8. Computations and covering spaces

...it is necessary, in order to affirm that a manifold is simply-connected, to study its fundamental group*, ...*

J. Henri Poincaré, 1904

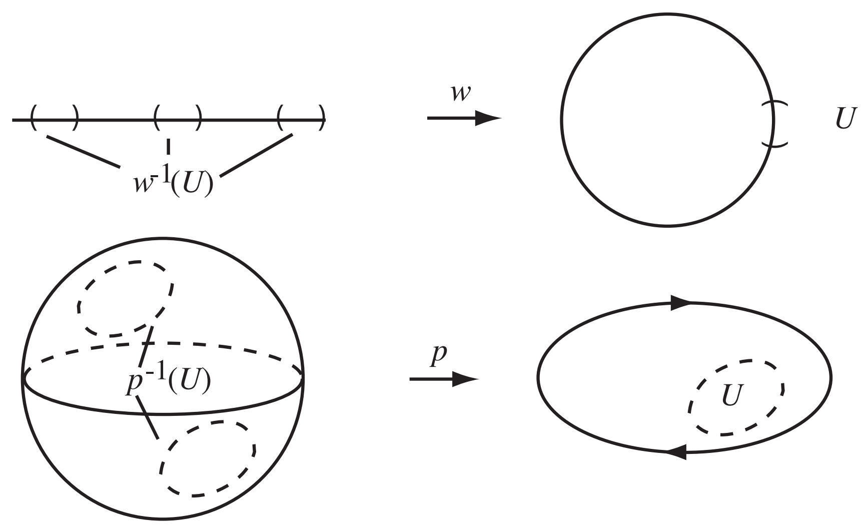
We have defined the fundamental group and showed that it is a topological invariant, that is, homeomorphic spaces have isomorphic fundamental groups. But we have yet to consider a space whose fundamental group is nontrivial. Two familiar spaces, S^1 and $\mathbb{R}\mathrm{P}^2$, will provide examples.

The method of computation focuses on the properties of the mappings,

$$\begin{array}{c} w\colon \mathbb{R} \to S^1 \\ w(r) = \cos(2\pi r) + i\sin(2\pi r) = e^{2\pi i r} \end{array} \quad \text{and} \quad \begin{array}{l} p\colon S^2 \to \mathbb{R}\mathrm{P}^2 \\ p(\mathbf{x}) = [\pm\mathbf{x}]. \end{array}$$

These mappings share certain important properties.

Definition 8.1. *Let X be a space. A* **covering space** *of X is a path-connected space $\tilde{X}$ and a mapping $p\colon \tilde{X} \to X$ such that, for every $x \in X$, there is an open, path-connected subset U with $x \in U$ for which each path component of $p^{-1}(U)$ is homeomorphic to U by restriction of the mapping p. Such open sets are called* **elementary neighborhoods***.*



For example, if $e^{i\theta} \in S^1$, then for $0 < \epsilon < \pi$, the open set $U = \{e^{i\alpha} \mid \theta - \epsilon < \alpha < \theta + \epsilon\}$ in S^1 has inverse image under w given by

$$w^{-1}(U) = \bigcup_{k\in\mathbb{Z}} \left(\frac{\theta}{2\pi} - \frac{\epsilon}{2\pi} + k, \frac{\theta}{2\pi} + \frac{\epsilon}{2\pi} + k \right).$$

Since $\epsilon/2\pi < 1/2$, the intervals in the union are all disjoint. Furthermore, w restricted to any one of these intervals has an inverse given by a branch of the logarithm. In the case of the quotient map $p\colon S^2 \to \mathbb{R}\mathrm{P}^2$, for a connected open set $V \subset S^2$ satisfying $V \cap -V = \emptyset$, we have $p(V)$ open in $\mathbb{R}\mathrm{P}^2$ and $p^{-1}(p(V)) = V \cup -V$. Since the components of $p^{-1}(p(V))$ are V and $-V$, it is an elementary neighborhood. For any $[\pm\mathbf{x}] \in \mathbb{R}\mathrm{P}^2$, there is such an elementary neighborhood containing $[\pm\mathbf{x}]$ and so $p\colon S^2 \to \mathbb{R}\mathrm{P}^2$ is a covering space.

Henceforth we will assume that all spaces are path-connected and locally path-connected to avoid pathological cases. The most useful property of covering spaces is the ability to lift paths in X to paths in $\tilde{X}$ while preserving the homotopy relation.

LEMMA 8.2. *Let $p: \tilde{X} \to X$ be a covering space and $\tilde{x}_0 \in \tilde{X}$ with $p(\tilde{x}_0) = x_0 \in X$. If $\lambda: [0,1] \to X$ is any path with $\lambda(0) = x_0$, then there exists a unique path $\hat{\lambda}: [0,1] \to \tilde{X}$ with $\hat{\lambda}(0) = \tilde{x}_0$ and $p \circ \hat{\lambda} = \lambda$.*

Proof: Cover X by elementary neighborhoods. If $\lambda([0,1]) \subset U$ for some elementary neighborhood, then $x_0 \in U$ and $\tilde{x}_0 \in p^{-1}(U)$. It follows that $\tilde{x}_0$ lies in some component C_0 of $p^{-1}(U)$ that is homeomorphic to U via $p|_{C_0}: C_0 \to U$. Let $(p|_{C_0})^{-1}: U \to C_0$ denote the inverse of this homeomorphism and let $\hat{\lambda} = (p|_{C_0})^{-1} \circ \lambda$. Then $\hat{\lambda}(0) = (p|_{C_0})^{-1}(x_0) = \tilde{x_0}$, since $\tilde{x_0}$ is the only point in $\tilde{X}$ corresponding to x_0 in this component. Finally, $p \circ \hat{\lambda} = p \circ (p|_{C_0})^{-1} \circ \lambda = \lambda$.

If $\lambda([0,1]) \not\subset U$, consider the collection $\{\lambda^{-1}(U') \subset [0,1] \mid U'$, an elementary neighborhood$\}$. This is a cover of $[0,1]$, which is a compact metric space, and so by Lebesgue's Lemma we can choose $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = 1$ with each $\lambda([t_{i-1}, t_i])$ a subset of some elementary neighborhood (take $t_i - t_{i-1} < \delta$, the Lebesgue number). Using the argument above, lift λ on $[0, t_1]$. Then take $\lambda(t_1)$ as x_0 and $\hat{\lambda}(t_1)$ as $\tilde{x}_0$ and lift λ to $[t_1, t_2]$. Continuing in this manner, we construct $\hat{\lambda}$ on $[0,1]$ with $\hat{\lambda}(0) = \tilde{x}_0$ and $p \circ \hat{\lambda} = \lambda$.

