§51 Homotopy of Paths

Definition If f and g are continuous maps of the space X into the space Y, we say that *f is homotopic to f'* if there is a continuous map $F : X \times I \to Y$, where $I = [0, 1]$, such that

$$F(x, 0) = f(x) \quad \text{and} \quad F(x, 1) = f'(x) \quad \text{for each } x \in X.$$

The map F is called a *homotopy* between f and f'. If f is homotopic to f', we write $f \simeq f'$. If $f \simeq f'$ and f is a constant map, we say that f is *nulhomotopic.*

We think of a homotopy as a continuous one-parameter family of maps from X to Y. If we imagine the parameter t as representing time, then the homotopy F represents a continuous “deforming” of the map f to the map f', as t goes from 0 to 1.

Definition If $f : [0, 1] \to X$ is a continuous map such that $f(0) = x_0$ and $f(1) = x_1$, then f is called a *path* in X from the *initial point* x_0 to the *final point* x_1. In the following, we shall for convenience use the interval $I = [0, 1]$ as the domain for all paths.

If f and f' are two paths in X, there is a stronger relation between them than mere homotopy. It is defined as follows:

Definition Two paths f, $f' : I \to X$, mapping the interval $I = [0, 1]]$ into the space X are said to be *path homotopic* if they have the same initial point x_0 and the same final point x_1 and there is a continuous map $F : I \times I \to X$ such that

$$\begin{aligned} F(s, 0) = f(s) \quad &\text{and} \qquad F(s, 1) = f'(s), \\ F(0, t) = x_0 \quad &\text{and} \qquad F(1, t) = x_1, \end{aligned}$$

for each $s \in I$ and each $t \in I$. We call F a *path homotopy* between f and f'. If f is path homotopic to f', we write *$f \simeq_p f'$.*

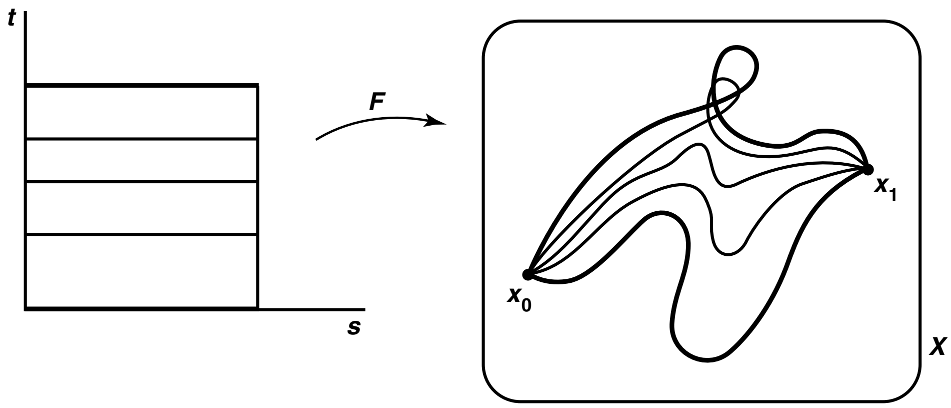


Figure 51.1

The first condition says simply that F is a homotopy between f and f', and the second says that for each t, the path defined by $f_t(s) = F(s, t)$ is a path from x_0 to x_1. Said differently, the first condition says that F represents a continuous way of deforming the path f to the path f', and the second condition says that the end points of the path remain fixed during the deformation.

Lemma 51.1 The relations $\simeq$ and $\simeq_p$ are equivalence relations.

If f is a path, we shall denote its *path-homotopy equivalence class by $[f]$.*

Proof Let us verify the properties of an equivalence relation.

Given f, it is trivial that $f \simeq f$ since the map $F(x,t) = f(x)$ is the required homotopy. If f is a path, F is a path homotopy.

If $f \simeq f'$ and if F is a homotopy between f and f', then $G(x,t) = F(x, 1-t)$ is a homotopy between f' and f, that is $f' \simeq f$. If F is a path homotopy, so is G.

If $f \simeq f'$, $f' \simeq f''$ and if F is a homotopy between f and f', F' is a homotopy between f' and f'', then the map $G : X \times I \to Y$, defined by

$$G(x,t) = \begin{cases} F(x, 2t) & \text{for } t \in \left[0, \frac{1}{2}\right] \\ F'(x, 2t-1) & \text{for } t \in \left[\frac{1}{2}, 1\right] \end{cases}$$

is a homotopy between f and f'', that is $f \simeq f''$. If F and F' are path homotopies, so is G.

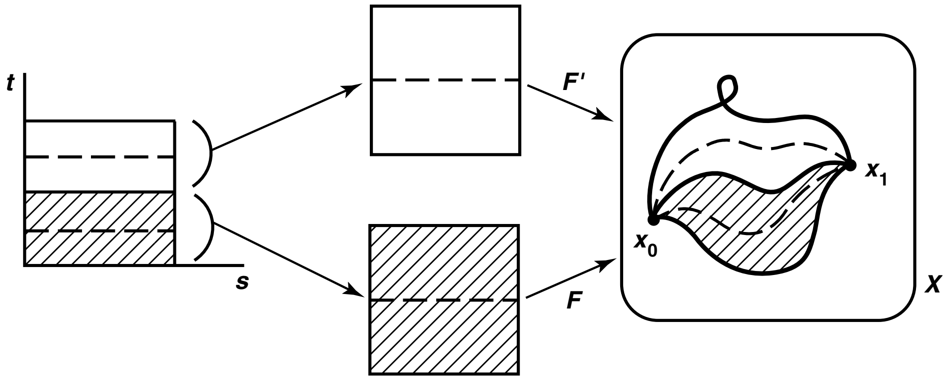


Figure 51.2

In general, we say that f is homotopic to g relative to A (a subset of X) and write $f \simeq g$ rel A if there is a homotopy F from f to g with the additional property that

$$F(a,t) = f(a) \quad \text{for all } a \in A, \text{ for all } t \in I.$$

So, f is path homotopic to g if and only if f is homotopic to g relative to $\{0,1\}$.

Examples

1. Let f and g be any two maps of a space X into $\mathbb{R}^2$. It is easy to see that f and g are homotopic; the map

   $$F(x,t) = (1-t)f(x) + tg(x) \quad \text{for each } x \in X,\ t \in I = [0,1]$$

   is a homotopy between them. It is called a *straight-line homotopy* because it moves the point $f(x)$ to the point $g(x)$ along the straight-line segment joining them.

   If f and g are paths from x_0 to x_1, then F will be a path homotopy, as you can check. This situation is pictured in Figure 51.3.

2. Let X denote the *punctured plane*, $\mathbb{R}^2 \setminus \{0\}$. The following paths in X,

   $$f(s) = (\cos \pi s, \sin \pi s), \quad \text{and} \quad g(s) = (\cos \pi s, 2 \sin \pi s)$$

   are path homotopic; the straight-line homotopy between them is an acceptable path homotopy. But the straight-line homotopy between f and the path

   $$h(s) = (\cos \pi s, -\sin \pi s)$$

   is not acceptable, for its image does not lie in the space $X = \mathbb{R}^2 \setminus \{0\}$. See Figure 51.4.

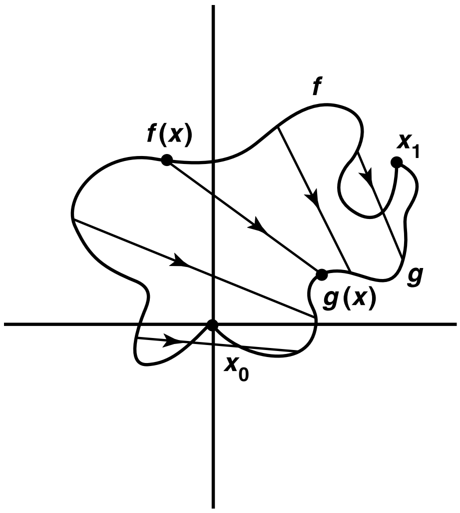


Figure 51.3

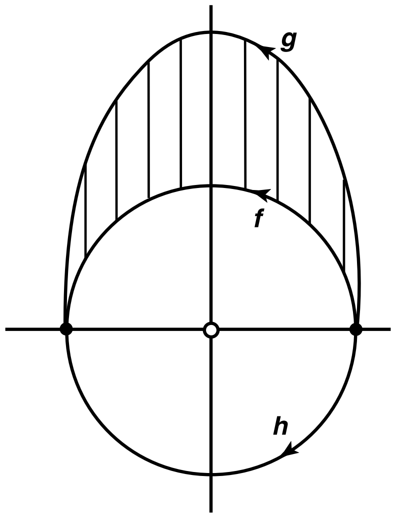


Figure 51.4

3. Let X be a topological space, $\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid \|x\|^2 = \sum_{i=1}^{n+1} x_i^2 = 1\}$ be the unit sphere in $\mathbb{R}^{n+1}$, and let $f, g : X \to \mathbb{S}^n \subset \mathbb{R}^{n+1}$ be continuous maps such that $f(x) + g(x) \neq 0 \in \mathbb{R}^n$ for all $x \in X$. Then the map $F : X \times I \to \mathbb{S}^n$ defined by
$$F(x,t) = \frac{(1-t)f(x) + tg(x)}{\|(1-t)f(x) + tg(x)\|}$$
is a homotopy from f to g.

Theorem 18.3 (Pasting Lemma) Let $f : X \to Y$ be a function mapping space X to space Y. Let $A \cup B = X$, where A and B are both open or both closed subsets of X. If $f|_A$ and $f|_B$ are both continuous, then f is continuous.

Proof Let U be an open subset of Y. We have that $f|_A^{-1}(U)$ is open in A and $f|_B^{-1}(U)$ is open in B. There exist open sets V_A and V_B of X such that $f|_A^{-1}(U) = A \cap V_A$ and $f|_B^{-1}(U) = B \cap V_B$. Since
$$f^{-1}(U) = f|_A^{-1}(U) \cup f|_B^{-1}(U) = (A \cap V_A) \cup (B \cap V_B),$$
and if both A and B are open subsets of X, then $f^{-1}(U) = (A \cap V_A) \cup (B \cap V_B)$ is open in X and f is continuous. It is also easy to show that f is continuous if sets A and B are closed.

Definition If f is a path in X from x_0 to x_1, and g is a path in X from x_1 to x_2, we define the *product $f * g$* of f and g to be the path h given by the equations
$$h(s) = \begin{cases} f(2s) & \text{for } s \in \left[0, \frac{1}{2}\right] \\ g(2s-1) & \text{for } s \in \left[\frac{1}{2}, 1\right] \end{cases}$$
The function h is well defined and continuous and it is a path in X from x_0 to x_2. We think of h as the path whose first half is the path f and whose second half is the path g.

The product operation on paths induces a well-defined operation on path-homotopy classes, defined by the equation
$$[f] * [g] = [f * g].$$

To verify this fact, let F be a path homotopy between f and f' and let G be a path homotopy between g and g'. Define

$$H(s,t) = \begin{cases} F(2s,t) & \text{for } s \in \left[0, \frac{1}{2}\right], \\ G(2s-1,t) & \text{for } s \in \left[\frac{1}{2}, 1\right]. \end{cases}$$

Because $F(1, t) = x_1 = G(0, t)$ for all t, the map H is well-defined; it is continuous by the pasting lemma. You can check that H is the required path homotopy between $f * g$ and $f' * g'$. It is pictured in Figure 51.5.

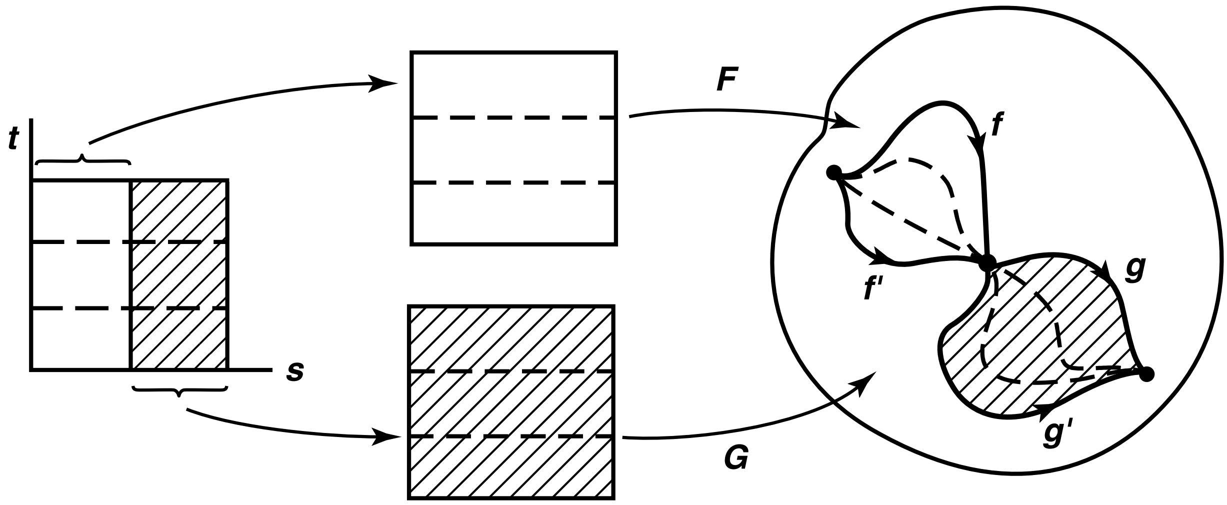


Figure 51.5

The operation $*$ on path-homotopy classes turns out to satisfy properties that look very much like the axioms for a group. They are called the *groupoid properties of* $*$. One difference from the properties of a group is that $[f] * [g]$ is not defined for every pair of classes, but only for those pairs $[f]$, $[g]$ for which $f(1) = g(0)$.

Theorem 51.2 The operation $*$ is well-defined on path-homotopy classes. It has the following properties:

(1) (Associativity) If $[f] * ([g] * [h])$ is defined, so is $([f] * [g]) * [h]$ and they are equal.

(2) (Right and left identities) Given $x \in X$, let e_x denote the constant path $e_x : I \to X$ carrying all of I to the point x. If f is a path in X from x_0 to x_1, then

$$[f] * [e_{x_1}] = [f] \quad \text{and} \quad [e_{x_0}] * [f] = [f].$$

(3) (Inverse) Given a path f in X from x_0 to x_1, let $\bar{f}$ be the path defined by $\bar{f}(s) = f(1-s)$. It is called the reverse of f. Then

$$[f] * [\bar{f}] = [e_{x_0}] \quad \text{and} \quad [\bar{f}] * [f] = [e_{x_1}].$$

Proof We shall make use of two elementary facts. The first is the fact that if $k : X \to Y$ is a continuous map, and if F is a path homotopy in X between the paths f and f', then $k \circ F$ is a path homotopy in Y between the paths $k \circ f$ and $k \circ f'$. See Figure 51.6.

