi_1(x) = x_1$$

$$U_2 = \{x \in S^1 \mid x_2 < 0\}, \quad \phi_2(x) = x_1$$

$$U_3 = \{x \in S^1 \mid x_1 > 0\}, \quad \phi_3(x) = x_2$$

$$U_4 = \{x \in S^1 \mid x_1 < 0\}, \quad \phi_4(x) = x_2.$$

The corresponding parameterizations are given by $\phi_i^{-1} : (-1, 1) \rightarrow U_i$, with

$$\phi_1^{-1}(z) = (z, 1 - z^2)$$

$$\phi_2^{-1}(z) = (z, z^2 - 1)$$

$$\phi_3^{-1}(z) = (1 - z^2, z)$$

$$\phi_4^{-1}(z) = (z^2 - 1, z).$$

It is clear that the U_i cover S^1 , so to verify that these charts form an atlas it is only necessary to show that they are C^∞ -related (figure 3.14). Note that

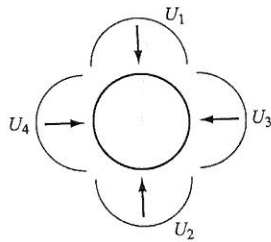


Figure 3.14 Four charts covering the circle S^1 .

$U_1 \cap U_2 = U_3 \cap U_4 = \emptyset$, so we need only check the four pairs of composite maps:

$$\begin{aligned} \phi_1 \circ \phi_3^{-1} &: (0, 1) \rightarrow (0, 1), & \phi_3 \circ \phi_1^{-1} &: (0, 1) \rightarrow (0, 1) \\ \phi_1 \circ \phi_4^{-1} &: (0, 1) \rightarrow (-1, 0), & \phi_4 \circ \phi_1^{-1} &: (-1, 0) \rightarrow (0, 1) \\ \phi_2 \circ \phi_3^{-1} &: (-1, 0) \rightarrow (0, 1), & \phi_3 \circ \phi_2^{-1} &: (0, 1) \rightarrow (-1, 0) \\ \phi_2 \circ \phi_4^{-1} &: (-1, 0) \rightarrow (-1, 0), & \phi_4 \circ \phi_2^{-1} &: (-1, 0) \rightarrow (-1, 0). \end{aligned}$$

In each case, $\phi_i \circ \phi_j^{-1}(z)$ is smooth on each of the open unit intervals that define the domains for the composite mappings given above. For example, $\phi_1 \circ \phi_3^{-1}(z) = \frac{1-z}{1+z}$.

This collection of four charts is not minimal; it is straightforward to find two charts to cover S^1 (see problem 9).

3.4.3 Connectedness and Compactness

We say that a manifold is *path-connected*, or just *connected*, if there exists a path between any two points of the manifold.⁷ All of the obstacle-free configuration spaces that we will consider in this text, e.g., \mathbb{R}^n , S^n , and T^n , are connected. The presence of obstacles, however, can disconnect the free configuration space Q_{free} . In this case, the free configuration space is broken into a set of *connected components*, the maximal connected subspaces. In figure 3.5(c), for example, obstacles break the mobile robot's free configuration space into two connected components. There can be no solution to a motion-planning problem if q_{start} and q_{goal} do not lie in the same connected component of Q_{free} .

⁷ For more general spaces, the concepts of path-connectedness and connectedness are not equivalent, but for a manifold they are the same. More generally, a space is connected if there is no continuous mapping from the space to a discrete set of more than one element.

A space is *compact*⁸ if it resembles a closed, bounded subset of \mathbb{R}^n . A space is closed if it includes all of its limit points. As examples, the half-open interval $[0, 1) \subset \mathbb{R}$ is bounded but not compact, while the closed interval $[0, 1]$ is bounded and compact. The space \mathbb{R}^n is not bounded and therefore not compact. The spaces S^n and T^n are both compact, as they can be expressed as closed and bounded subsets of Euclidean spaces. The unit circle S^1 , e.g., can be expressed as a closed and bounded subset of \mathbb{R}^2 .

In configuration spaces with obstacles or joint limits, the modeling of the obstacles may affect whether the space is compact or not. For a revolute joint subject to joint limits, the set of joint configurations is compact if the joint is allowed to hit the limits, but not compact if the joint can only approach the limits arbitrarily closely.

The product of compact configuration spaces is also compact. For a noncompact space $\mathcal{M}_1 \times \mathcal{M}_2$, if \mathcal{M}_1 is compact, then it is called the *compact factor* of the space. Compact and noncompact spaces cannot be diffeomorphic.

3.4.4 Not All Configuration Spaces Are Manifolds

We are focusing on configuration spaces that are manifolds, and more specifically differentiable manifolds, but it is important to keep in mind that not all configuration spaces are manifolds. As a simple example, the closed unit square $[0, 1] \times [0, 1] \subset \mathbb{R}^2$ is not a manifold, but a *manifold with boundary* obtained by pasting the one-dimensional boundary on the two-dimensional open set $(0, 1) \times (0, 1)$. Also, some parallel mechanisms with one degree of freedom have configurations from which there are two distinct possible motion directions (i.e., the configuration space is a self-intersecting figure eight). It is beyond the scope of this chapter to discuss other types of configuration spaces, but be aware: if you cannot show it to be a manifold, it may not be!

3.5 Embeddings of Manifolds in \mathbb{R}^n

Although a k -dimensional manifold can be represented using as few as k parameters, we have seen above that doing so may require multiple charts. An alternative is to use a representation with "extra" numbers, subject to constraints, to achieve a single global representation. As an example, S^1 is a one-dimensional manifold that we can

⁸ In topology, a space is defined to be compact if every open cover of the space admits a finite subcover, but we will not use these concepts here.