To show that $\hat{\lambda}$ constructed in this manner is unique, we prove a more general result that implies uniqueness.

LEMMA 8.3. *Let $p: \tilde{X} \longrightarrow X$ be a covering space and Y, a connected, locally connected space. Given two mappings f_1, $f_2: Y \to \tilde{X}$ with $p \circ f_1 = p \circ f_2$, then the set $\{y \in Y \mid f_1(y) = f_2(y)\}$ is either empty or all of Y.*

Proof: Consider the subset of Y given by $B = \{y \in Y \mid f_1(y) = f_2(y)\}$. We show that B is both open and closed. If $y \in \text{cls } B$, consider $x = p \circ f_1(y) = p \circ f_2(y)$ and U an elementary neighborhood containing x. Consider $(p \circ f_1)^{-1}(U) \cap (p \circ f_2)^{-1}(U)$ which contains y. Because Y is locally connected, there is an open set W for which $y \in W \subset (p \circ f_1)^{-1}(U) \cap (p \circ f_2)^{-1}(U)$ with W connected. Then $f_1(W)$ and $f_2(W)$ are connected subsets of $p^{-1}(U) \subset \tilde{X}$. Since W is open and $y \in \text{cls } B$, there is a point $z \in W$ with $z \in B$. Thus $f_1(z) = f_2(z)$ and $f_1(W) \cap f_2(W) \neq \emptyset$; therefore, $f_1(W)$ and $f_2(W)$ must lie in the same component of $p^{-1}(U)$. Since $p \circ f_1(y) = p \circ f_2(y)$ and the component in which we find both $f_1(y)$ and $f_2(y)$ is homeomorphic to U by the restriction of p, we have $f_1(y) = f_2(y)$. Thus $y \in B$ and B is closed.

If we let $y \in B$, the argument above shows that the sets $f_1(W)$ and $f_2(W)$ lie in the same component C_0 of $p^{-1}(U)$. It follows that, for all $w \in W$,

$$f_1(w) = (p|_{C_0})^{-1} \circ p \circ f_1(w) = (p|_{C_0})^{-1} \circ p \circ f_2(w) = f_2(w)$$

and so W is contained in B. Thus B is open.

The only subsets of Y that are both open and closed are Y itself and $\emptyset$ and so we have proved the lemma. ♢

Using Lemma 8.3, two lifts of a path $\lambda: [0,1] \to X$ which begin at the same point in $\tilde{X}$ must be the same lift. This is the uniqueness part of Lemma 8.2. ♢

Having lifted paths in X to paths in $\tilde{X}$, we next lift certain homotopies between paths.

LEMMA 8.4. *Let* $p\colon \tilde{X} \to X$ *be a covering space and* η_0, $\eta_1\colon [0,1] \to \tilde{X}$ *be two paths in* $\tilde{X}$ *with* $\eta_0(0) = \eta_1(0) = \tilde{x}_0$. *If* $p\circ\eta_0(1) = x_1 = p\circ\eta_1(1)$ *and* $p\circ\eta_0 \simeq p\circ\eta_1$ *via a homotopy that fixes the endpoints of the paths in* X, *then* $\eta_1 \simeq \eta_2$ *in* $\tilde{X}$ *and, in particular,* $\eta_0(1) = \eta_1(1)$.

Proof: Let $H\colon [0,1]\times[0,1] \to X$ be a homotopy between $p\circ\eta_0$ and $p\circ\eta_1$. In this case, we have, for all $s, t \in [0,1]$,

$$\begin{array}{c} H(s,0) = p\circ\eta_0(s) \\ H(s,1) = p\circ\eta_1(s) \end{array} \quad \text{and} \quad \begin{array}{c} H(0,t) = p(\tilde{x}_0) \\ H(1,t) = p\circ\eta_0(1) = p\circ\eta_1(1). \end{array}$$

Since $[0,1]\times[0,1]$ is a compact metric space, when we cover it by the collection $\{H^{-1}(U) \mid U$, an elementary neighborhood of $X\}$, we can apply Lebesgue's Lemma to get $\delta > 0$ for which any subset of $[0,1]\times[0,1]$ of diameter $< \delta$ lies in some $H^{-1}(U)$. If we subdivide the interval $[0,1]$,

$$0 = s_0 < s_1 < \cdots < s_{m-1} < s_m = 1 \quad \text{and} \quad 0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = 1$$

so that $s_i - s_{i-1} < \delta/2$ and $t_j - t_{j-1} < \delta/2$, then H maps each subrectangle $[s_{i-1}, s_i] \times [t_{j-1}, t_j]$ into an elementary neighborhood for all i and j.

To construct the lifting $\hat{H}\colon [0,1]\times[0,1] \to \tilde{X}$, and show it is a homotopy between η_0 and η_1, begin by lifting H on $[0,s_1]\times[0,t_1]$ to $\tilde{X}$ by using $\hat{H} = (p|_{C_{11}})^{-1}\circ H$, where C_{11} is the component of $p^{-1}(U_{11})$ containing $\eta_0(0)$ and $H([0,s_1]\times[0,t_1]) \subset U_{11}$, an elementary neighborhood. Having done this, extend $\hat{H}$ next to $[s_1,s_2]\times[0,t_1]$. Notice that $\hat{H}$ has been defined on the line segment $\{s_1\}\times[0,t_1]$ which is connected and this determines the component of $p^{-1}(U_{21})$ for the elementary neighborhood U_{21} which contains $H([s_1,s_2]\times[0,t_1])$. Once the component, say C_{21}, is determined, extend $\hat{H}$ by $\hat{H} = (p|_{C_{21}})^{-1}\circ H$. Continue in this manner until $\hat{H}$ is defined on $[0,1]\times[0,t_1]$. Next, extend to $[0,1]\times[t_1,t_2]$ using the fact that the value of $\hat{H}$ has been determined on each succesive subrectangle along the left and bottom edges, as a connected subset. Continue along each row until $\hat{H}$ is defined on $[0,1]\times[0,1]$. By Lemma 8.3, $\hat{H}$ is unique fulfilling the condition $\hat{H}(0,0) = \eta(0)$. Since $\eta_0(s)$ is also a lift of $H(s,0)$, we have that $\hat{H}(s,0) = \eta_0(s)$. The condition $H(0,t) = p\circ\eta_0(0)$ implies that $\hat{H}(0,t) = \eta_0(0)$, that is, the homotopy $\hat{H}$ is constant on the subset $\{0\}\times[0,1]$. Thus, the lift $\hat{H}(s,1)$ of the path $p\circ\eta_1(s)$ in X begins at $\eta_0(0) = \eta_1(0)$, and $\eta_1(s)$ is also such a lift. By uniqueness, $\hat{H}(s,1) = \eta_1(s)$. Finally, $H(1,t) = p\circ\eta_0(1) = p\circ\eta_1(1)$ for all $t \in [0,1]$, $\hat{H}(1,t) = \eta_0(1)$ and we conclude that $\eta_0(1) = \eta_1(1)$ since $\hat{H}(1,t)$ is constant. ♢