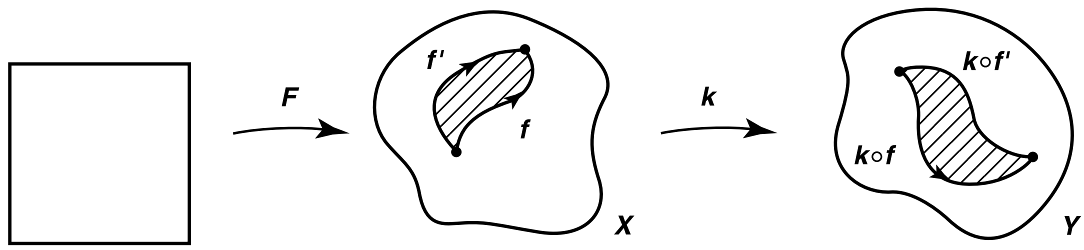


Figure 51.6

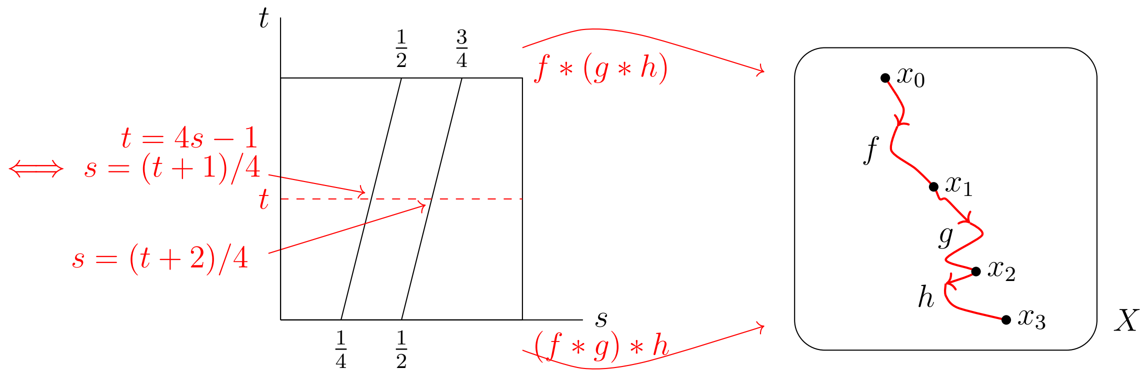
The second is the fact that if $k : X \to Y$ is a continuous map and if f and g are paths in X with $f(1) = g(0)$, then

$$k \circ (f * g) = (k \circ f) * (k \circ g)\,.$$

This equation follows at once from the definition of the product operation $*$.

Associativity : For each $t \in [0,\, 1]$, since the map

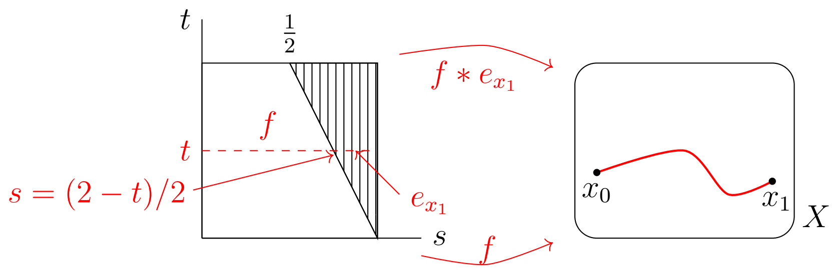
$$F(s,t) = \begin{cases} f(4s/(t+1)) & \text{for } s \in [0, (t+1)/4], \\ g(4s-t-1) & \text{for } s \in [(t+1)/4, (t+2)/4], \\ h((4s-t-2)/(2-t)) & \text{for } s \in [(t+2)/4, 1] \end{cases}$$



is a path homotopy between $(f * g) * h$ and $f * (g * h)$, $([f] * [g]) * [h] = [f] * ([g] * [h])$ and the product $*$ is associative on path-homotopy classes.

Right and left identities : For each $t \in [0, 1]$, since the map

$$G(s,t) = \begin{cases} f(2s/(2-t)) & \text{for } s \in [0, (2-t)/2] \quad (\text{from } x_0 \text{ to } x_1) \\ e_{x_1}(s) = x_1 & \text{for } s \in [(2-t)/2, 1] \quad (\text{from } x_1 \text{ to } x_1) \end{cases}$$

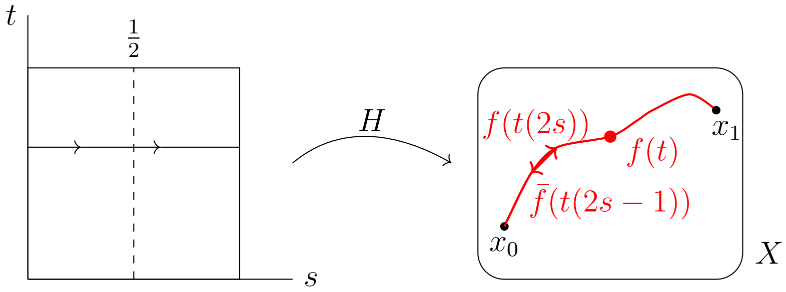


is a path hompotpy between f and $f * e_{x_1}$, $f \simeq_p f * e_{x_1}$ and $[f] = [f] * [e_{x_1}]$

A similar argument could be used to show that $[e_{x_0}] * [f] = [f]$.

Inverse : For each $t \in [0, 1]$, since the map

$$H(s,t) = \begin{cases} f(t(2s)) = f(2ts) & \text{for } s \in [0,1/2] \quad (\text{from } x_0 \text{ to } f(t)) \\ \bar{f}(t(2s-1)) = f(t - t(2s-1)) = f(2t(1-s)) & \text{for } s \in [1/2,1] \quad (\text{from } f(t) \text{ to } x_0) \end{cases}$$



is a path hompotpy between e_{x_0} and $f * \bar{f}$, we have $[e_{x_0}] = [f] * [\bar{f}]$.

A similar argument could be used to show that $[\bar{f}] * [f] = [e_{x_1}]$. But better yet, note the following: We have shown that for any path g, we have $g * \bar{g} \simeq_p e_x$, where x is the initial point of g. In particular, $\bar{f} * \bar{\bar{f}} \simeq_p e_{x_1}$, where $\bar{\bar{f}}$ is the reverse of $\bar{f}$. But the reverse of $\bar{f}$ is just f. Thus $\bar{f} * f \simeq_p e_{x_1}$, as desired.

Theorem 51.3 Let f be a path in X, and let $a_0, \ldots, a_n$ be numbers such that $0 = a_0 < a_1 < \cdots < a_n = 1$. Let $f_i : I \to X$ be the path that equals the positive linear map of I onto $[a_{i-1}, a_i]$ followed by f, i.e. $f_i(s) = f(a_{i-1} + (a_i - a_{i-1})\, s)$ for $s \in [0, 1]$, $1 \le i \le n$. Then

$$[f] = [f_1] * \cdots * [f_n].$$

§52 The Fundamental Groups

Let us first review some terminology from group theory. Suppose G and G' groups, written multiplicatively. A *homomorphism* $f : G \to G'$ is a map such that $f(x \cdot y) = f(x) \cdot f(y)$ for all x, y; it automatically satisfies the equations $f(e) = e'$ and $f(x^{-1}) = f(x)^{-1}$, where e and e' are the identities of G and G', respectively, and the exponent -1 denotes the inverse. The *kernel* of f is the set $f^{-1}(e')$; it is a subgroup of G. The image of f, similarly, is a subgroup of G. The homomorphism f is called a *monomorphism* if it is injective (or equivalently, if the kernel of f consists of e alone). It is called an *epimorphism* if it is surjective; and it is called an *isomorphism* if it is bijective.

Suppose G is a group and H is a subgroup of G. Let xH denote the set of all products xh, for $h \in H$; it is called a *left coset* of H in G. The collection of all such cosets forms a partition of G. Similarly, the collection of all right cosets Hx of H in G forms a partition of G. We call H a *normal subgroup* of G if $x \cdot h \cdot x^{-1} \in H$ for each $x \in G$ and each $h \in H$. In this case, we have $xH = Hx$ for each x, so that our two partitions of G are the same. We denote this partition by G/H; if one defines

$$(xH) \cdot (yH) = (x \cdot y)\, H,$$

one obtains a well-defined operation on G/H that makes it a group. This group is called the *quotient* of G by H. The map $f : G \to G/H$ carrying x to xH is an epimorphism with kernel H. Conversely, if $f : G \to G'$ is an epimorphism, then its kernel N is a normal subgroup of G, and f induces an isomorphism $G/N \to G'$ that carries xN to $f(x)$ for each $x \in G$.

If the subgroup H of G is not normal, it will still be convenient to use the symbol G/H; we will use it to denote the collection of right cosets of H in G.

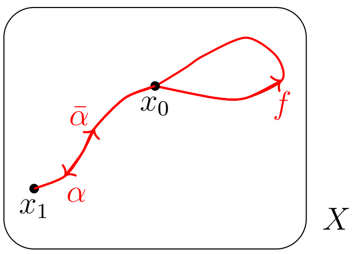
Now we define the fundamental group.

Definition Let X be a space; let x_0 be a point of X. A path in X that begins and ends at x_0 is called a *loop* based at x_0. The set of path homotopy classes of loops based at x_0, with the operation $*$, is called the *fundamental group* of X relative to the *base point* x_0. It is denoted by $\pi_1(X, x_0)$.

Examples

1. Let $\mathbb{R}^n$ denote the euclidean n-space. Then $\pi_1(\mathbb{R}^n, x_0) = \{e_{x_0}\}$, i.e $\pi_1(\mathbb{R}^n, x_0)$ is the trivial group (the consisting of identity alone). For if f is a loop in $\mathbb{R}^n$ based at x_0, the straight-line homotopy
$$F(s, t) = tx_0 + (1 - t)f(s)$$
is a path homotopy between f and the constant loop e_{x_0}.
2. More generally, if X is any convex subset of $\mathbb{R}^n$, then $\pi_1(X, x_0)$ is the trivial group. The straight-line homotopy will work once again, for convexity of X means that for any x, $y \in X$, the straight-line segment
$$\{tx + (1 - t)y \mid 0 \le t \le 1\}$$
between them lies in X. In particular, the unit ball B^n in $\mathbb{R}^n$,
$$B^n = \{x \mid x_1^2 + \cdots + x_n^2 \le 1\},$$
has trivial fundamental group.

Definition Let α be a path in X from x_0 to x_1. We define a map
$$\widehat{\alpha} : \pi_1(X, x_0) \to \pi_1(X, x_1)$$
by the equation
$$\widehat{\alpha}([f]) = [\bar{\alpha}] * [f] * [\alpha]$$



The map $\widehat{\alpha}$ is well-defined because the operation $*$ is well defined. If f is a loop based at x_0, then $\bar{\alpha} * (f * \alpha)$ is a loop based at x_1. Hence $\widehat{\alpha}$ maps $\pi_1(X, x_0)$ into $\pi_1(X, x_1)$, as desired.

Theorem 52.1 The map $\hat{\alpha}$ is a group isomorphism.

Proof For any $[f]$, $[g] \in \pi_1(X, x_0)$, since

$$\begin{aligned} \widehat{\alpha}([f]) * \widehat{\alpha}([g]) &= ([\bar{\alpha}] * [f] * [\alpha]) * ([\bar{\alpha}] * [g] * [\alpha]) \\ &= ([\bar{\alpha}] * [f] * [g] * [\alpha]) \quad \text{since } \alpha * \bar{\alpha} \simeq_p e_{x_0} \text{ and } e_{x_0} * g \simeq_p g \\ &= \widehat{\alpha}([f] * [g]) \end{aligned}$$

the map $\widehat{\alpha} : \pi_1(X, x_0) \to \pi_1(X, x_1)$ is a homomorphism.

Let β denote the path $\bar{\alpha}$, i.e. the reverse of α. For each $[h] \in \pi_1(X, x_1)$ and $[f] \in \pi_1(X, x_0)$, since

$$\widehat{\beta}([h]) = [\bar{\beta}] * [h] * [\beta] = [\alpha] * [h] * [\bar{\alpha}] \implies \widehat{\alpha}(\widehat{\beta}([h])) = [\bar{\alpha}] * ([\alpha] * [h] * [\bar{\alpha}]) * [\alpha] = [h],$$

and

$$\widehat{\beta}(\widehat{\alpha}([f])) = [\bar{\beta}] * ([\bar{\alpha}] * [f] * [\alpha]) * [\beta] = [\bar{\beta}] * ([\beta] * [f] * [\bar{\beta}]) * [\beta] = [f],$$

the map $\widehat{\alpha} : \pi_1(X, x_0) \to \pi_1(X, x_1)$ is an isomorphism.

Corollary 52.2 If X is path connected and x_0 and x_1 are two points of X, then $\pi_1(X, x_0)$ is isomorphic to $\pi_1(X, x_1)$.

Remark If X is path connected, all the groups $\pi_1(X, x)$ are isomorphic, so it is tempting to try to "identify" all these groups with one another, and to speak simply of the fundamental group of the space X, without reference to base point. The difficulty with this approach is that there is no natural way of identifying $\pi_1(X, x_0)$ with $\pi_1(X, x_1)$; different paths α and β from x_0 to x_1 may give rise to different isomorphisms between these groups. For this reason, omitting the base point can lead to error.

Theorem The isomorphism of $\pi_1(X, x_0)$ with $\pi_1(X, x_1)$ is independent of path if and only if the fundamental group $\pi_1(X, x_0)$ is abelian.