Uniqueness of liftings of homotopies provides considerable control over the fundamental group through a covering space, giving us a toehold for computation.

COROLLARY 8.5. *Suppose* $p\colon \hat{X} \to X$ *is a covering space:* (1) *If* $\eta\colon [0,1] \to \tilde{X}$ *is a loop at* $\tilde{x}_0$ *and* $p\circ\eta$ *is homotopic to the constant loop* c_{x_0} *for* $x_0 = p(\tilde{x}_0)$, *then* $\eta \simeq c_{\tilde{x}_0}$. (2) *The induced homomorphism* $p_*\colon \pi_1(\tilde{X},\tilde{x}_0) \to \pi_1(X,x_0)$ *is injective.* (3) *For all* $x \in X$, *the subsets* $p^{-1}(\{x\})$ *of* $\tilde{X}$ *have the same cardinality.*

Proof: (1) One lift of c_{x_0} is simply the constant path $c_{\tilde{x}_0}$. By Lemma 8.4 $p\circ\eta \simeq p\circ c_{\tilde{x}_0} = c_{x_0}$ implies $\eta \simeq c_{\tilde{x}_0}$.

(2) If $p_*([\lambda]) = p_*([\mu])$, then, because p_* is a homomorphism, $p_*([\lambda] * [\mu^{-1}]) = [c_{x_0}]$, that is, $p \circ (\lambda * \mu^{-1}) \simeq c_{x_0}$. By (1) $\lambda * \mu^{-1} \simeq c_{\tilde{x}_0}$ or $\lambda \simeq \mu$, that is, $[\lambda] = [\mu]$.

(3) Suppose x_0 and x_1 are in X and $\lambda\colon [0,1] \to X$ is a path joining x_0 to x_1. Suppose $y \in p^{-1}(\{x_0\})$. We define a mapping $\Lambda\colon p^{-1}(\{x_0\}) \to p^{-1}(\{x_1\})$ by lifting λ to $\lambda_y\colon [0,1] \to \tilde{X}$ with $\lambda_y(0) = y$. Define $\Lambda(y) = \lambda_y(1)$. Since λ_y is uniquely determined by being a lift of $p \circ \lambda_y = \lambda$ with $\lambda_y(0) = y$, the function Λ is well-defined. By Lemma 8.3, lifts of λ beginning at different elements in $p^{-1}(\{x_0\})$ must end at different points in $p^{-1}(\{x_1\})$ and so Λ is injective. Using lifts of λ^{-1} we deduce that Λ is surjective.(Notice that a different choice of λ might give a different one-one correspondence Λ.) $\diamondsuit$

For $w\colon \mathbb{R} \to S^1$, $w(r) = e^{2\pi i r}$, we find that $w^{-1}(1 + 0i) = \mathbb{Z} \subset \mathbb{R}$ and so $w^{-1}(\{z\})$ is countably infinite for each $z \in S^1$. For $p\colon S^2 \to \mathbb{R}\mathrm{P}^2$, $p^{-1}(\{[\pm \mathbf{x}_0]\})$ contains two elements, $\mathbf{x}_0$ and $-\mathbf{x}_0$. In general, if we lift a loop $\omega\colon [0,1] \to X$ at x_0 in X, the proof of (3) of Corollary 8.5 obtains a mapping $\Omega\colon p^{-1}(\{x_0\}) \to p^{-1}(\{x_0\})$ by lifting the loop. By remark (1) of the corollary, if Ω is nontrivial, then the loop ω is not homotopic to the constant map. This observation is enough to prove the following.

Theorem 8.6. **A.** $\pi_1(S^1) \cong \mathbb{Z}$. **B.** $\pi_1(\mathbb{R}\mathrm{P}^2) \cong \mathbb{Z}/2\mathbb{Z}$.

Proof of **A***:* If $\beta\colon [0,1] \to S^1$ is any loop at $1 \in S^1$, then the lift of β to $\hat{\beta}\colon [0,1] \to \mathbb{R}$ satisfies $\hat{\beta}(1) \in \mathbb{Z}$. The properties of liftings determine a function $\Xi\colon \pi_1(S^1) \to \mathbb{Z}$ given by $[\beta] \mapsto \hat{\beta}(1)$.

Let $\alpha\colon [0,1] \to S^1$ given by $\alpha(t) = (\cos(2\pi t), \sin(2\pi t))$. Since $\alpha = w|_{[0,1]}$, we see that one lift of α to $\mathbb{R}$ is just the identity and $\hat{\alpha}(1) = 1$. It follows that α is not homotopic to the constant map at 1, c_1. Next consider α^n for $n \in \mathbb{Z}$, given by $\alpha^n(t) = (\cos(2\pi n t), \sin(2\pi n t))$. By the same argument for α, $\hat{\alpha}^n(1) = n$ and so the mapping $\Xi\colon \pi_1(S^1) \to \mathbb{Z}$ is surjective. Since each $\alpha^n \not\simeq c_1$ for $n \neq 0$, the subgroup generated by $[\alpha]$, isomorphic to $\mathbb{Z}$, is a subgroup of $\pi_1(S^1)$.

To finish the proof of **A**, we show that if β is any loop based at 1 in S^1, then $\beta \simeq \alpha^n$ for some $n \in \mathbb{Z}$. Let $U_1 = \{(x,y) \in S^1 \mid y > -1/10\}$, and $U_2 = \{(x,y) \in S^1 \mid y < 1/10\}$. The pair $\beta^{-1}(U_1)$, $\beta^{-1}(U_2)$ is an open cover of $[0,1]$ and by Lebesgue's Lemma we can subdivide $[0,1]$ as $0 = t_0 < t_1 < \ldots < t_{m-1} < t_m = 1$ so that

i) $\beta([t_i, t_{i+1}]) \subset U_1$ or $\beta([t_i, t_{i+1}]) \subset U_2$ for $0 \leq i < m$.

Form the union of consecutive subintervals when both are mapped to the same U_j $j = 1$ or 2. In detail, let $s_0 = 0$ and $s_1 = t_{i_1}$ where $\beta([0, t_{i_1}] \subset U_{j_1}$ for j_1 is one of 1 or 2 and $\beta([t_{i_1}, t_{i_1+1}]) \not\subset U_{j_1}$. Let $U_{j_2} \neq U_{j_1}$ and $\beta([t_{i_1}, t_{i_1+1}]) \subset U_{j_2}$. Then let $s_2 = t_{i_2}$ where $\beta([t_{i_1}, t_{i_2}]) \subset U_{j_2}$ but $\beta([t_{i_2}, t_{i_2+1}]) \not\subset U_{j_2}$. Continue in this manner to get

$$0 = s_0 < s_1 < \cdots < s_{k-1} < s_k = 1$$

so that

ii) $\beta([s_{j-1}, s_j])$ and $\beta([s_j, s_{j+1}])$ are not both contained in the same U_k, for $k = 1, 2$.

Let $\beta_j\colon [0,1] \to S^1$ denote the reparametrization of $\beta|_{[s_j, s_{j+1}]}$ so that $\beta \simeq \beta_0 * \beta_1 * \ldots * \beta_{k-1}$, and each β_j is a path in exactly one of U_1 or U_2. Furthermore, $\beta(s_j) \in U_1 \cap U_2$, a subspace of two components, one of which contains $1 = e^{2\pi i 0}$ and the other $-1 = e^{\pi i}$. For $0 < j < m$

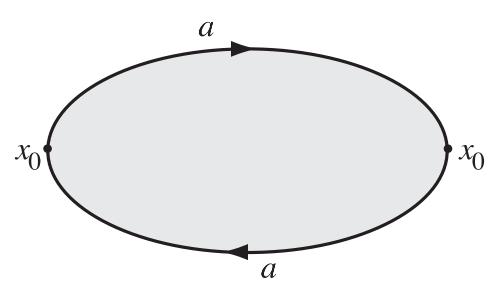
choose a path $\lambda_j\colon [0,1] \to U_1 \cap U_2$ with $\lambda_j(0) = \beta(s_j) = \beta_{j-1}(s_j)$ and $\lambda_j(1) = 1$ or -1, depending on the component. Define

$$\begin{aligned} \gamma_1 &= \beta_0 * \lambda_1 \\ \gamma_j &= \lambda_{j-1}^{-1} * \beta_{j-1} * \lambda_j \text{ for } 1 < j < k \\ \gamma_k &= \lambda_{k-1}^{-1} * \beta_{k-1}. \end{aligned}$$

By canceling $\lambda_j * \lambda_j^{-1}$, $\beta \simeq \gamma_1 * \gamma_2 * \ldots * \gamma_k$. Consider the paths γ_k. If γ_k is a closed path, it lies in U_1 or U_2 which are simply-connected and so $\gamma_k \simeq c_1$ or $\gamma_k \simeq c_{-1}$. If γ_k is not closed, then it crosses between the components of $U_1 \cap U_2$. It follows that $\gamma_k \simeq \eta_1^{\pm 1}$ or $\gamma_k \simeq \eta_2^{\pm 1}$ where $\eta_1(t) = (\cos(\pi t), \sin(\pi t))$, the path joining 1 to -1 in U_1, and $\eta_2(t) = (\cos(\pi t + \pi), \sin(\pi t + \pi))$, the path joining -1 to 1 in U_2. Making the cancellations of the type $\eta_1 \eta_1^{-1} \simeq c_1$ or $\eta_2 \eta_2^{-1} \simeq c_{-1}$, we are left with three possibilities:

$$\beta \simeq c_1, \quad \beta \simeq \eta_1 * \eta_2 * \eta_1 * \eta_2 * \ldots * \eta_1 * \eta_2, \text{ or } \beta \simeq \eta_2^{-1} * \eta_1^{-1} * \eta_2^{-1} * \ldots * \eta_2^{-1} * \eta_1^{-1},$$

after cancelling out $c_{\pm 1}$. The ordering is determined by the fact that β begins and ends at 1, and each γ_k either joins 1 to -1, joins -1 to 1, or it simply stays put. After cancellation of the paths that stay put or products of paths that are homotopic to the constant path, we are left with such a product in that order. Finally, $w|_{[0,1]} = \alpha \simeq \eta_1 * \eta_2$ and so $\beta \simeq \alpha^n$ for some $n \in \mathbb{Z}$. $\diamondsuit$



Proof of **B***:* Consider the model of the projective plane given by the *di-gon*, a disk with a point on the boundary identified with the point symmetric with respect to the origin. Let $x_0 \in \mathbb{R}\mathrm{P}^2$ be the point $x_0 = [\pm(1,0,0)]$. Let $p\colon S^2 \to \mathbb{R}\mathrm{P}^2$ denote the covering space $p(\mathbf{x}) = [\pm \mathbf{x}]$. Let the loop a in $\mathbb{R}\mathrm{P}^2$ denote *half of the equator*, and lift a to S^2. We get a path $\hat{a}$ from $(1,0,0)$ to $(-1,0,0)$ along the equator of S^2. By Corollary 8.5, $a \not\simeq c_{x_0}$. In the di-gon representation of $\mathbb{R}\mathrm{P}^2$, $a * a = a^2$ surrounds the disk, and so a^2 can be contracted to c_{x_0} by shrinking to the center of the disk and translating over to x_0. It follows that $\pi_1(\mathbb{R}\mathrm{P}^2)$ contains $\mathbb{Z}/2\mathbb{Z}$. To finish, we need show that any loop at x_0 is homotopic to a^n for some $n \in \mathbb{Z}$. Using the di-gon we see that away from the image of the path a^2 a path lies in the contractible interior of a disk. The disk can be used to push any loop onto a as often as it crosses between the copies of x_0. Thus we see that any loop based at x_0 is homotopic to a^n for some $n \in \mathbb{Z}$ and so homotopic to a or c_{x_0}. This implies that

$$\pi_1(\mathbb{R}\mathrm{P}^2) = \langle [a] \rangle / ([a]^2 = [c_{x_0}]) \cong \mathbb{Z}/2\mathbb{Z}.$$