Proof Let α and β be two paths from x_0 to x_1. For each $[f] \in \pi_1(X, x_0)$, since

$$\widehat{\alpha}([f]) = [\bar{\alpha}] * [f] * [\alpha] \quad \text{and} \quad \widehat{\beta}([f]) = [\bar{\beta}] * [f] * [\beta]$$

we have

$$\widehat{\beta}^{-1}(\widehat{\alpha}([f])) = [\beta] * [\bar{\alpha}] * [f] * [\alpha] * [\bar{\beta}] = [\beta * \bar{\alpha}] * [f] * [\alpha * \bar{\beta}],$$

where note that $[\beta * \bar{\alpha}] \in \pi_1(X, x_0)$. Hence, the isomorphism of $\pi_1(X, x_0)$ with $\pi_1(X, x_1)$ is independent of path if for any $[\beta * \bar{\alpha}]$, $[f] \in \pi_1(X, x_0)$,

$$\widehat{\alpha}([f]) = \widehat{\beta}([f]) \iff \widehat{\beta}^{-1}(\widehat{\alpha}([f])) = [f] \iff [\beta * \bar{\alpha}] * [f] * [\alpha * \bar{\beta}] = [f] \iff [\beta * \bar{\alpha}] * [f] = [f] * [\beta * \bar{\alpha}]$$

i.e. $\pi_1(X, x_0)$ is abelian.

Definition A space X is said to be *simply connected* if it is a path-connected space and if $\pi_1(X, x_0)$ is the trivial (one-element) group for some $x_0 \in X$, and hence for every $x_0 \in X$.

We often express the fact that $\pi_1(X, x_0)$ is the trivial group by writing $\pi_1(X, x_0) = 0$.

Lemma 52.3 In a simply connected space X, any two paths having the same initial and final points are path homotopic.

Proof Let f and g be two paths from x_0 to x_1. Then $f * \bar{g}$ is defined and is a loop on X based at x_0. Since X is simply connected, $f * \bar{g} \simeq_p e_{x_0}$. Applying the groupoid properties, we see that

$$[(f * \bar{g}) * g] = [e_{x_0} * g] = [g].$$

But

$$[(f * \bar{g}) * g] = [f * (\bar{g} * g)] = [f * e_{x_1}] = [f].$$

Thus f and g are path homotopic.

Definition Let $h : (X, x_0) \to (Y, y_0)$ be a continuous map that carries the point x_0 of X to the point y_0 of Y. Define

$$h_* : \pi_1(X, x_0) \to \pi_1(Y, y_0)$$

by the equation

$$h_*([f]) = [h \circ f] \quad \text{for each loop } f \text{ in } X \text{ based at } x_0.$$

The map h_* is called the *homomorphism induced by h,* relative to the base point x_0.

Remark It is easy to see that h_* is well defined. If f and f' are path homotopic, and $F : I \times I \to X$ is the path homotopy between them, then $h \circ F$ is a path homotopy between the loops $h \circ f$ and $h \circ f'$.

It is easy to check that h_* is a homomorphism. Since

$$(f * g)(s) = \begin{cases} f(2s) & \text{for } s \in \left[0, \frac{1}{2}\right] \\ g(2s - 1) & \text{for } s \in \left[\frac{1}{2}, 1\right] \end{cases}$$

it follows that

$$h(f * g)(s) = \begin{cases} h(f(2s)) & \text{for } s \in \left[0, \frac{1}{2}\right] \\ h(g(2s - 1)) & \text{for } s \in \left[\frac{1}{2}, 1\right] \end{cases} = \begin{cases} (h \circ f)(2s) & \text{for } s \in \left[0, \frac{1}{2}\right] \\ (h \circ g)(2s - 1)) & \text{for } s \in \left[\frac{1}{2}, 1\right] \end{cases}$$

i.e. $h(f * g) = (h \circ f) * (h \circ g)$ which implies that

$$h_*([f] * [g]) = h_*([f * g]) = [h(f * g)] = [(h \circ f) * (h \circ g)] = h_*([f]) * h_*([g]),$$

i.e. h_* is a homomorphism.

Theorem 52.4 If $h : (X, x_0) \to (Y, y_0)$ and $k : (Y, y_0) \to (Z, z_0)$ are continuous maps carry respectively $x_0 \in X$ to $y_0 \in Y$ and $y_0 \in Y$ to $z_0 \in Z$, then $(k \circ h)_* = k_* \circ h_*$. If $i : (X, x_0) \to (X, x_0)$ is the identity map, then i_* is the identity homomorphism.

Proof The proof is a triviality. By definition,

$$\begin{aligned} (k \circ h)_*([f]) &= [(k \circ h) \circ f] \\ (k_* \circ h_*)([f]) &= k_*(h_*([f])) = k_*([h \circ f]) = [k \circ (h \circ f)] \end{aligned}$$

Similarly, $i_*([f]) = [i \circ f] = [f]$.

Corollary 52.5 If $h : (X, x_0) \to (Y, y_0)$ is a homeomorphism of X with Y carries $x_0 \in X$ to $y_0 \in Y$, then h_* is an isomorphism of $\pi_1(X, x_0)$ with $\pi_1(Y, y_0)$.

Proof Let $k : (Y, y_0) \to (X, x_0)$ be the inverse of h. Then

$$k_* \circ h_* = (k \circ h)_* = i_*,$$

where i is the identity map of (X, x_0); and

$$h_* \circ k_* = (h \circ k)_* = j_*,$$

where j is the identity map of (Y, y_0). Since i_* and j_* are the identity homomorphisms of the groups $\pi_1(X, x_0)$ and $\pi_1(Y, y_0)$, respectively, k_* is the inverse of h_*.

§53 Covering Spaces

Definition Let $p : E \to B$ be a continuous surjective map. The open set U of B is said to be *evenly covered* by p if the inverse image $p^{-1}(U) = \bigcup_{\alpha \in J} V_\alpha$ is the union of disjoint open sets V_α in E such that for each α, the restriction of p to V_α is a homeomorphism of V_α onto U. The collection $\{V_\alpha\}$ will be called a partition of $p^{-1}(U)$ into *slices*.

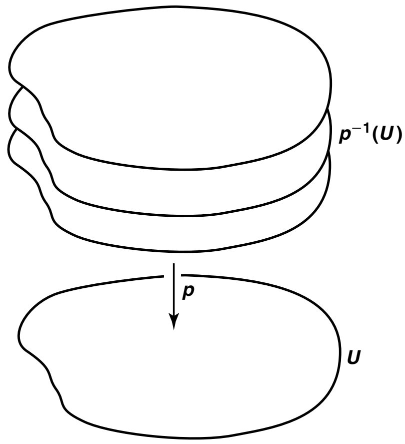


Figure 53.1

Definition Let $p : E \to B$ be continuous and surjective. If every point b of B has an open neighborhood U_b that is evenly covered by p, then p is called a *covering map*, and E is said to be a *covering space* of B.

Remark Note that if $p : E \to B$ is a covering map, then for each $b \in B$ the subset $p^{-1}(b)$ of E necessarily has the discrete topology. For each slice V_α is open in E and intersects $p^{-1}(b)$ in a single point; therefore this point is open in the subspace topology on $p^{-1}(b)$.

Exmaple 1. Let X be any space; let $i : X \to X$ be the identity map. Then i is a covering map (of the most trivial sort). More generally let E be the space $X \times \{1, \dots, n\}$ consisting of n disjoint copies of X. The map $p : E \to X$ given by $p(x, i) = x$ for all $1 \leq i \leq n$ is again a (rather trivial) covering map.

Theorem 53.1 The map $p : \mathbb{R} \to S^1$ given by the equation

$$p(x) = (\cos 2\pi x, \sin 2\pi x)$$

is a covering map.

One can picture p as a function that wraps the real line $\mathbb{R}$ around the circle S^1, and in the process maps each interval $[n, n+1]$ onto S^1.

Proof The fact that p is a covering map comes from elementary properties of the sine and cosine functions. Consider, for example, the subset U of S^1 consisting of those points having positive first coordinate. The set $p^{-1}(U)$ consists of those points x for which $\cos 2\pi x$ is positive; that is, it is the union of the intervals

$$V_n = \left(n - \frac{1}{4}, n + \frac{1}{4}\right),$$

for all $n \in \mathbb{Z}$. See Figure 53.2. Now, restricted to any closed interval $\bar{V}_n$, the map p is injective because $\sin 2\pi x$ is strictly monotonic on such an interval. Furthermore, p carries $\bar{V}_n$ surjectively onto $\bar{U}$, and V_n, to U, by the intermediate value theorem. Since $\bar{V}_n$ is compact, $p|_{\bar{V}_n}$ is a homeomorphism of $\bar{V}_n$ with $\bar{U}$. In particular, $p|_{\bar{V}_n}$ is a homeomorphism of V_n with U.

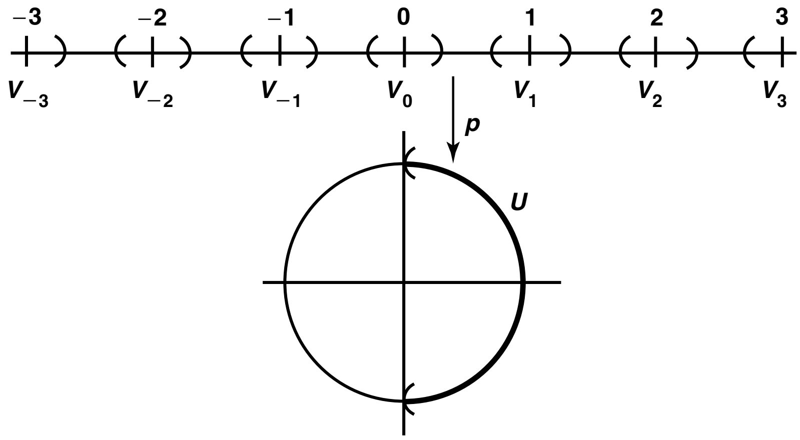


Figure 53.2

Similar arguments can be applied to the intersections of S^1 with the upper and lower open half-planes, and with the open left-hand half-plane. These open sets cover S^1, and each of them is evenly covered by p. Hence $p : \mathbb{R} \to S^1$ is a covering map.

If $p : E \to B$ is a covering map, then p is a *local homeomorphism* of E with B. That is, each point e of E has a neighborhood that is mapped homeomorphically by p onto an open subset of B. The condition that p be a local homeomorphism does not suffice, however, to ensure that p is a covering map, as the following example shows.

Exmaple 2. The map $p : \mathbb{R}_+ \to \mathbb{S}^1$ given by the equation

$$p(x) = (\cos 2\pi x, \sin 2\pi x)$$

is surjective, and it is a local homeomorphism. See Figure 53.3. But it is not a covering map, for the point $b_0 = (1, 0)$ has no neighborhood U that is evenly covered by p. The typical neighborhood U of b_0 has an inverse image consisting of small neighborhoods V_n of each integer n for $n > 0$, along with a small interval V_0 of the form $(0, \varepsilon)$. Each of the intervals V_n for $n > 0$ is mapped homeomorphically onto U by the map p, but the interval V_0 is only imbedded in U by p.

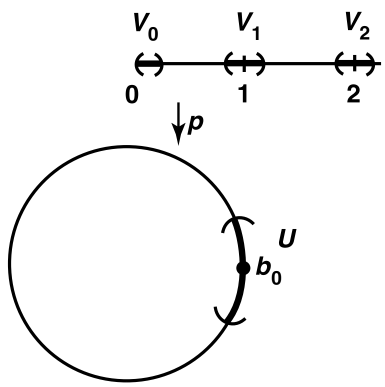


Figure 53.3

Exmaple 3. The preceding example might lead you to think that the real line $\mathbb{R}$ is the only connected covering space of the circle S^1. This is not so. Consider, for example, the map $p : S^1 \to S^1$ given in equations by

$$p(z) = z^2.$$

[Here we consider S^1 as the subset of the complex plane $\mathbb{C}$ consisting of those complex numbers z with $|z| = 1$.] We leave it to you to check that p is a covering map.

Example 2 shows that the map obtained by restricting a covering map may not be a covering map. Here is one situation where it will be a covering map:

Theorem 53.2 Let $p : E \to B$ be a covering map. If B_0 is a subspace of B, and if $E_0 = p^{-1}(B_0)$, then the map $p_0 : E_0 \to B_0$ obtained by restricting p is a covering map.

Proof Given $b_0 \in B_0$, let U be an open set in B containing b_0 that is evenly covered by p; let $\{V_\alpha\}$ be a partition of $p^{-1}(U)$ into slices. Then $U \cap B_0$ is a neighborhood of b_0 in B_0, and the sets $V_\alpha \cap E_0$ are disjoint open sets in E_0 whose union is $p^{-1}(U \cap B_0)$, and each is mapped homeomorphically onto $U \cap B_0$ by p.

Theorem 53.3 If $p : E \to B$ and $p' : E' \to B'$ are covering maps, then

$$p \times p' : E \times E' \to B \times B'$$

is a covering map.

Proof Given $b \in B$ and $b' \in B'$, let U and U' be open neighborhoods of b and b', respectively, that are evenly covered by p and p', respectively. Let $\{V_\alpha\}$ and $\{V'_\beta\}$ be partitions of $p^{-1}(U)$ and $p^{-1}(U')$, respectively, into slices. Then the inverse image under $p \times p'$ of the open set $U \times U'$ is the union of all the sets $V_\alpha \times V'_\beta$. These are disjoint open sets of $E \times E'$, and each is mapped homeomorphically onto $U \times U'$ by $p \times p'$.

Exmaple 4. Consider the space $T = \mathbb{S}^1 \times \mathbb{S}^1$; it is called the torus. The product map

$$p \times p : \mathbb{R} \times \mathbb{R} \to \mathbb{S}^1 \times \mathbb{S}^1$$

is a covering of the torus by the plane $\mathbb{R}^2$, where p denotes the covering map of Theorem 53.1. Each of the unit squares $[n, n+1] \times [m, m+1]$ gets wrapped by $p \times p$ entirely around the torus. See Figure 53.4.

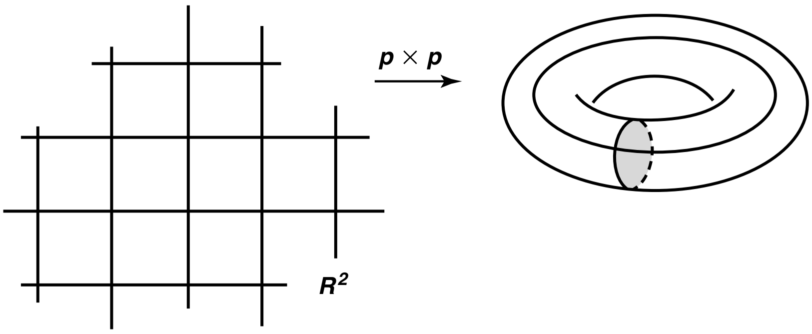


Figure 53.4

§54 The Fundamental Group of the Circle

Definition Let $p : E \to B$ be a map. If f is a continuous mapping of some space X into B, a *lifting* of f is a map $\tilde{f} : X \to E$ such that $p \circ \tilde{f} = f$.

$$\begin{tikzcd} & E \arrow[d, "p"] \\ X \arrow[ur, "\tilde{f}"] \arrow[r, "f"] & B \end{tikzcd}$$

Example Consider the covering $p : \mathbb{R} \to \mathbb{S}^1$ defined by the equation

$$p(x) = (\cos 2\pi x, \sin 2\pi x).$$

The path $f : [0, 1] \to \mathbb{S}^1$ beginning at $b_0 = (1, 0)$ given by $f(s) = (\cos \pi s, \sin \pi s)$ lifts to the path $\tilde{f}(s) = \dfrac{s}{2}$ beginning at 0 and ending at $\dfrac{1}{2}$.

The path $g(s) = (\cos \pi s, -\sin \pi s)$ lifts to the path $\tilde{g}(s) = -\dfrac{s}{2}$ beginning at 0 and ending at $-\dfrac{1}{2}$.

The path $h(s) = (\cos 4\pi s, \sin 4\pi s)$ lifts to the path $\tilde{h}(s) = 2s$ beginning at 0 and ending at 2.

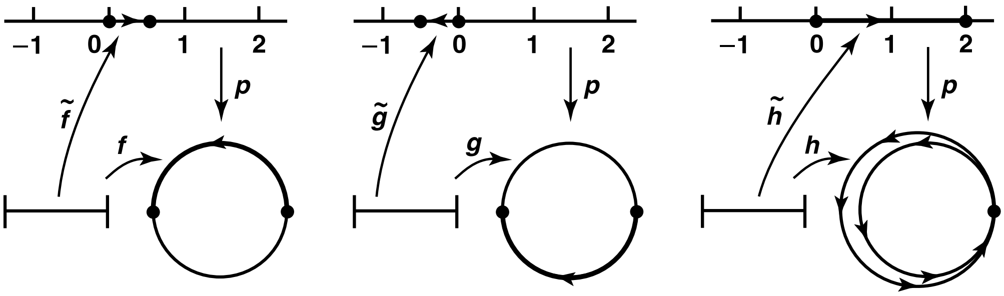


Figure 54.1

Lemma 54.1 Let $p : E \to B$ be a covering map; let $p(e_0) = b_0$. Any path $f : [0, 1] \to B$ beginning at b_0 has a unique lifting to a path $\tilde{f}$ in E beginning at e_0.

Remark Note that

- for each $b \in B$, since $p : E \to B$ is a covering map, there is an open neighborhood U_b of b that is evenly covered by p;
- since the collection $\mathscr{F} = \{U_b \mid b \in B\}$ is an open cover of B, and $f([0, 1])$ is a compact subspace of B, $\mathscr{F}$ is an open cover of $f([0, 1])$;
- by the Lebesgue number lemma, there exists an $\varepsilon > 0$ such that if A is a subset of $f([0, 1])$ with $\operatorname{diam}(A) < \varepsilon$, then there exists an $U_{b_i} \in \mathscr{F}$ containing A;
- with this (Lebesgue number of $\mathscr{F}$) $\varepsilon > 0$ and since $f : [0, 1] \to B$ is uniformly continuous, there exists a $\delta > 0$ such that if x, $y \in [0,\ 1]$ satisfying that $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon/2$.

This implies that if $0 = s_0 < s_1 < \cdots < s_n = 1$ is a subdivision of $[0,\ 1]$ such that $|s_{i+1} - s_i| < \delta$ for each i, then there exists an $U_{b_i} \in \mathscr{F}$ containing the set $f([s_i, s_{i+1}])$.

Proof Cover B by open sets U each of which is evenly covered by p. Use the Lebesgue number lemma to choose a subdivision of $[0, 1]$, say $0 = s_0 < s_1 < \cdots < s_n = 1$, such that for each i the set $f([s_i, s_{i+1}])$ lies in some open set U_i that is evenly covered by p.

Existence : Let $\tilde{f}(0) = e_0$ and suppose $\tilde{f}(s)$ is defined for $0 \le s \le s_i$. Since

- the set $f([s_i, s_{i+1}])$ lies in some open set U_i that is evenly covered by p,
- where $p^{-1}(U_i) = \bigcup_{\alpha} V_{i,\alpha}$ is the disjoint union of open subsets $\{V_{i,\alpha}\}$ in E,
- and each $V_{i,\alpha}$ is mapped homeomorphically onto U_i by p,

$\tilde{f}(s_i)$ lies in one of these sets, say in $V_{i,0}$. Define $\tilde{f}(s)$ for $s \in [s_i, s_{i+1}]$ by the equation

$$\tilde{f}(s) = \left(p|_{V_{i,0}}\right)^{-1}(f(s)).$$

Since $p|_{V_{i,0}} : V_{i,0} \to U_i$ is a homeomorphism and $f : [s_i, s_{i+1}] \to B$ is continuous, $\tilde{f}$ will be continuous on $[s_i, s_{i+1}]$.

Continuing in this way, we define $\tilde{f}$ on all of $[0,1]$. Continuity of $\tilde{f}$ follows from the pasting lemma; the fact that $p \circ \tilde{f} = f$ is immediate from the definition of $\tilde{f}$.

Uniqueness : Suppose that $\tilde{\tilde{f}}$ is another lifting of f beginning at e_0 such that $\tilde{\tilde{f}}(s) = \tilde{f}(s)$ for all $s \in [0,\, s_i]$, and suppose that $\tilde{f}(s) = \left(p|_{V_{i,0}}\right)^{-1}(f(s))$ for $s \in [s_i, s_{i+1}]$. Since $\tilde{\tilde{f}}(s_i) = \tilde{f}(s_i) \in V_{i,0}$, $\{V_{i,\alpha}\}$ are open, disjoint and $\tilde{\tilde{f}}([s_i, s_{i+1}])$ is connected, $\tilde{\tilde{f}}([s_i, s_{i+1}])$ must lie entirely in $V_{i,0}$. Thus, for each $s \in [s_i, s_{i+1}]$, $\tilde{\tilde{f}}(s) = y$ for some point $y \in p^{-1}(f(s)) \cap V_{i,0}$, and since $p^{-1}(f(s)) \cap V_{i,0} = \left(p|_{V_{i,0}}\right)^{-1}(f(s))$, $\tilde{\tilde{f}}(s) = \left(p|_{V_{i,0}}\right)^{-1}(f(s)) = \tilde{f}(s)$ for all $s \in [s_i, s_{i+1}]$, i.e. the lifting $\tilde{f}$ is unique.

Lemma 54.2 Let $p : E \to B$ be a covering map; let $p(e_0) = b_0$. Let the map $F : I \times I \to B$ be continuous, with $F(0,0) = b_0$. There is a lifting of F to a continuous map

$$\tilde{F} : I \times I \to E$$

such that $\tilde{F}(0,0) = e_0$. If F is a path homotopy, then $\tilde{F}$ is a path homotopy.

Proof Given F, we first define $\tilde{F}(0,0) = e_0$. Next, we use the preceding lemma to extend $\tilde{F}$ to the left-hand edge $0 \times I$ and the bottom edge $I \times 0$ of $I \times I$. Then we extend $\tilde{F}$ to all of $I \times I$ as follows:

Step 1 : Use the Lebesgue number lemma to choose subdivisions

$$0 = s_0 < s_1 < \cdots < s_m = 1, \quad 0 = t_0 < t_1 < \cdots < t_n = 1$$

of I fine enough that each rectangle

$$I_i \times J_j = [s_{i-1}, s_i] \times [t_{j-1}, t_j]$$

is mapped by F into an open set U_{ij} of B that is evenly covered by p.

Step 2 : We define the lifting $\tilde{F}$ step by step, beginning with the rectangle $I_1 \times J_1$, continuing with the other rectangles $I_i \times J_1$ in the "bottom row", then with the rectangles $I_i \times J_2$ in the next row, and so on.

In general, given $0 < i_0 < m$, $0 < j_0 < n$, assume that

- $\tilde{F}$ is defined on the set A which is the union of $0 \times I$ and $I \times 0$ and all the rectangles "previous" to $I_{i_0} \times J_{j_0}$, i.e.

$$A = (0 \times I) \cup (I \times 0) \cup \left(\bigcup_{j<j_0} \bigcup_{i=1}^{m} I_i \times J_j \right) \cup \left(\bigcup_{i<i_0} I_i \times J_{j_0} \right)$$

- $\tilde{F}$ is a continuous lifting of $F|_A$.

Choose an open set U of B that is evenly covered by p and contains the set $F(I_{i_0} \times J_{j_0})$. Let $\{V_\alpha\}$ be a partition of $p^{-1}(U)$ into slices; each set V_α is mapped homeomorphically onto U by p. Now $\tilde{F}$ is already defined on the set $C = A \cap (I_{i_0} \times J_{j_0})$. This set is the union of the left and

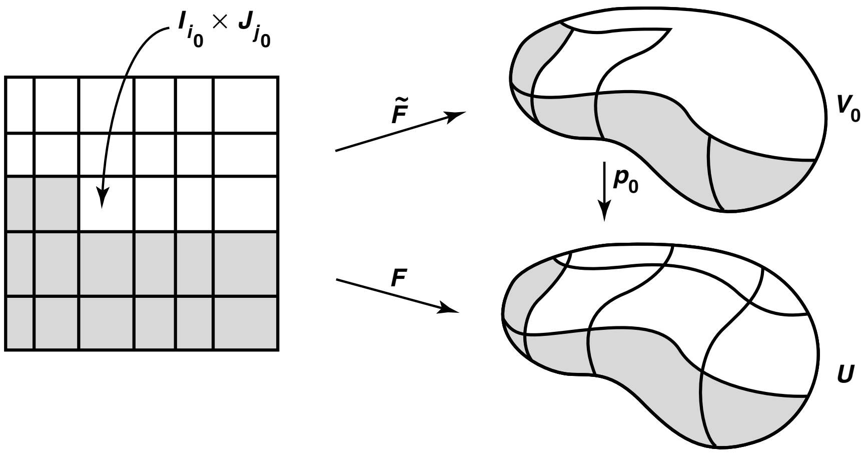


Figure 54.2

bottom edges of the rectangle $I_{i_0} \times J_{j_0}$, so it is connected. Therefore, $\tilde{F}(C)$ is connected and must lie entirely within one of the sets V_α. Suppose it lies in V_0. Then, the situation is as pictured in Figure 54.2.

Let $p_0 : V_0 \to U$ denote the restriction of p to V_0. Since $\tilde{F}$ is a lifting of $F|_A$, we know that for $x \in C$,

$$p_0\left(\tilde{F}(x)\right) = p\left(\tilde{F}(x)\right) = F(x) \implies \tilde{F}(x) = p_0^{-1}(F(x)) \quad \text{for } x \in C.$$

Hence we may extend $\tilde{F}$ by defining

$$\tilde{F}(x) = p_0^{-1}(F(x)) \quad \text{for } x \in I_{i_0} \times J_{j_0}.$$

Since $p_0 : V_0 \to U$ is a homeomorphism and F is continuous, the extended map $\tilde{F}$ will be continuous by the pasting lemma.

Continuing in this way, we define $\tilde{F}$ on all of I^2.

To check uniqueness, note that at each step of the construction of $\tilde{F}$, as we extend $\tilde{F}$ first to the bottom and left edges of I^2, and then to the rectangles $I_i \times J_j$, one by one, there is only one way to extend $\tilde{F}$ continuously. Thus, once the value of $\tilde{F}$ at (0, 0) is specified, $\tilde{F}$ is completely determined.

Now suppose that F is a path homotopy. We wish to show that $\tilde{F}$ is a path homotopy. The map F carries the entire left edge $0 \times I$ of I^2 into a single point b_0 of B. Because $\tilde{F}$ is a lifting of F, it carries this edge into the set $p^{-1}(b_0)$. But this set has the discrete topology as a subspace of E. Since $0 \times I$ is connected and $\tilde{F}$ is continuous, $\tilde{F}(0 \times I)$ is connected and thus must equal a one-point set. Similarly, $\tilde{F}(1 \times I)$ must be a one-point set. Thus $\tilde{F}$ is a path homotopy.

Theorem 54.3 Let $p : E \to B$ be a covering map; let $p(e_0) = b_0$. Let f and g be two paths in B from b_0 to b_1; let $\tilde{f}$ and $\tilde{g}$ be their respective liftings to paths in E beginning at e_0. If f and g are path homotopic, then $\tilde{f}$ and $\tilde{g}$ end at the same point of E and are path homotopic.

Proof Let $F : I \times I \to B$ be a path homotopy between f and g. Then $F(0, 0) = b_0$. Let $\tilde{F} : I \times I \to E$ be a lifting of F to E such that $\tilde{F}(0, 0) = e_0$. By the preceding lemma, $\tilde{F}$ is a path homotopy, so that $\tilde{F}(0 \times I) = \{e_0\}$ and $\tilde{F}(1 \times I)$ is a one-point set $\{e_1\}$.

The restriction $\tilde{F}|_{I\times 0}$ of $\tilde{F}$ to the bottom edge of $I \times I$ is a path on E beginning at e_0 that is a lifting of $F|_{I\times 0}$. By uniqueness of path liftings, we must have $\tilde{F}(s, 0) = \tilde{f}(s)$.

Similarly, $\tilde{F}|_{I\times 1}$ is a path on E that is a lifting of $F|_{I\times 1}$, and it begins at e_0 because $\tilde{F}(0 \times I) = \{e_0\}$. By uniqueness of path liftings, $\tilde{F}(s, 1) = \tilde{g}(s)$.

Therefore, both $\tilde{f}$ and $\tilde{g}$ end at e_1, and $\tilde{F}$ is a path homotopy between them.

Definition Let $p : E \to B$ be a covering map; let $b_0 \in B$. Choose e_0 so that $p(e_0) = b_0$. Given $[f] \in \pi_1(B, b_0)$, let $\tilde{f}$ be the lifting of f to a path in E that begins at e_0. Let $\phi([f])$ denote the end point $\tilde{f}(1)$ of $\tilde{f}$. Then ϕ is a well-defined set map

$$\phi : \pi_1(B, b_0) \to p^{-1}(b_0).$$

We call ϕ the *lifting correspondence* derived from the covering map p. It depends of course on the choice of the point e_0.

Theorem 54.4 Let $p : E \to B$ be a covering map; let $p(e_0) = b_0$. If E is path connected, then the lifting correspondence

$$\phi : \pi_1(B,\, b_0) \to p^{-1}(b_0)$$

is surjective. If E is simply connected, it is bijective.