This completes the proof of Theorem 8.6. $\diamondsuit$

Covering spaces can be developed much further. We refer the reader to [Massey] or [Lima] for thorough treatments. Let's turn now to applications. We first return to the central question of the text:

Invariance of Dimension for $(2,n)$: *For $n \neq 2$, $\mathbb{R}^n$ and $\mathbb{R}^2$ are not homeomorphic.*

Proof: We assume that $n \geq 2$ since the case of $n = 1$ is covered in Chapter 5. If $\mathbb{R}^n \cong \mathbb{R}^2$, then, by composing with a translation if needed, we can choose a homeomorphism $f\colon \mathbb{R}^n \to \mathbb{R}^2$ for which $f(\mathbf{0}) = (0,0)$. Such a mapping induces a homeomorphism $\mathbb{R}^n - \{\mathbf{0}\} \cong \mathbb{R}^2 - \{(0,0)\}$. Since S^{n-1} is a deformation retract of $\mathbb{R}^n - \{\mathbf{0}\}$, by Theorem 7.10, $\pi_1(\mathbb{R}^n - \{\mathbf{0}\}) \cong \pi_1(S^{n-1})$. For $n > 2$, Corollary 7.13 states that $\pi_1(S^{n-1}) \cong \{e\}$, while, for $n = 2$, $\pi_1(S^1) \cong \mathbb{Z}$. Since the fundamental group is a topological invariant, it must be the case that $n = 2$. ♢

This argument is characteristic of the power of introducing algebraic structures as topological invariants of spaces. Our goal in later chapters is to generalize these ideas.

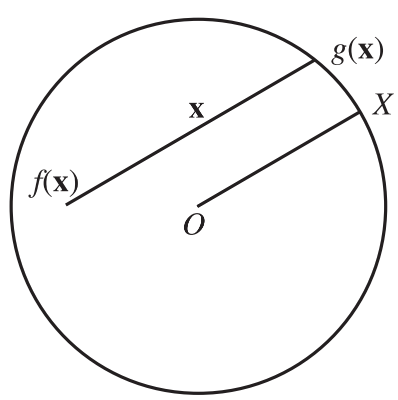
Recall the somewhat unexpected topological property introduced in the exercises of Chapter 2: A space X has the *fixed point property* (**FPP**) if any continuous mapping $f\colon X \to X$ has a fixed point, that is, there exists a point $x_0 \in X$ with $f(x_0) = x_0$. By the Intermediate Value Theorem we can prove that the interval $[0,1]$ has the FPP: if $f\colon [0,1] \to [0,1]$ is continuous, then define $g(x) = f(x) - x\colon [0,1] \to \mathbb{R}$. If $f(0) \neq 0$ and $f(1) \neq 1$, then $g(0) > 0$ and $g(1) < 0$ and so g must take the value 0 somewhere. If $g(x) = 0$, then $f(x) = x$.

What is the generalization of the space $[0,1]$ to higher dimensions? Candidates include $[0,1] \times [0,1]$ in dimension 2 or maybe the **two-disk** $e^2 = \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| \leq 1\} = \operatorname{cls} B(\mathbf{0}, 1)$. The choice between these two candidates is irrelevant since the fixed point property is a topological property and they are homeomorphic. (Can you prove it?) We generalize the fixed point property for the interval $[0,1]$ to the two-disk.

Theorem 8.7. (Brouwer's Theorem in dimension 2). *The two-disk $e^2 = \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| \leq 1\} \subset \mathbb{R}^2$ has the fixed point property.*

Proof: Suppose $f\colon e^2 \to e^2$ is a continuous function without a fixed point. Then for each $\mathbf{x} \in e^2$, $f(\mathbf{x}) \neq \mathbf{x}$. Define $g\colon e^2 \to S^1$ by

$$g(\mathbf{x}) = \text{ intersection of the ray from } f(\mathbf{x}) \text{ to } \mathbf{x} \text{ with } S^1.$$



To see that $g(\mathbf{x})$ is continuous on e^2, we apply some vector geometry: write $Q = f(\mathbf{x})$, $Z = g(\mathbf{x})$. Let $O = (0,0)$ and define $X = (\mathbf{x} - Q)/\|\mathbf{x} - Q\|$. Then, $g(\mathbf{x}) = Z = Q + tX$ for some $t \geq 0$ for which $Q + tX \in S^1$, that is, $(Q + tX) \cdot (Q + tX) = 1$. This condition can be rewritten to solve for t, namely,

$$(Q + tX) \cdot (Q + tX) = t^2(X \cdot X) + 2t(Q \cdot X) + Q \cdot Q = 1.$$

The quadratic formula gives

$$\begin{aligned} t_{\mathbf{x}} &= -Q \cdot X + \sqrt{(Q \cdot X)^2 + 1 - Q \cdot Q} \\ &= -f(\mathbf{x}) \cdot \frac{\mathbf{x} - f(\mathbf{x})}{\|\mathbf{x} - f(\mathbf{x})\|} + \sqrt{\left(f(\mathbf{x}) \cdot \frac{\mathbf{x} - f(\mathbf{x})}{\|\mathbf{x} - f(\mathbf{x})\|}\right)^2 + 1 - f(\mathbf{x}) \cdot f(\mathbf{x})}. \end{aligned}$$

Note that this choice of signs gives $t_{\mathbf{x}} \geq 0$ and $t_{\mathbf{x}} = 0$ implies $f(\mathbf{x}) = \mathbf{x}$. Since we have assumed that this doesn't happen, $t_{\mathbf{x}} > 0$. Furthermore, $t_{\mathbf{x}}$ is a continuous function of $\mathbf{x}$. We can write $g(\mathbf{x})$ explicitly as

$$g(\mathbf{x}) = f(\mathbf{x}) + t_{\mathbf{x}} \frac{\mathbf{x} - f(\mathbf{x})}{\|\mathbf{x} - f(\mathbf{x})\|}.$$

and so $g(\mathbf{x})$ is continuous.