Proof If E is path connected, then, given $e_1 \in p^{-1}(b_0)$, there is a path $\tilde{f}$ in E from e_0 to e_1. Then $f = p \circ \tilde{f}$ is a loop in B at b_0, and $\phi([f]) = \tilde{f}(1) = e_1$ by definition.

If E is simply connected and suppose that $[f]$ and $[g]$ are elements of $\pi_1(B,\, b_0)$ such that $\phi([f]) = \phi([g])$. Let $\tilde{f}$ and $\tilde{g}$ be the liftings of f and g, respectively, to paths in E that begin at e_0. Since $\tilde{f}(1) = \phi([f]) = \phi([g]) = \tilde{g}(1)$ and E is simply connected, there is a path homotopy $\tilde{F}$ in E between $\tilde{f}$ and $\tilde{g}$. Thus $p \circ \tilde{F}$ is a path homotopy in B between f and g, so ϕ is injective.

Definition If E is a simply connected space, and if $p : E \to B$ is a covering map, then we say that E is a *universal covering space* of B.

Theorem 54.5 The fundamental group $(\pi_1(S^1), *)$ of the circle is isomorphic to $(\mathbb{Z}, +)$.

Proof Let $p : \mathbb{R} \to S^1$ be a covering map defined by

$$p(x) = (\cos 2\pi x, \sin 2\pi x) \quad \text{for } x \in \mathbb{R},$$

let $e_0 = 0$, and let $b_0 = p(e_0) = (1, 0) \in S^1$. Then $p^{-1}(b_0)$ is the set $\mathbb{Z}$ of integers. Since $\mathbb{R}$ is simply connected, the lifting correspondence

$$\phi : \pi_1(S^1, b_0) \to \mathbb{Z}$$

is bijective. We show that ϕ is a homomorphism, and the theorem is proved.

Given $[f]$ and $[g]$ in $\pi_1(S^1, b_0)$, let $\tilde{f}$ and $\tilde{g}$ be their respective liftings to paths on $\mathbb{R}$ beginning at 0. Let $n = \tilde{f}(1)$ and $m = \tilde{g}(1)$; then $\phi([f]) = n$ and $\phi([g]) = m$, by definition.

Let $\tilde{\tilde{g}}$ be the path

$$\tilde{\tilde{g}}(s) = n + \tilde{g}(s) \quad \text{on } \mathbb{R}.$$

Because $p(n + x) = p(x)$ for $x \in \mathbb{R}$, the path $\tilde{\tilde{g}}$ is a lifting of g; it begins at n. Then the product $\tilde{f} * \tilde{\tilde{g}}$ is defined, and it is the lifting of $f * g$ that begins at 0, as you can check. The end point of this path is $\tilde{\tilde{g}}(1) = n + m$. Then by definition,

$$\phi([f] * [g]) = n + m = \phi([f]) + \phi([g]).$$

Definition Let G be a group; let x be an element of G. we denote the inverse of x by x^{-1}. The symbol x^n denotes the n-fold product of x with itself, x^{-n} denotes the n-fold product of x^{-1} with itself, and x^0 denotes the identity element of G. If the set of all elements of the form x^m, for $m \in \mathbb{Z}$, equals G, then G is said to be a *cyclic* group, and x is said to be a *generator* of G.

The cardinality of a group is also called the *order* of the group. A group is cyclic of infinite order if and only if it is isomorphic to the additive group of integers; it is cyclic of order k if and only if it is isomorphic to the group $\mathbb{Z}/k$ of integers modulo k. The preceding theorem implies that the fundamental group of the circle is infinite cyclic.

Note that if x is a generator of the infinite cyclic group G, and if y is an element of the arbitrary group H, then there is a unique homomorphism h of G into H such that $h(x) = y$; it is defined by setting $h(x^n) = y^n$ for all n.

Theorem 54.6 Let $p : E \to B$ be a covering map; let $p(e_0) = b_0$.

(a) The homomorphism $p_* : \pi_1(E, e_0) \to \pi_1(B, b_0)$ is a monomorphism, i.e. an injective homomorphism.

(b) Let $H = p_*(\pi_1(E, e_0))$. The lifting correspondence $\phi : \pi_1(B, b_0) \to p^{-1}(b_0)$ induces an injective map

$$\Phi : \pi_1(B, b_0)/H \to p^{-1}(b_0)$$

of the collection of right cosets $\pi_1(B, b_0)/H = \{H * [g] \mid [g] \in \pi_1(B, b_0)\}$ of H into $p^{-1}(b_0)$, which is bijective if E is path connected.

$$\begin{array}{ccc} p^{-1}(b_0) & \phi([f]) = \tilde{f}(1) \in p^{-1}(b_0) & \Phi([f]) = \phi([f]) = \tilde{f}(1) = \tilde{h} * \tilde{g}(1) = \tilde{g}(1) = \phi([g]) \\ \phi\Big\uparrow & \phi\Big\uparrow & \Phi\Big\uparrow \\ \pi_1(B, b_0) & [f] \in \pi_1(B, b_0) & [f] = [h * g] \in H * [g] \subset \pi_1(B, b_0)/H \end{array}$$

(c) If f is a loop in B based at b_0, then $[f] \in H$ if and only if f lifts to a loop in E based at e_0.

Proof (a) Suppose $\tilde{h}$ is a loop in E at e_0, and $p_*([\tilde{h}])$ is the identity element $[e_{b_0}]$. Let F be a path homotopy between $p \circ \tilde{h}$ and the constant loop e_{b_0}. If $\tilde{F}$ is the lifting of F to E such that $\tilde{F}(0, 0) = e_0$, then $\tilde{F}$ is a path homotopy between $\tilde{h}$ and the constant loop e_{e_0} at e_0.

(b) Given loops f and g in B, let $\tilde{f}$ and $\tilde{g}$ be liftings of them to E that begin at e_0. Then $\phi([f]) = \tilde{f}(1)$ and $\phi([g]) = \tilde{g}(1)$. To show the map $\Phi : \pi_1(B, b_0)/H \to p^{-1}(b_0)$ is injective is equivalent to show that $\phi([f]) = \phi([g])$ if and only if $[f] \in H * [g]$.

First, suppose that $[f] \in H * [g]$. Then $[f] = [h * g]$, where $h = p \circ \tilde{h}$ for some loop $\tilde{h}$ in E based at e_0. Now the product $\tilde{h} * \tilde{g}$ is defined, and it is a lifting of $h * g$. Because $[f] = [h * g]$, the liftings $\tilde{f}$ and $\tilde{h} * \tilde{g}$, which begin at e_0, must end at the same point of E. Then $\tilde{f}$ and $\tilde{g}$ end at the same point of E, so that $\phi([f]) = \phi([g])$. See Figure 54.3.

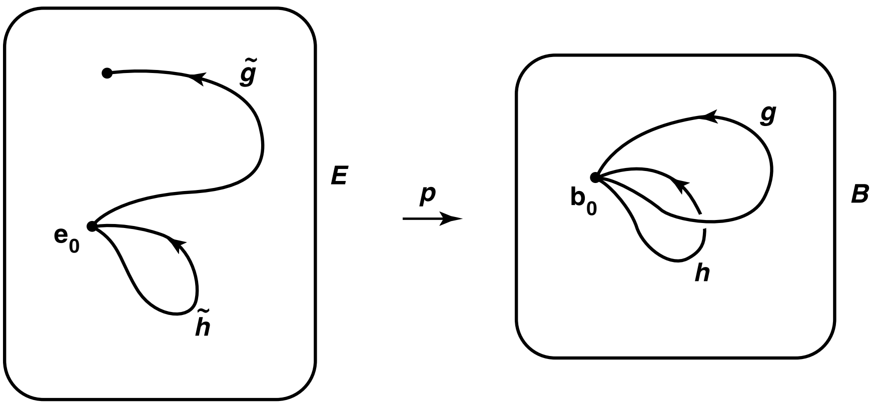


Figure 54.3

Now suppose that $\phi([f]) = \phi([g])$. Then $\tilde{f}$ and $\tilde{g}$ end at the same point of E. The product of $\tilde{f}$ and the reverse of $\tilde{g}$ is defined, and it is a loop $\tilde{h}$ in E based at e_0. By direct computation, $[\tilde{h} * \tilde{g}] = [\tilde{f}]$. If $\tilde{F}$ is a path homotopy in E between the loops $\tilde{h} * \tilde{g}$ and $\tilde{f}$, then $p \circ \tilde{F}$ is a path homotopy in B between $h * g$ and f, where $h = p \circ \tilde{h}$. Thus $[f] \in H * [g]$, as desired.

If E is path connected, then the lifting correspondence $\phi : \pi_1(B, b_0) \to p^{-1}(b_0)$ is surjective, so that Φ is surjective as well.

(c) Injectivity of Φ means that $\phi([f]) = \phi([g])$ if and only if $[f] \in H * [g]$. Applying this result in the case where g is the constant loop, we see that $\phi([f]) = e_0$ if and only if $[f] \in H$. But $\phi([f]) = e_0$ precisely when the lift of f that begins at e_0 also ends at e_0.

Theorem 5.13 (Basic Topology by Armstrong) If G acts as a group of homeomorphisms on a simply connected space X, and if each point $x \in X$ has an open neighborhood U which satisfies $U \cap g(U) = \emptyset$ for all $g \in G \setminus \{e\}$, then $\pi_1(X/G)$ is isomorphic to G.

Proof Fix a point $x_0 \in X$ and, given $g \in G$, join x_0 to $g(x_0)$ by a path γ. If $p : X \to X/G$ denotes the projection, $p \circ \gamma$ is a loop based at $p(x_0) \in X/G$.

Given $y \in X/G$, choose a point $x \in p^{-1}(y)$ and an open neighborhood U of x in X such that $U \cap g(U) = \emptyset$ for all $g \in G \setminus \{e\}$. If we set $V = p(U)$ and take $V_g = g(U)$ for each $g \in G$, then

- each $V_g = g(U)$ is homeomorphic to V, so it is open in X;
- V is open in X/G since $p : X \to X/G$ is a quotient map and $p^{-1}(V) = \bigcup_{g \in G} V_g$ is open in X.

This shows that $p : X \to X/G$ is a covering map and $\pi_1(X/G)$ is isomorphic to G since X is simply connected.

$$\begin{tikzcd} X \arrow[d, "p"] \\ X/G \end{tikzcd} \qquad \begin{tikzcd} O(x) = \{g(x) \mid g \in G\} \subset X \arrow[d, "p"] \\ \{x\} \subset X \end{tikzcd} \qquad \begin{tikzcd} \bigcup_{\{V_g = g(U) \mid g \in G\}} V_g \subset X \arrow[d, "p"] \\ V = p(U) \subseteq X/G \end{tikzcd}$$

Remark The subset $O(x) = \{g(x) \mid g \in G\}$ is called the *orbit of x*, and the quotient space $X/G = \{O(x) \mid x \in X\}$ of X is called the *orbit space*.

Remark In the proof, we have shown that the quotient map $p : X \to X/G$ is an open map, so it is a covering map. In particular, we have proved the following lemma.

Lemma If G acts as a group of homeomorphisms on a topological space X, the projection map $p : X \to X/G$ is an open map.

Proof For each $g \in G$ and any subset $U \subseteq X$, we define a set $g(U) \subseteq X$ by

$$g(U) = \{g(x) \mid x \in U\} \implies p^{-1}(p(U)) = \bigcup_{g \in G} g(U).$$

If $U \subseteq X$ is open, since each g is a homeomorphism from X onto X, each $g(U)$ is open, and therefore $p^{-1}(p(U)) = \bigcup_{g \in G} g(U)$ is open in X. Because p is a quotient map, this implies that $p(U)$ is open in X/G, and therefore p is an open map.

Example 1 $\mathbb{Z} \times \mathbb{Z}$ on $\mathbb{R}^2$ with the orbit space the torus $\mathbb{T}^2$, giving $\pi_1(\mathbb{T}^2) \cong \mathbb{Z} \times \mathbb{Z}$.

Example 2 $\mathbb{Z}_2$ on $\mathbb{S}^n$ with the orbit space $\mathbb{P}^n$, giving $\pi_1(\mathbb{P}^n) \cong \mathbb{Z}_2$.

§55 Retractions and Fixed Points

Definition Let $A \subset X$. A *retraction* of X onto A is a continuous map $r : X \to A$ such that $r|_A$ is the identity map of A. If such a map r exists, we say that A is a *retract* of X.

Lemma 55.1 If A is a retract of X, then the homomorphism of fundamental groups induced by inclusion $j : A \to X$ is injective.

Proof If $r : X \to A$ is a retraction, then the composite map $r \circ j$ equals the identity map of A. It follows that $r_* \circ j_*$ is the identity map of $\pi_1(A, a)$, so that j_* must be injective.

Theorem 55.2 **(No-retraction theorem)** There is no retraction of B^2 onto S^1.

Proof If S^1 were a retract of B^2, then the homomorphism induced by inclusion $j : S^1 \to B^2$ would be injective. But the fundamental group of S^1 is nontrivial and the fundamental group of B^2 is trivial.

Lemma 55.3 Let $h : S^1 \to X$ be a continuous map. Then the following conditions are equivalent:

(1) h is nulhomotopic.

(2) h extends to a continuous map $k : B^2 \to X$.

(3) h_* is the trivial homomorphism of fundamental groups.

Proof (1) $\Longrightarrow$ (2). Let $H : S^1 \times I \to X$ be a homotopy between h and a constant map. Let $\pi : S^1 \times I \to B^2$ be the map

$$\pi\,(x,\,t) = (1-t)x.$$

Then π is continuous, closed and surjective, so it is a quotient map; it collapses $S^1 \times 1$ to the point $\mathbf{0} = (0,0)$ and is otherwise injective. Because H is constant on $S^1 \times 1$, it induces, via the quotient map π, a continuous map $k : B^2 \to X$ that is an extension of h. See Figure 55.1.