By the definition of the mapping g, if $\mathbf{x} \in S^1 \subset e^2$, then $g(\mathbf{x}) = \mathbf{x}$. We have constructed a continuous mapping $g\colon e^2 \to S^1$ for which $g \circ i = \mathrm{id}_{S^1}$, that is, the identity mapping on S^1 can be factored:

$$\mathrm{id}_{S^1}\colon S^1 \xrightarrow{i} e^2 \xrightarrow{g} S^1.$$

This composite leads to a composite of group homomorphisms and fundamental groups:

$$\mathrm{id}\colon \pi_1(S^1) \xrightarrow{i_*} \pi_1(e^2) \xrightarrow{g_*} \pi_1(S^1).$$

However, $\pi_1(e^2) = \{[c_1]\}$ and so $g_* \circ i_*([\alpha]) = [c_1] \neq [\alpha]$ and $g_* \circ i_* \neq \mathrm{id}$, a contradiction. Therefore, a continuous function $f\colon e^2 \to e^2$ without fixed points is not possible. ♢

Corollary 8.8. *S^1 is not a retract of e^2.*

More powerful tools will be developed in later chapters to prove a generalization of Theorem 8.7 and its corollary. Brouwer proved this general result around 1911 [11].

We next apply the fact that $\pi_1(\mathbb{R}\mathrm{P}^2) \cong \mathbb{Z}/2\mathbb{Z}$. Recall that $\mathbb{R}\mathrm{P}^2$ is the space of lines through the origin in $\mathbb{R}^3$. The lower dimensional analogue is the space $\mathbb{R}\mathrm{P}^1$ consisting of lines through the origin in $\mathbb{R}^2$. We can identify a line with the angle it makes with the x-axis. To obtain every line through the origin, we only need angles $0 \leq \theta \leq \pi$ where the x-axis is identified with the angles 0 and π. Hence $\mathbb{R}\mathrm{P}^1 \cong [0, \pi]/(0 \sim \pi) \cong S^1$. Thus $\pi_1(\mathbb{R}\mathrm{P}^1) \cong \mathbb{Z}$. The analogue of the covering map $p\colon S^2 \to \mathbb{R}\mathrm{P}^2$ in this case is $\overline{p}\colon S^1 \to \mathbb{R}\mathrm{P}^1$ given by $e^{2\pi i\theta} \mapsto [\pm e^{2\pi i\theta}]$. In fact, $\overline{p}_*\colon \pi_1(S^1) \to \pi_1(\mathbb{R}\mathrm{P}^1)$ is described as a homomorphism $\mathbb{Z} \to \mathbb{Z}$ given by multiplication by two, because the generator $[\alpha]$ wraps around $\mathbb{R}\mathrm{P}^1$ twice.

In Chapter 5 we proved that a continuous mapping $f: S^1 \to \mathbb{R}$ must send some point and its negative to the same value, that is, there is always a point $x_0 \in S^1$ with $f(x_0) = f(-x_0)$. We can generalize that result to S^2.

THEOREM 8.9. *If $f: S^2 \to \mathbb{R}^2$ is a continuous function, then there exists a point $\mathbf{x} \in S^2$ with $f(\mathbf{x}) = f(-\mathbf{x})$.*

We proceed by proving an associated result.

PROPOSITION 8.10. (THE BORSUK-ULAM THEOREM FOR $n = 2$.) *There does not exist a continuous function $f: S^2 \to S^1$ that satisfies $f(-\mathbf{x}) = -f(\mathbf{x})$ for all $\mathbf{x} \in S^2$.*

Proof of the Borsuk-Ulam theorem: Assume such a function exists. The condition satisfied by f can be written $f(\pm\mathbf{x}) = \pm f(\mathbf{x})$. It follows that f induces $\hat{f}: \mathbb{RP}^2 \to \mathbb{RP}^1$ and $\hat{f}$ fits into a diagram:

$$\begin{array}{ccc} S^2 & \xrightarrow{f} & S^1 \\ \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle \bar{p}} \\ \mathbb{RP}^2 & \xrightarrow[\hat{f}]{} & \mathbb{RP}^1. \end{array} \quad \text{for which } \bar{p} \circ f = \hat{f} \circ p.$$

Consider the induced homomorphism $\hat{f}_*: \pi_1(\mathbb{RP}^2) \to \pi_1(\mathbb{RP}^1)$. By Theorem 8.6, $\hat{f}_*$ is a homomorphism $\mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}$. However, any such homomorphism must be the trivial homomorphism. (Do you know why?) Let $\lambda: [0,1] \to S^2$ denote a path from the north pole to the south pole along a meridian of constant longitude. It follows that $[p \circ \lambda] = [\alpha]$, a generator for $\mathbb{Z}/2\mathbb{Z} \cong \pi_1(\mathbb{RP}^2)$. Since the north and south pole are antipodal, these points are identified in $\mathbb{RP}^1$ after passage through f and $\bar{p}$. Hence $[\bar{p} \circ f \circ \lambda]$ is nontrivial in $\pi_1(\mathbb{RP}^1)$. But $[\bar{p} \circ f \circ \lambda] = [\hat{f} \circ p \circ \lambda] = \hat{f}_*([p \circ \lambda]) = 0$, a contradiction. ♢

COROLLARY 8.11. *If $f: S^2 \to \mathbb{R}^2$ is a continuous function such that $f(-\mathbf{x}) = -f(\mathbf{x})$ for all $\mathbf{x} \in S^2$, then $f(\mathbf{x}) = (0,0)$ for some $\mathbf{x} \in S^2$.*

Proof: If not, then $g(\mathbf{x}) = f(\mathbf{x})/\|f(\mathbf{x})\|$ would be a continuous function $g: S^2 \to S^1$ with $g(-\mathbf{x}) = -g(\mathbf{x})$ for all $\mathbf{x} \in S^2$. ♢

Proof of Theorem 8.9: Suppose for every $\mathbf{x} \in S^2$, that $f(\mathbf{x}) \neq f(-\mathbf{x})$. Then define $g(\mathbf{x}) = f(\mathbf{x}) - f(-\mathbf{x})$. Notice that g is continuous, $g(-\mathbf{x}) = -g(\mathbf{x})$, and $g(\mathbf{x}) \neq \mathbf{0}$ for all $\mathbf{x} \in S^2$, a contradiction. ♢

COROLLARY 8.12. *No subset of $\mathbb{R}^2$ is homeomorphic to S^2.*

The corollary tells us that there is no cartographic map homeomorphic to the entire sphere.

Finally, we derive an unexpected corollary of our analysis of the fundamental group of the circle, namely, the Fundamental Theorem of Algebra. This topological proof gives a complete proof avoiding the difficulties in the approach of Gauss in Chapter 5 based on connectedness.