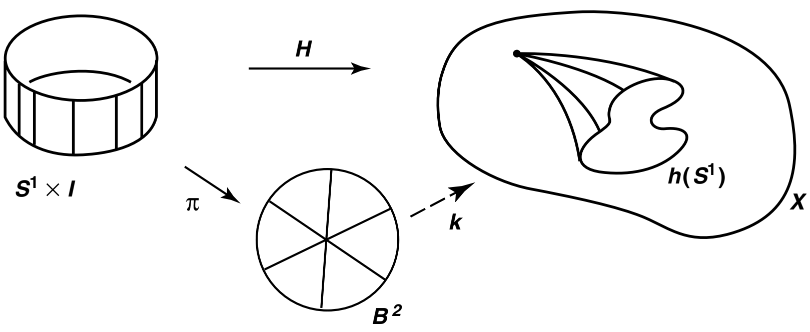


Figure 55.1

(2) $\Longrightarrow$ (3). If $j : S^1 \to B^2$ is the inclusion map, then h equals the composite $k \circ j$. Hence $h_* = k_* \circ j_*$. But

$$j_* : \pi_1\left(S^1,\, b_0\right) \to \left(B^2,\, b_0\right)$$

is trivial because the fundamental group of B^2 is trivial. Therefore h_* is trivial.

(3) $\Longrightarrow$ (1). Let $p : \mathbb{R} \to S^1$ be the standard covering map defined by $p(t) = (\cos 2\pi t, \sin 2\pi t)$, $t \in \mathbb{R}$, and let $p_0 : I \to S^1$ be its restriction to the unit interval $I = [0,1]$. Then $[p_0]$ generates $\pi_1\left(S^1,\, b_0\right)$ because p_0 is a loop in S^1 whose lift to $\mathbb{R}$ begins at 0 and ends at 1.

Let $x_0 = h(b_0)$, and let $f = h \circ p_0$ be a loop based at x_0. Since h_* is trivial,

$$[f] = [h \circ p_0] = h_*\left([p_0]\right) = \text{identity}_{\pi_1(X,x_0)} \implies f \simeq_p e_{x_0},$$

i.e. there is a path homotopy F in X between f and the constant path at x_0. Next, since the map $p_0 \times \mathrm{id} : I \times I \to S^1 \times I$ is continuous, closed, and surjective, it is a quotient map that maps $0 \times t$ and $1 \times t$ to $b_0 \times t$ for each t, and it is injective elsewhere; also since the path homotopy F maps $0 \times I$ and $1 \times I$ and $I \times 1$ to the point x_0 of X, it induces a continuous map $H : S^1 \times I \to X$ that is a homotopy between h and a constant map. See Figure 55.2.

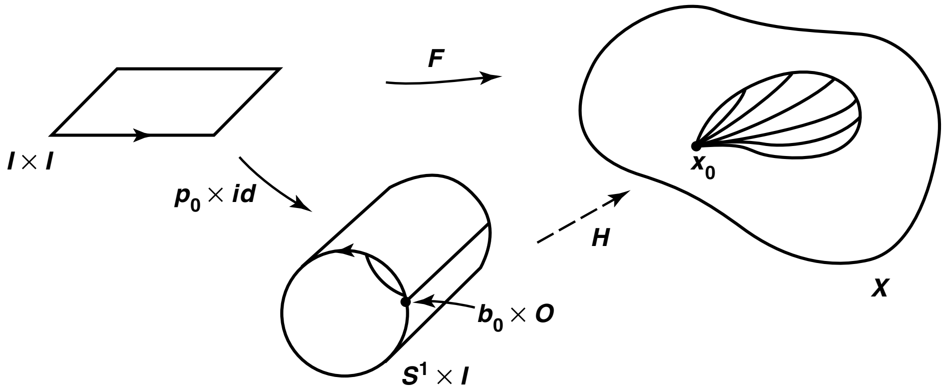


Figure 55.2

Corollary 55.4 The inclusion map $j : S^1 \to \mathbb{R}^2 \setminus \mathbf{0}$ is not nulhomotopic. The identity map $i : S^1 \to S^1$ is not nulhomotopic.

Proof There is a retraction of $\mathbb{R}^2 \setminus \mathbf{0}$ onto S^1 given by the equation $x/\|x\|$. Therefore, j_* is injective, and hence nontrivial. Similarly, i_* is the identity homomorphism, and hence nontrivial.

Theorem 55.5 Given a nonvanishing vector field on B^2, there exists a point of S^1 where the vector field points directly inward and a point of S^1 where it points directly outward.

Definition A *vector field* on B^2 is an ordered pair $(x, v(x))$, where x is in B^2 and v is a continuous map of B^2 into $\mathbb{R}^2$. In calculus, one often uses the notation

$$v(x) = v_1(x)\,\mathbf{i} + v_2(x)\,\mathbf{j}$$

for the function v, where $\mathbf{i} = (1, 0)$ and $\mathbf{j} = (0, 1)$ are the standard unit basis vectors in $\mathbb{R}^2$. But we shall stick with simple functional notation. To say that a vector field is *nonvanishing* means that $v(x) \neq \mathbf{0} = (0, 0)$ for every x; in such a case v actually maps B^2 into $\mathbb{R}^2 \setminus \mathbf{0}$.

Proof Consider the map $v : B^2 \to \mathbb{R}^2 \setminus \mathbf{0}$. Suppose that $v(x)$ does not point directly inward at any point x of S^1 and let w be its restriction to S^1. Because the map w extends to a map of B^2 into $\mathbb{R}^2 \setminus \mathbf{0}$, it is nulhomotopic.

On the other hand, the map F defined by

$$F(x, t) = tx + (1 - t)w(x), \quad \text{for } x \in S^1$$

clearly satisfies that $F(x, t) \neq \mathbf{0}$ for $t = 0$ or $t = 1$; and if $F(x, t) = \mathbf{0}$ for some t with $0 < t < 1$, then $tx + (1 - t)w(x) = \mathbf{0}$, so that $w(x)$ equals a negative scalar multiple of x, contradicting to the assumption that v does not point directly inward at any point of S^1. Hence $F(x, t) \neq \mathbf{0}$ for all $(x, t) \in S^1 \times I$, and $F : S^1 \times I \to \mathbb{R}^2 \setminus \mathbf{0}$ is a homotopy between $w : S^1 \to \mathbb{R}^2 \setminus \mathbf{0}$ and the inclusion map $j : S^1 \to \mathbb{R}^2 \setminus \mathbf{0}$.

It follows that j is nulhomotopic, contradicting the preceding corollary.

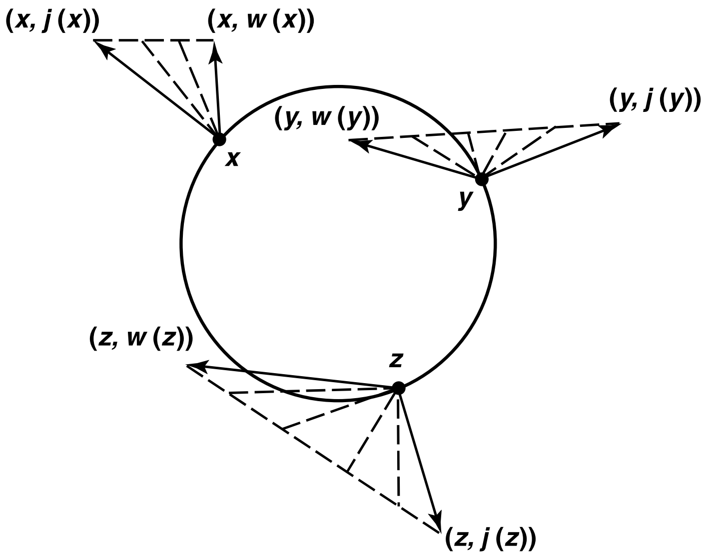


Figure 55.3

To show that v points directly outward at some point of S^1, we apply the result just proved to the vector field $(x, -v(x))$.

Theorem 55.6 (Brouwer fixed-point theorem for the disc) If $f : B^2 \to B^2$ is continuous, then there exists a point $x \in B^2$ such that $f(x) = x$.

Proof We proceed by contradiction. Suppose that $f(x) \neq x$ for every x in B^2. Then defining $v(x) = f(x) - x$ gives us a nonvanishing vector field $(x, v(x))$ on B^2. But the vector field v cannot point directly outward at any point x of S^1, for that would mean

$$f(x) - x = ax \quad \text{for some positive real number } a \implies f(x) = (1+a)x \notin B^2.$$

We thus arrive at a contradiction.

§58 Deformation Retracts and Homotopy Type

Theorem 58.2 Let $x_0 \in \mathbb{S}^1$. The inclusion mapping

$$j : (\mathbb{S}^1, x_0) \to (\mathbb{R}^2 \setminus \mathbf{0}, x_0), \quad \text{where } \mathbf{0} = (0,0) \in \mathbb{R}^2$$

induces an isomorphism of fundamental groups.

Proof Let $r : \mathbb{R}^2 \setminus \mathbf{0} \to \mathbb{S}^1$ be the continuous map defined by

$$r(x) = \frac{x}{\|x\|}, \quad \text{where } \|x\| \text{ is the Euclidean distance from } x \text{ to the origin } \mathbf{0}.$$

Given any loop f in $\mathbb{S}^1$ based at x_0, since $r \circ j(x) = x$ for all $x \in \mathbb{S}^1$, we have $r \circ j \circ f = f$ and

$$r_* \circ j_* ([f]) = [f] \implies r_* \circ j_* = \text{identity}_{\pi_1(\mathbb{S}^1, x_0)},$$

that is, $r_* : \pi_1 \left(\mathbb{R}^2 \setminus \mathbf{0}, x_0\right) \to \pi_1 \left(\mathbb{S}^1, x_0\right)$ is a left inverse for $j_* : \pi_1 \left(\mathbb{S}^1, x_0\right) \to \pi_1 \left(\mathbb{R}^2 \setminus \mathbf{0}, x_0\right)$.
On the other hand, given any loop f in $\mathbb{R}^2 \setminus \mathbf{0}$ based at x_0, let g be defined by

$$g(s) = j \circ r \circ f(s) = \frac{f(s)}{\|f(s)\|} \quad \text{for } s \in I.$$

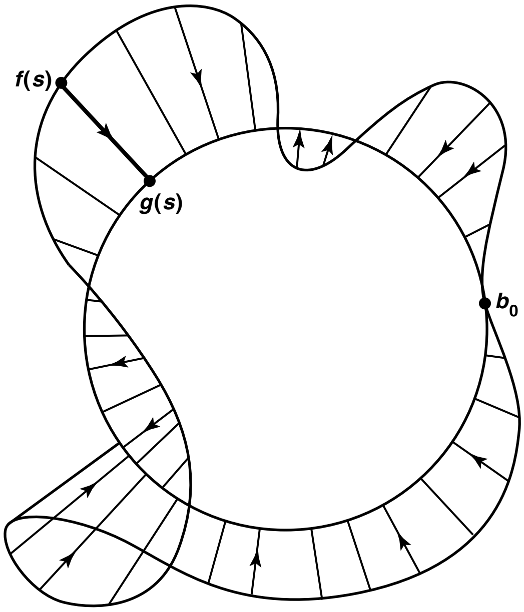


Figure 58.1

Then g is a loop in $\mathbb{S}^1$ based at x_0, and the continuous map F defined by

$$F(s,t) = t\,g(s) + (1-t)\,f(s) = t\,\frac{f(s)}{\|f(s)\|} + (1-t)\,f(s) \quad \text{for } s,\, t \in I \quad \text{(See Figure 58.1)}$$

is a path homotopy between f and g, since $F(0,t) = F(1,t) = x_0$ for all $t \in I$, and since

$$f(s) \in \mathbb{R}^2 \setminus \mathbf{0} \implies \|f(s)\| > 0 \implies \frac{t}{\|f(s)\|} + (1-t) \neq 0 \implies F(s,t) \neq \mathbf{0} \quad \text{for all } s,\, t \in I,$$

where we note that $t/\|f(s)\| + (1-t)$ is a number between 1 and $1/\|f(s)\|$. Hence,

$$[f] = [g] = [j \circ r \circ f] = j_* \circ r_*([f]) \implies j_* \circ r_* = \text{identity}_{\pi_1(\mathbb{R}^2 \setminus \mathbf{0}, x_0)},$$

that is, $r_* : \pi_1\left(\mathbb{R}^2 \setminus \mathbf{0}, x_0\right) \to \pi_1\left(\mathbb{S}^1, x_0\right)$ is a right inverse for $j_* : \pi_1\left(\mathbb{S}^1, x_0\right) \to \pi_1\left(\mathbb{R}^2 \setminus \mathbf{0}, x_0\right)$.
So, r_* is the inverse for j_* and $\pi_1(\mathbb{S}^1, x_0)$ is isomorphic to $\pi_1(\mathbb{R}^2 \setminus \mathbf{0}, x_0)$.

Theorem If $x_0 \in \mathbb{S}^{n-1}$, the inclusion mapping

$$j : (\mathbb{S}^{n-1}, x_0) \to (\mathbb{R}^n \setminus \mathbf{0}, x_0), \quad \text{where } \mathbf{0} = (0, \ldots, 0) \in \mathbb{R}^n$$

induces an isomorphism of fundamental groups.

Definition Let A be a subspace of X. Then A is said to be a *deformation retract* of X if there is a continuous map $H : X \times I \to X$ such that

$$\begin{aligned} &H(x,0) = x \quad \text{and} \quad H(x,1) \in A && \text{for all } x \in X, \\ \text{and} \quad &H(a,t) = a && \text{for all } a \in A,\ t \in I. \end{aligned}$$

The map H is called a *deformation retraction.*

Example The map $H : \mathbb{R}^n \setminus \mathbf{0} \times I \to \mathbb{R}^n \setminus \mathbf{0}$ defined by

$$H(x,t) = t\,\frac{x}{\|x\|} + (1-t)\,x$$

is a deformation retract of $\mathbb{R}^n \setminus \mathbf{0}$ onto $\mathbb{S}^{n-1}$; it gradually collapses each radial line into the point where it intersects $\mathbb{S}^{n-1}$.

Example Illustration in Figure 58.2 sketches the "figure eight" space as a three-stage deformation retract of the doubly punctured plane $\mathbb{R} \setminus \{p, q\}$.

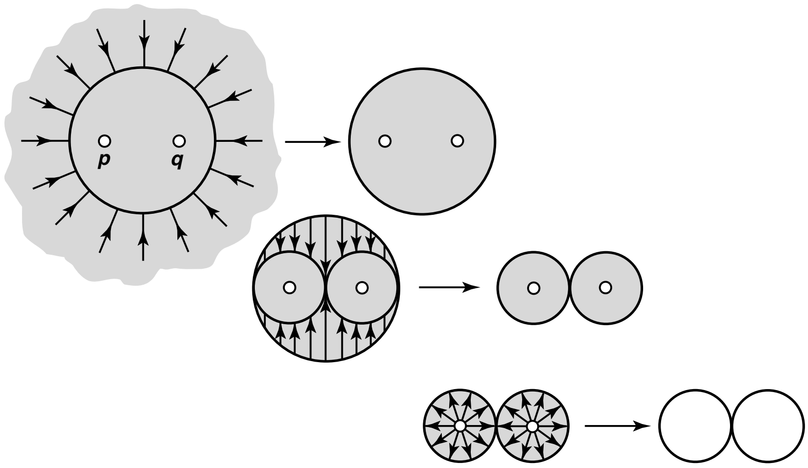


Figure 58.2

Theorem 58.3 Let A be a deformation retract of X, and let $a_0 \in A$. Then the inclusion map

$$j : (A, a_0) \to (X, a_0)$$

induces an isomorphism of fundamental groups.

The proof is similar to that of the preceding Theorem of $\pi_1(\mathbb{S}^1, x_0) \simeq \pi_1(\mathbb{R}^2 \setminus \mathbf{0}, x_0)$.

Definition Let $f : X \to Y$ and $g : Y \to X$ be continuous maps. Suppose that the map $g \circ f : X \to X$ is homotopic to the identity map of X, and the map $f \circ g : Y \to Y$ is homotopic to the identity map of Y. Then the maps f and g are called *homotopy equivalences,* and each is said to be a *homotopy inverse* of the other.

It is straightforward to show that if $f : X \to Y$ is a homotopy equivalence of X with Y and $h : Y \to Z$ is a homotopy equivalence of Y with Z, then $h \circ f : X \to Z$ is a homotopy equivalence of X with Z. It follows that the relation of homotopy equivalence is an equivalence relation. Two spaces that are homotopy equivalent are said to have the same *homotopy type.*

Note that if A is a deformation retract of X, then A has the same homotopy type as X. For let $j : A \to X$ be the inclusion mapping and let $r : X \to A$ be the retraction mapping. Then the composite $r \circ j$ equals the identity map of A, and the composite $j \circ r$ is by hypothesis homotopic to the identity map of X (and in fact each point of A remains fixed during the homotopy).

We now show that two spaces having the same homotopy type have isomorphic fundamental groups. For this purpose, we need to study what happens when we have a homotopy between two continuous maps of X into Y such that the base point of X does not remain fixed during the homotopy.

Lemma 58.4 Let $h,\ k : X \to Y$ be continuous maps; let $h(x_0) = y_0$ and $k(x_0) = y_1$. If h and k are homotopic, there is a path α in Y from y_0 to y_1 such that $k_* = \widehat{\alpha} \circ h_*$. Indeed, if $H : X \times I \to Y$

is the homotopy between h and k, then α is the path $\alpha(t) = H(x_0, t)$.

$$\begin{tikzcd} \pi_1(X, x_0) \arrow[r, "h_*"] \arrow[dr, "k_*"'] & \pi_1(Y, y_0) \arrow[d, "\widehat{\alpha}"] \\ & \pi_1(Y, y_1) \end{tikzcd}$$

Proof Let $f : I \to X$ be a loop in X based at x_0. We must show that

$$k_*([f]) = \widehat{\alpha}\,(h_*([f]))\,.$$

This equation states that $[k \circ f] = [\bar{\alpha}] * [h \circ f] * [\alpha]$, or equivalently, that

$$[\alpha] * [k \circ f] = [h \circ f] * [\alpha].$$

This is the equation we shall verify. To begin, consider the loops f_0 and f_1 in the space $X \times I$ given by the equations

$$f_0(s) = (f(s), 0) \quad \text{and} \quad f_1(s) = (f(s), 1).$$

Consider also the path c in $X \times I$ given by the equation

$$c(t) = (x_0, t).$$

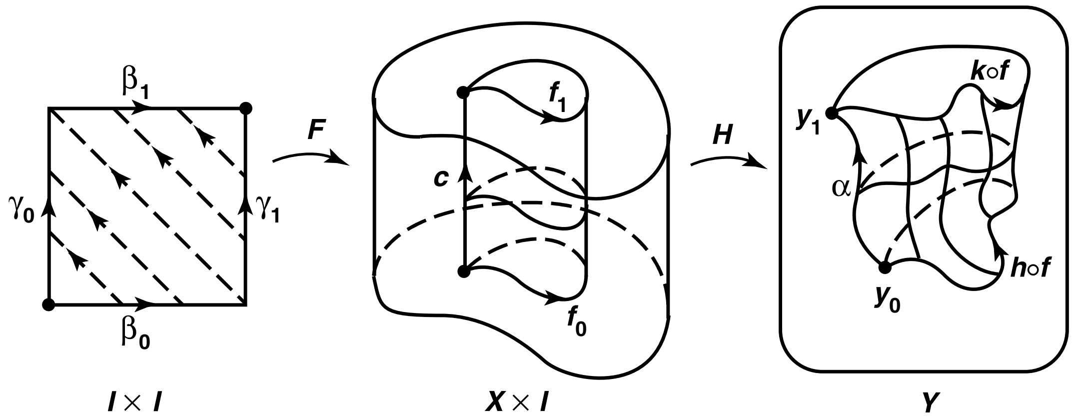


Figure 58.3

Then $H \circ f_0 = h \circ f$ and $H \circ f_1 = k \circ f$, while $H \circ c$ equals the path α. See Figure 58.3.
Let $F : I \times I \to X \times I$ be the map $F(s, t) = (f(s), t)$. Consider the following paths in $I \times I$, which run along the four edges of $I \times I$:

$$\begin{aligned} \beta_0(s) &= (s, 0) \quad \text{and} \quad \beta_1(s) = (s, 1), \\ \gamma_0(t) &= (0, t) \quad \text{and} \quad \gamma_1(t) = (1, t). \end{aligned}$$

Then $F \circ \beta_0 = f_0$ and $F \circ \beta_1 = f_1$, while $F \circ \gamma_0 = F \circ \gamma_1 = c$.

The broken-line paths $\beta_0 * \gamma_1$ and $\gamma_0 * \beta_1$ are paths in $I \times I$ from $(0,\, 0)$ to $(1,\, 1)$; since $I \times I$ is convex, there is a path homotopy G between them. Then $F \circ G$ is a path homotopy in $X \times I$ between $f_0 * c$ and $c * f_1$. And $H \circ (F \circ G)$ is a path homotopy in Y between

$$\begin{aligned}&(H \circ f_0) * (H \circ c) = (h \circ f) * \alpha \quad \text{and}\\ &(H \circ c) * (H \circ f_1) = \alpha * (k \circ f),\end{aligned}$$

as desired.

Corollary 58.5 Let $h,\, k : X \to Y$ be homotopic continuous maps; let $h(x_0) = y_0$ and $k(x_0) = y_1$. If h_* is injective, or surjective, or trivial, so is k_*.

Corollary 58.6 Let $h : X \to Y$. If h is nulhomotopic, then h_* is the trivial homomorphism.

Proof The constant map induces the trivial homomorphism.

Theorem 58.7 Let $f : X \to Y$; be continuous; let $f(x_0) = y_0$. If f is a homotopy equivalence, then

$$f_* : \pi_1(X,\, x_0) \to \pi_1(Y,\, y_0)$$

is an isomorphism.

Proof Let $g : Y \to X$ be a homotopy inverse for f. Consider the maps

$$(X,\, x_0) \xrightarrow{f} (Y,\, y_0) \xrightarrow{g} (X,\, x_1) \xrightarrow{f} (Y,\, y_1),$$

while $x_1 = g(y_0)$ and $y_1 = f(x_1)$. We have the corresponding induced homomorphisms:

$$\begin{tikzcd} \pi_1(X,\, x_0) \arrow[r, "(f_{x_0})_*"] & \pi_1(Y,\, y_0) \arrow[dl, "g_*"] \\ \pi_1(X,\, x_1) \arrow[r, "(f_{x_1})_*"'] & \pi_1(Y,\, y_1) \end{tikzcd}$$

[Here we have to distinguish between the homomorphisms induced by f relative to two different base points.] Now

$$g \circ f : (X, x_0) \to (X, x_1)$$

is by hypothesis homotopic to the identity map, so there is a path α in X such that

$$(g \circ f)_* = \widehat{\alpha} \circ (i_X)_* = \widehat{\alpha}.$$

It follows that $(g \circ f)_* = g_* \circ (f_{x_0})_*$ is an isomorphism.

Similarly, because $f \circ g : (Y, y_0) \to (Y, y_1)$ homotopic to the identity map i_Y, the homomorphism $(f \circ g)_* = (f_{x_1})_* \circ g_*$ is an isomorphism.

The first fact implies that g_* is surjective, and the second implies that g_* is injective. Therefore, g_* is an isomorphism. Applying the first equation once again, we conclude that

$$(f_{x_0})_* = (g_*)^{-1} \circ \widehat{\alpha},$$

so that $(f_{x_0})_*$ is also an isomorphism.

Note that although g is a homotopy inverse for f, the homomorphism g_* is not an inverse for the homomorphism $(f_{x_0})_*$.

§59 The Fundamental Group of S^n

Theorem 59.1 (Van Kampen Theorem) Let $X = U \cup V$, where U and V are open sets of X. Suppose that $U \cap V$ is path connected, and that $x_0 \in U \cap V$. Let i and j be the inclusion mappings of U and V, respectively, into X. Then the images of the induced homomorphisms

$$i_* : \pi_1(U, x_0) \to \pi_1(X, x_0) \quad \text{and} \quad j_* : \pi_1(V, x_0) \to \pi_1(X, x_0) \quad \text{generate } \pi_1(X, x_0).$$

Proof This theorem states that, given any loop f in X based at x_0, it is path homotopic to a product of the form $(g_1 * (g_2 * (\cdots * g_n)))$, where each g_i is a loop in X based at x_0 that lies either in U or in V.

Let $f : I \to X$ be a loop based at x_0. We wish to show that f is path homotopic to a constant loop.

Step 1. By the Lebesgue number lemma, there is a subdivision

$$0 = a_0 < a_1 < \cdots < a_n = 1$$

of the interval $[0, 1]$ such that for each i, the set $f([a_{i-1}, a_i])$ lies entirely in either U or V. Among all such subdivisions, choose one for which the number n of subintervals is minimal. Then it follows that for each i, the point $f(a_i) \in U \cap V$.

Suppose that $f(a_i) \notin U$, for instance. Then neither $f([a_{i-1}, a_i])$ nor $f([a_i, a_{i+1}])$ lies entirely in U. Therefore, both of them must lie entirely in V. We can then discard a_i from the subdivision, and still have a subdivision of $[0, 1]$ for which the image of each subinterval lies either in U or in V. This contradicts to minimality of the subdivisions. Hence $f(a_i)$ must belong to U.

Step 2. Let f_i be the restriction of f to the interval $[a_{i-1}, a_i]$ defined by

$$f_i(s) = f((1-s)a_{i-1} + sa_i) \quad \text{for } s \in [0, 1].$$

Then f_i is a path that lies either in U or in V, and

$$[f] = [f_1] * [f_2] * \cdots * [f_n]$$

For each i, choose a path α_i in $U \cap V$ from x_0 to $f(a_i)$ (Here we use the fact that $U \cap V$ is path connected.) Since $f(a_0) = f(a_n) = x_0$, we can choose α_0 and α_n to be the constant path at x_0.

Now we set

$$g_i = (\alpha_{i-1} * f_i) * \bar{\alpha}_i \quad \text{for each } i.$$

Then g_i is a loop in X based at x_0 whose image lies either in U or in V. Direct computation shows that

$$[g_1] * [g_2] * \cdots * [g_n] = [f_1] * [f_2] * \cdots * [f_n].$$

Corollary 59.2 (The special Van Kampen theorem) Suppose $X = U \cup V$, where U and V are open sets of X; suppose $U \cap V$ is nonempty and path connected. If U and V are simply connected, then X is simply connected.

Theorem 59.3 For $n \geq 2$, the n-sphere $\mathbb{S}^n$ is simply connected.

Proof Let $p = (0, \ldots, 0, 1) \in \mathbb{R}^{n+1}$ be the "north pole" of $\mathbb{S}^n$; let $q = (0, \ldots, 0, -1) \in \mathbb{R}^{n+1}$ be the "south pole" of $\mathbb{S}^n$.

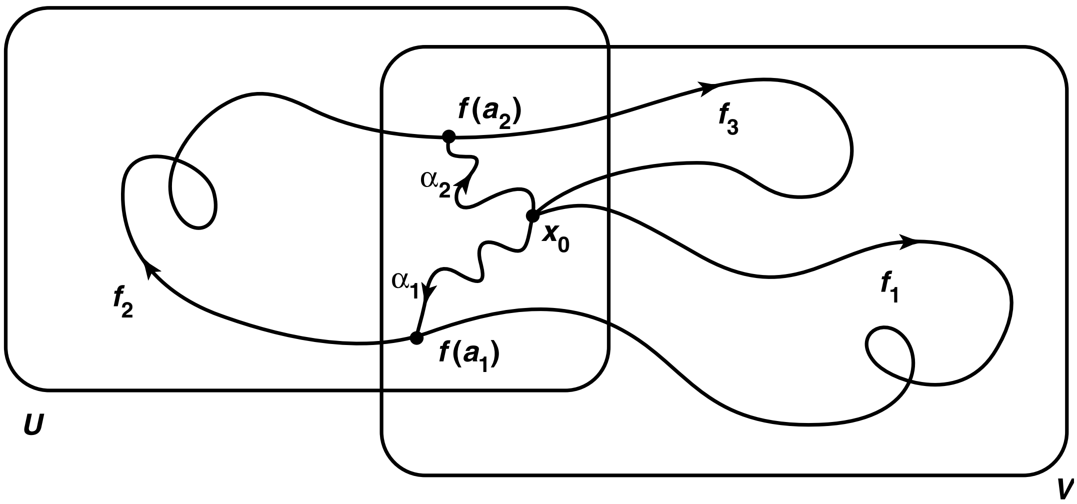


Figure 59.1

Step 1. For each $x = (x_1, \ldots, x_{n+1}) \neq p$, $x \in \mathbb{S}^n$, let the *stereographic projection* map $f : \mathbb{S}^n \setminus p \to \mathbb{R}^n$ be define by

$$f(x) = \frac{1}{1 - x_{n+1}} (x_1, \ldots, x_n).$$

Obviously, f is continuous and it is in fact a homeomorphism since the map $g : \mathbb{R}^n \to \mathbb{S}^n \setminus p$ given by

$$g(y_1, \ldots, y_n) = (ty_1,\, ty_2, \ldots,\, ty_n,\, 1 - t), \quad \text{where } t = \frac{2}{1 + y_1^2 + \cdots + y_n^2}$$

is an inverse for f. So, $\mathbb{S}^n \setminus p$ is homeomorphic to $\mathbb{R}^n$.