THE FUNDAMENTAL THEOREM OF ALGEBRA. *If $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ is a polynomial with complex coefficients, then there is a complex number z_0 with $p(z_0) = 0$.*

Proof: Recall that $\mathbb{C} \cong \mathbb{R}^2$ and the nth power mapping $h: z \mapsto z^n$ induces a mapping $h: S^1 \to S^1$ which can be written as $e^{i\theta} \mapsto e^{in\theta}$. Lifting this mapping to the covering space

$w\colon \mathbb{R} \to S^1$, it represents $n \in \mathbb{Z} \cong \pi_1(S^1)$ via the identification of $\pi_1(S^1)$ with $\mathbb{Z}$ given by $[\beta] \mapsto \hat{\beta}(1)$.

Viewed as a mapping, $h\colon S^1 \to S^1$, h induces the homomorphism $h_*\colon \pi_1(S^1) \to \pi_1(S^1)$. The law of exponents implies that $h_*(\theta \mapsto e^{\pi i m\theta}) = (\theta \mapsto (e^{\pi i m\theta})^n = e^{\pi i n m\theta})$, that is, h_* is multiplication by n.

We first consider a special case of the theorem—suppose

$$|a_{n-1}| + |a_{n-2}| + \cdots + |a_0| < 1.$$

Suppose $p(z)$ has no root in $e^2 = \{z \in \mathbb{C} \mid |z| \leq 1\}$. Define the mapping $\hat{p}\colon e^2 \to \mathbb{R}^2 - \{\mathbf{0}\}$ by $\hat{p}(z) = p(z)$. Restricting to $S^1 = \partial e^2$ we get $\hat{p}|\colon S^1 \to \mathbb{R}^2 - \{\mathbf{0}\}$. Since $\hat{p}|$ can be extended to e^2, it follows (exercise) that $\hat{p}|$ is homotopic to a constant map. However, consider the mapping

$$F(z,t) = z^n + t(a_{n-1}z^{n-1} + \cdots + a_0),$$

which gives a homotopy between $F(z,0) = z^n$ and $F(z,1) = p(z)$. If $F(z,t)$ never vanishes on S^1, the homotopy implies $\hat{p}| \simeq z^n$. To establish this condition, we estimate for $|z| = 1$,

$$\begin{aligned} |F(z,t)| &\geq |z^n| - |t(a_{n-1}z^{n-1} + \cdots + a_0)| \\ &\geq 1 - t(|a_{n-1}z^{n-1}| + \cdots + |a_0|) \\ &= 1 - t(|a_{n-1}| + \cdots + |a_0|) > 0. \end{aligned}$$

As a class in $\pi_1(S^1)$, $[(z \mapsto z^n)]$ is not homotopic to the constant map while $\hat{p}|$ is, so we get a contradiction.

To reduce the general case to this special case, let $t \in \mathbb{R}$, $t \neq 0$, and let $u = tz$. So

$$\begin{aligned} p(u) &= u^n + a_{n-1}u^{n-1} + \cdots + a_1 u + a_0 \\ &= (tz)^n + a_{n-1}(tz)^{n-1} + \cdots + a_1 tz + a_0. \end{aligned}$$

If $p(u) = 0$, then

$$z^n + \frac{a_{n-1}}{t}z^{n-1} + \cdots + \frac{a_1}{t^{n-1}}z + \frac{a_0}{t^n} = 0.$$

So given a zero for $p(u)$ we get a zero for $\tilde{p}_t(z)$ with $\tilde{p}_t(z) = z^n + \dfrac{a_{n-1}}{t}z^{n-1} + \cdots + \dfrac{a_0}{t^n}$ and vice versa. Taking t large enough we can guarantee

$$\left|\frac{a_{n-1}}{t}\right| + \cdots + \left|\frac{a_1}{t^{n-1}}\right| + \left|\frac{a_0}{t^n}\right| < 1$$

and we can apply the special case. $\diamondsuit$

In Chapter 7 we proved that a subspace A of a space X, which is a deformation retract of X, shares the same fundamental group as X . Furthermore, if X and Y are homeomorphic spaces, they share the same fundamental group. We generalize these conditions to identify an important relation between spaces.

Definition 8.13. *Two spaces are* **homotopy equivalent***, denoted* $X \simeq Y$*, if there are mappings* $f : X \to Y$ *and* $g : Y \to X$ *with* $g \circ f \simeq \mathrm{id}_Y$ *and* $f \circ g \simeq \mathrm{id}_X$*.*

If $A \subset X$ is a deformation retract, then there is a mapping $r{:}\, X \to A$ for which $\mathrm{id}_A = r \circ i{:}\, A \to A$ and $\mathrm{id}_X \simeq i \circ r{:}\, X \to X$. Thus A is homotopy equivalent to X and homotopy equivalence generalizes the relation of deformation retraction. Contractible spaces are homotopy equivalent to a one-point space so homotopy equivalence is a weaker notion than homeomorphism.

Proposition 8.14. *In a set of topological spaces, homotopy equivalence is an equivalence relation.*

Proof: It suffices to check transitivity since the other properties are clear. Suppose $X \simeq Y$ and $Y \simeq Z$ via mappings $f{:}\, X \to Y$, $g{:}\, Y \to X$; $t{:}\, Y \to Z$ and $u{:}\, Z \to Y$. Consider $t \circ f{:}\, X \to Z$ and $g \circ u{:}\, Z \to X$. Then

$$\begin{aligned} (g \circ u) \circ (t \circ f) &\simeq g \circ (u \circ t) \circ f \\ &\simeq g \circ \mathrm{id}_Y \circ f = g \circ f \simeq \mathrm{id}_X \\ \text{and } (t \circ f) \circ (g \circ u) &\simeq t \circ (f \circ g) \circ u \\ &\simeq t \circ \mathrm{id}_X \circ u = t \circ u \simeq \mathrm{id}_Z. \end{aligned}$$

Fixing a universe, that is, a set in which all relevant spaces are elements, the equivalence class of a space X is called its **homotopy type**. The effectiveness of the fundamental group to distinguish spaces is limited by homotopy equivalence.

Proposition 8.15. *If* X *and* Y *are homotopy-equivalent spaces via mappings* $f{:}\, X \to Y$ *and* $g{:}\, Y \to X$*, then the induced mappings* $f_*{:}\, \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ *and* $g_*{:}\, \pi_1(Y, y_0) \to \pi_1(X, g(y_0))$ *are isomorphisms.*

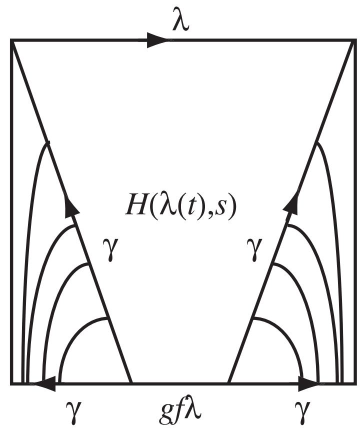
Proof: Let $H{:}\, X \times [0,1] \to X$ be a homotopy between $g \circ f$ and id_X. Let $\gamma{:}\, [0,1] \to X$ be the path $\gamma(t) = H(x_0, t)$, so that $\gamma(0) = g \circ f(x_0)$ and $\gamma(1) = x_0$. We can write the induced homomorphisms:

$$\pi_1(X, x_0) \xrightarrow{f_*} \pi_1(Y, f(x_0)) \xrightarrow{g_*} \pi_1(X, g \circ f(x_0)) \xrightarrow{u_\gamma} \pi_1(X, x_0).$$

We claim that this composite is the identity homomorphism. Consider $[\lambda] \in \pi_1(X, x_0)$. The result of the composite on this element is the following

$$[\lambda] \mapsto [f \circ \lambda] \mapsto [g \circ f \circ \lambda] \mapsto [\gamma^{-1} * (g \circ f \circ \lambda) * \gamma].$$

Apply the homotopy H to get a homotopy from $g \circ f \circ \lambda$ to λ by $H(\lambda(t), s)$. We use this homotopy to construct one from $\gamma^{-1} * (g \circ f \circ \lambda) * \gamma$ to λ by reparametrizising according to the diagram:



In the triangles, we have taken γ and opened it into a triangle with the pictured curves given by isobars (constant paths). It follows from the homotopy that $[\gamma^{-1} * (g \circ f \circ \lambda) * \gamma] = [\lambda]$. This implies that f_* is injective and g_* surjective. To finish the proof consider the composite

$$\pi_1(Y, f(x_0)) \xrightarrow{g_*} \pi_1(X, g \circ f(x_0)) \xrightarrow{f_*} \pi_1(Y, f \circ g \circ f(x_0)) \xrightarrow{u_\eta} \pi_1(Y, f(x_0)),$$

where $\eta\colon [0,1] \to Y$ is the path $\eta(t) = \bar{H}(f(x_0), t)$ in the homotopy $\bar{H}$ between $f \circ g$ and id_Y. The same argument applies *mutatis mutandis* to show that f_* is surjective and g_* is injective and hence both homomorphisms are isomorphisms. $\diamondsuit$

Homotopy equivalence is cruder than homeomorphism but includes it as a special case. To give an idea of how crude homotopy equivalence is, notice, for all n, $\mathbb{R}^n$ is homotopy equivalent to a point. The letters of the alphabet as subspaces of $\mathbb{R}^2$ show other failures to distinguish between different topological spaces.

$$A \simeq D \simeq S^1, B \simeq S^1 \vee S^1, C \simeq E \simeq F \simeq *, \ldots$$

Proposition 8.15 shows that the fundamental group is a **homotopy invariant**, that is, if $X \simeq Y$, then $\pi_1(X) \cong \pi_1(Y)$. Thinking of the fundamental group as a filter that distinguishes spaces, it can only hope to catch homotopy inequivalent spaces. In later chapters we will consider other homotopy invariants. Poincaré [64] introduced the fundamental group to distinguish certain manifolds that were indistinguishable via other combinatorial invariants.

Exercises

1. Suppose that $f\colon S^1 \to S^1$ has an extension $\hat{f}\colon e^2 \to S^1$, that is, the mapping $\hat{f}$ satisfies $\hat{f} \circ i = f$ where $i\colon S^1 \to e^2$ is the inclusion. Show that f is **null-homotopic**, that is, f is homotopic to the constant mapping.

2. Though we will not prove it, one of the useful theorems for computing the fundamental groups of spaces is the **Seifert-van Kampen Theorem** [53]. A special case of this theorem is the following: *If a path-connected space X is a union $X = U \cup V$ with V simply-connected and $x_0 \in U \cap V$, then the inclusion $i\colon U \to X$ induces a surjection $i_*\colon \pi_1(U, x_0) \to \pi_1(X, x_0)$ with kernel given by the smallest subgroup of*

$\pi_1(U, x_0)$ containing $j_*(\pi_1(U \cap V, x_0))$, where $j: U \cap V \hookrightarrow U$ denotes the inclusion. Use the descriptions of $\mathbb{R}\mathrm{P}^2$ of previous chapters and this theorem to make another computation of $\pi_1(\mathbb{R}\mathrm{P}^2)$.

3. Suppose that X is simply-connected and $p: \tilde{X} \to X$ is a covering space of X. Show that p is a homeomorphism.

4. Let $\Omega(X, x_0)$ denote the based loop space of X given by

$$\Omega(X, x_0) = \{\lambda: [0,1] \to X \mid \lambda \text{ is continuous and } \lambda(0) = \lambda(1) = x_0\}.$$

This subspace of $\mathrm{map}(I, X)$ is topologized with the compact-open topology. Show that

i) $\pi_0(\Omega(X, x_0))$, the collection of path-components of $\Omega(X, x_0)$ is in one-to-one correspondence with $\pi_1(X, x_0)$.

ii) Show that the loop multiplication $m: \Omega(X, x_0) \times \Omega(X, x_0) \to \Omega(X, x_0)$ given by $m(\lambda, \mu) = \lambda * \mu$ is a continuous multiplication on $\Omega(X, x_0)$.

5. We know from Theorem 7.15 and Theorem 8.6 that the fundamental group of the torus, $S^1 \times S^1$ is $\mathbb{Z} \times \mathbb{Z}$. Use the argument for the computation of $\pi_1(\mathbb{R}\mathrm{P}^2) \cong \mathbb{Z}/2\mathbb{Z}$ to prove $\pi_1(S^1 \times S^1) \cong \mathbb{Z} \times \mathbb{Z}$ by viewing the torus as a quotient of $[0,1] \times [0,1]$.

6. Let's make a space—take two distinct 2-spheres, S^2, and join them by a line segment—kinda like dumbbells, but with a very thin connector. Denote this space by X and show that it is simply-connected.