Since $\mathbb{S}^n \setminus q$ is homeomorphic to $\mathbb{S}^n \setminus p$ under the reflection map

$$p(x_1, \ldots,\, x_n,\, x_{n+1}) = (x_1, \ldots,\, x_n,\, -x_{n+1}),$$

$\mathbb{S}^n \setminus q$ is also homeomorphic to $\mathbb{R}^n$.

Hence both $\mathbb{S}^n \setminus p$ and $\mathbb{S}^n \setminus p$ are simply connected since $\mathbb{R}^n$ is simply connected.

Step 2. Let $U = \mathbb{S}^n \setminus p$ and $V = \mathbb{S}^n \setminus q$. Since U and V are simply connected open sets of $\mathbb{S}^n$ such that $U \cup V = \mathbb{S}^n$, and since $U \cap V \simeq \mathbb{R}^n \setminus \mathbf{0}$ is path connected, $\mathbb{S}^n$ is simply connected by the special Van Kampen theorem.

Corollary $\mathbb{R}^n \setminus \mathbf{0}$ is simply connected if $n > 2$.

Proof If $n > 2$ and if $x_0 \in \mathbb{S}^{n-1}$, since $\mathbb{S}^{n-1}$ is simply connected and, by a preceding theorem, the inclusion map

$$j : (\mathbb{S}^{n-1}, x_0) \to (\mathbb{R}^n \setminus \mathbf{0}, x_0)$$

induces an isomorphism of fundamental groups, $\mathbb{R}^n \setminus \mathbf{0}$ is simply connected.

Corollary $\mathbb{R}^n$ and $\mathbb{R}^2$ are not homeomorphic for $n > 2$.

Proof Deleting a point from $\mathbb{R}^n$ leaves a simply connected space, while deleting a point from $\mathbb{R}^2$ does not.

§60 Fundamental Groups of Some Surfaces

Definition A *surface* is a Hausdorff space with a countable basis, every point of which has a neighborhood that is homeomorphic with an open subset of $\mathbb{R}^2$.

Surfaces are of interest in various parts of mathematics, including geometry, topology, and complex analysis. We consider here several surfaces, including the torus and double torus, and show by comparing their fundamental groups that they are not homeomorphic. In a later chapter, we shall classify up to homeomorphism all compact surfaces.

First, we consider the torus. In an earlier exercise, we asked you to compute its fundamental group using the theory of covering spaces. Here, we compute its fundamental group by using a theorem about the fundamental group of a product space.

Recall that if A and B are groups with operation $\cdot$, then the cartesian product $A \times B$ is given a group structure by using the operation

$$(a \times b) \cdot (a' \times b') = (a \cdot a') \times (b \cdot b')\,.$$

Recall also that if $h : C \to A$ and $k : C \to B$ are group homomorphisms, then the map $\Phi : C \to A \times B$ defined by $\Phi(c) = h(c) \times k(c)$ is a group homomorphism.

Theorem 60.1 Let X, Y be topological spaces, and let $x_0 \in X$, $y_0 \in Y$. Then $\pi_1(X \times Y, x_0 \times y_0)$ is isomorphic with $\pi_1(X, x_0) \times \pi_1(Y, y_0)$.

Remark Recall that if A and B are groups with operation $\cdot$, then the Cartesian product $A \times B$ is given a group structure by using the operation

$$(a \times b) \cdot (a' \times b') = (a \cdot a') \times (b \cdot b').$$

Recall also that if $h : C \to A$ and $k : C \to B$ are group homomorphisms, then the map $\Phi : C \to A \times B$ defined by $\Phi(c) = h(c) \times k(c)$ is a group homomorphism.

Let $p : X \times Y \to X$ and $q : X \times Y \to Y$ be the projection mappings, $p_* : \pi_1(X \times Y, x_0 \times y_0) \to \pi_1(X, x_0)$ and $q_* : \pi_1(X \times Y, x_0 \times y_0) \to \pi_1(Y, y_0)$ be the induced homomorphisms. Then we define a homomorphism

$$\Phi : \pi_1(X \times Y, x_0 \times y_0) \to \pi_1(X, x_0) \times \pi_1(Y, y_0)$$

by the equation

$$\Phi([f]) = p_*([f]) \times q_*([f]) = [p \circ f] \times [q \circ f].$$

Let $g : I \to X$ be a loop in X based at x_0, and let $h : I \to Y$ be a loop in Y based at y_0. Since the map $f : I \to X \times Y$ defined by

$$f(s) = g(s) \times h(s), \quad \text{for } s \in I$$

is a loop in $X \times Y$ based at $x_0 \times y_0$ and satisfies that

$$\Phi([f]) = [p \circ f] \times [q \circ f] = [g] \times [h],$$

the map Φ is surjective.

Suppose that $f : I \to X \times Y$ is a loop in $X \times Y$ based at $x_0 \times y_0$ such that

$$\Phi([f]) = [e_{x_0}] \times [e_{y_0}].$$

Since $\Phi([f]) = [p \circ f] \times [q \circ f]$, we have $p \circ f \simeq_p e_{x_0}$ and $q \circ f \simeq_p e_{y_0}$; let G and H be the respective path homotopies. Then the map $F : I \times I \to X \times Y$ defined by

$$F(s,t) = G(s,t) \times H(s,t)$$

is a path homotopy between f and the constant loop based at $x_0 \times y_0$.

Hence Φ is an isomorphism.

Corollary 60.2 The fundamental group of the torus $T = \mathbb{S}^1 \times \mathbb{S}^1$ is isomorphic to the group $\mathbb{Z} \times \mathbb{Z}$.

Definition The *projective plane* $\mathbb{P}^2$ is the space obtained from $\mathbb{S}^2$ by identifying each point x of $\mathbb{S}^2$ with its antipodal point $-x$.

Theorem 60.3 The projective plane $\mathbb{P}^2$ is a surface, and the map $p : \mathbb{S}^2 \to \mathbb{P}^2$ is a covering map.

Corollary 60.4 $\pi_1(\mathbb{P}^2, y)$ is a group of order 2.

Lemma 60.5 The fundamental group of the figure eight is not abelian.

Van Kampen's Theorem (revisit) Let X be a topological space and let U, $V \subset X$ be open subsets such that $U \cap V$ is nonempty and path-connected. Let $x \in U \cap V$ be a basepoint. Then

$$\pi_1(X, x) = \pi_1(U, x) *_{\pi_1(U\cap V)} \pi_1(V, x).$$

Here, $A *_C B$ denotes the *amalgamated product.* Suppose you have groups A, B, C and homomorphisms $f : C \to A$ and $g : C \to B$. In our case, $A = \pi_1(U, x)$, $B = \pi_1(V, x)$ and $C = \pi_1(U \cap V, x)$, the map f is the pushforward map i_* where $i : U \cap V \to U$ is the inclusion, and g is the pushforward j_* where $j : U \cap V \to V$ is the inclusion.

Given A, B, C, f, g, you can define the amalgamated product

$$A *_C B = \langle \text{generators of } A,\ \text{generators of } B \mid \text{relations of } A, \text{relations of } B, \text{amalgamated relations}\rangle.$$

The amalgamated relations come from elements $c \in C$: each $c \in C$ gives a relation $f(c) = g(c)$.

Amalgamated relations

What does this mean? We have $f(c) \in A$ and we have already in the presentation of $A *_C B$ the generators and relations of A, so I can make sense of the element $f(c)$ in the amalgamated product. Similarly, $g(c) \in B$ and we can think of this as an element in the amalgamated product. The corresponding amalgamated relation is just saying that these two elements agree.

Geometrically, in the context of Van Kampen's theorem, C consists of loops living in the intersection $U \cap V$; the amalgamated relations are then saying "you can think of these loops as living in U; you can think of them living in V; it makes no difference".

Example 1. The 2-sphere X can be written as $U \cup V$ where U is a neighbourhood of the Northern hemisphere and V is a neighbourhood of the Southern hemisphere. The overlap $U \cap V$ is an annular neighbourhood of the equator. Van Kampen's theorem tells us that $\pi_1(X, x) = \pi_1(U, x) *_{\pi_1(U\cap V, x)} \pi_1(V, x)$.

We have $\pi_1(U) = \pi_1(V) = \{1\}$ as both U and V are simply-connected discs. Since $U \cap V$ is homotopy equivalent to the circle, $\pi_1(U \cap V) = \mathbb{Z} = \langle c\rangle$ (i.e. one generator, c, and no relations).

The amalgamated product $\pi_1(U) *_{\pi_1(U\cap V)} \pi_1(V)$ has the empty set of generators coming from $\pi_1(U)$ and the empty set of generators coming from $\pi_1(V)$. Therefore it has no generators and the amalgamated product is trivial.

The fact that $\pi_1(U \cap V)$ is nontrivial doesn't affect this computation: it would only enter into the relations in the group. Indeed, for each $c \in \pi_1(U \cap V)$, we have $i_*c \in \pi_1(U)$ and $j_*c \in \pi_1(V)$, so the amalgamated relations all become $1 = 1$, which we don't need to include as a relation because it holds in any group.

The upshot of all of this is the statement that $\pi_1(\mathbb{S}^2) = \{1\}$.

Example 2. Let $X = S^1 \vee S^1$ be the wedge of two circles (i.e. the disjoint union of two circles modulo an equivalence relation which identifies one point on the first circle with one point on the second. We take U to be a neighbourhood of the first circle and V to be a neighbourhood of the second. The intersection $U \cap V$ is a cross-shaped neighbourhood of the point where the two circles intersect (the wedge point).

Since $U \simeq S^1$ and $V \simeq S^1$ we have $\pi_1(U) = \pi_1(V) = \mathbb{Z}$. Moreover, $U \cap V$ is contractible, so $\pi_1(U \cap V) = \{1\}$. The amalgamated product has a generator a coming from $\pi_1(U)$ and a generator b coming from $\pi_1(V)$. There are no relations coming from $\pi_1(U)$, none coming from $\pi_1(V)$ and also no amalgamated relations (since $\pi_1(U \cap V) = \{1\}$).

A presentation for $\pi_1(S^1 \vee S^1)$ is therefore

$$\langle a, b \rangle$$

This has no relations: we call a group with no relations a *free group* on its generators. The elements of a free group $\langle a, b \rangle$ are words on its generators, written using as, bs, a^{-1}s, b^{-1}s, for example

$$a^2 b^{-1} a^{14} bab^2.$$

The only simplifications one may perform with these elements are things like $aa^{-1} = 1$ or $bb^{-1} = 1$. In particular, the free group on two generators is *not abelian* (the relation $ab = ba$ does not hold).

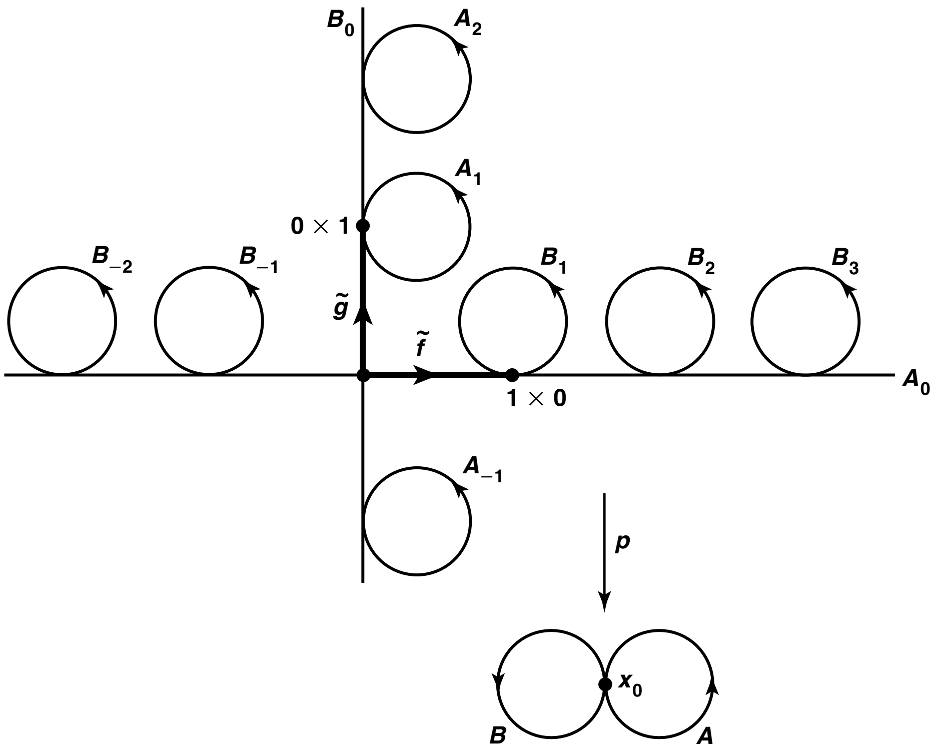


Figure 60.1

Example 3. Let $X = \mathbb{T}^2$, thought of as a square with its opposite sides identified. Let U be an open disc in the middle of the square. Let V be (a small open thickening of) the complement of U. The intersection $U \cap V$ is a circle. We have

- $\pi_1(U)$ is trivial, as U is a disc;

- $\pi_1(U \cap V) = \mathbb{Z}$;
- $\pi_1(V) = \langle a, b \rangle$, as $V \simeq S^1 \vee S^1$.

To see $V \simeq S^1 \vee S^1$, note that the square minus a disc is homotopy equivalent to the boundary of the square, which becomes a wedge of two circles in the quotient space.

Van Kampen's theorem then tells us

$$\pi_1(\mathbb{T}^2) = \langle a, b \mid i_*(c) = j_*(c) \rangle,$$

where

- $c \in \mathbb{Z}$ is a generator,
- $i_* : \mathbb{Z} \to \{1\}$ is the trivial map,
- $j_* : \mathbb{Z} \to \langle a, b \rangle$ sends a generator for $\mathbb{Z}$ to some word in a, b

This amalgamated relation $j_*(c) = 1$ is equivalent to the boundary of the circle is homotopic to the loop $b^{-1}a^{-1}ba = 1$ in V. Therefore the amalgamated relation for the generator $c \in \mathbb{Z}$ is $b^{-1}a^{-1}ba = 1$.

This relation is equivalent to $ab = ba$, so the group we get is just the abelian group $\mathbb{Z}^2 = \mathbb{Z} \times \mathbb{Z}$.