

THE HOMOTOPY CATEGORY, AND THE HOMOTOPY GROUPS

1. THE HOMOTOPY CATEGORY

Throughout, $I = [0, 1]$ and all maps are continuous.

Definition 1.1. Let $f, g: X \rightarrow Y$ be maps. A *homotopy* from f to g is a map

$$h: X \times I \longrightarrow Y \quad \text{with} \quad h(x, 0) = f(x), \quad h(x, 1) = g(x)$$

for all $x \in X$. We write $h_t(x) = h(x, t)$, so that a homotopy is a family of maps h_t depending continuously on t and interpolating between $h_0 = f$ and $h_1 = g$; and we write $f \simeq g$, read *f is homotopic to g*, when such an h exists.

If $A \subseteq X$ and $h(a, t) = f(a)$ for all $a \in A$ and all t , the homotopy is said to be *rel A*, and we write $f \simeq g \text{ rel } A$.

Proposition 1.2. *For all spaces X, Y , homotopy is an equivalence relation on the set of maps $X \rightarrow Y$. Moreover, it is compatible with composition: if $f \simeq f': X \rightarrow Y$ and $g \simeq g': Y \rightarrow Z$ then $g \circ f \simeq g' \circ f'$.*

Proof. Reflexive. $h(x, t) = f(x)$ is a homotopy from f to f .

Symmetric. If h is a homotopy from f to g then $\bar{h}(x, t) = h(x, 1 - t)$ is one from g to f .

Transitive. If h is a homotopy from f to g and k is one from g to j , define

$$(h * k)(x, t) = \begin{cases} h(x, 2t), & 0 \leq t \leq \frac{1}{2}, \\ k(x, 2t - 1), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

The two formulas give continuous functions on the closed subsets $X \times [0, \frac{1}{2}]$ and $X \times [\frac{1}{2}, 1]$, respectively, and agree on the intersection $X \times \{\frac{1}{2}\}$, since both give $g(x)$. Thus $h * k$ is continuous, and it is a homotopy from f to j .

Composition. Let h be a homotopy from f to f' and k one from g to g' . Then $g \circ h$ is a homotopy from $g \circ f$ to $g \circ f'$, and $(x, t) \mapsto k(f'(x), t)$ is a homotopy from $g \circ f'$ to $g' \circ f'$. Concatenating them by transitivity gives $g \circ f \simeq g' \circ f'$. $\square$

Definition 1.3. Write $[X, Y]$ for the set of homotopy classes of maps $X \rightarrow Y$, that is, the quotient of the set of continuous maps $\text{Hom}_{\mathbf{Top}}(X, Y)$ by the equivalence relation $\simeq$. The *homotopy category* $\mathbf{hTop}$ has

- (i) as objects all topological spaces, and
- (ii) $\text{Hom}_{\mathbf{hTop}}(X, Y) = [X, Y]$, with composition $[g] \circ [f] = [g \circ f]$.

Proposition 1.4. *$\mathbf{hTop}$ is a category, and $\gamma: \mathbf{Top} \rightarrow \mathbf{hTop}$ which is the identity on objects and sends f to $[f]$ is a functor.*

Proof. Composition is well defined by the last clause of Proposition 1.2: if $[f] = [f']$ and $[g] = [g']$ then $[g \circ f] = [g' \circ f']$. It is associative, and $[\text{id}_X]$ is a two-sided unit, because these identities already hold before passing to classes. Each $[X, Y]$ is a set, being a quotient of one. That γ is a functor is the definition of composition in $\mathbf{hTop}$ together with $\gamma(\text{id}_X) = [\text{id}_X]$. $\square$

The following says that $\mathbf{hTop}$ is exactly the category obtained from $\mathbf{Top}$ by decreeing that homotopic maps are equal, and nothing more.

Proposition 1.5 (universal property). *Let $F: \mathbf{Top} \rightarrow \mathcal{C}$ be a functor with the property that $f \simeq g$ implies $F(f) = F(g)$. Then there is a unique functor $\bar{F}: \mathbf{hTop} \rightarrow \mathcal{C}$ with $\bar{F} \circ \gamma = F$.*

Proof. Uniqueness is forced: γ is the identity on objects and surjective on morphisms, so $\bar{F}(X) = F(X)$ and $\bar{F}([f]) = F(f)$ are the only possible definitions. They are legitimate: $\bar{F}([f])$ is independent of the representative precisely by the hypothesis on F . And $\bar{F}$ is a functor, since $\bar{F}([g][f]) = \bar{F}([gf]) = F(gf) = F(g)F(f) = \bar{F}([g])\bar{F}([f])$ and $\bar{F}([\text{id}]) = F(\text{id}) = \text{id}$. $\square$

Definition 1.6. A map $f: X \rightarrow Y$ is a *homotopy equivalence* if $[f]$ is an isomorphism in $\mathbf{hTop}$. More directly f is a homotopy equivalence if there is $g: Y \rightarrow X$ with $g \circ f \simeq \text{id}_X$ and $f \circ g \simeq \text{id}_Y$. Spaces X and Y are then *homotopy equivalent*, $X \simeq Y$. A space is *contractible* if it is homotopy equivalent to a point, equivalently if id_X is homotopic to a constant map to a point.

Example 1.7. A convex subset $C \subseteq \mathbb{R}^n$ is contractible: $h(x, t) = (1 - t)x + tc_0$ is a homotopy from id_C to the constant map at any $c_0 \in C$. In particular $\mathbb{R}^n$ is contractible, so in $\mathbf{hTop}$ it is isomorphic to a point — although of course it is not homeomorphic to one. The inclusion $S^1 \hookrightarrow \mathbb{R}^2 \setminus \{0\}$ is a homotopy equivalence, with inverse $x \mapsto x/|x|$; so is the inclusion of the core circle into the Möbius band. Thus isomorphism in $\mathbf{hTop}$ is a much coarser relation than homeomorphism, and this coarsening is precisely what makes the invariants of algebraic topology computable.

A homotopy equivalence is most often produced by deforming a space onto a subspace.

Definition 1.8. Let $A \subseteq X$. A *retraction* of X onto A is a map $\rho: X \rightarrow A$ with $\rho|_A = \text{id}_A$. The subspace A is a *deformation retract* of X if there is a homotopy $F: X \times I \rightarrow X$ with

$$F_0 = \text{id}_X, \quad F_1(X) \subseteq A, \quad F_1|_A = \text{id}_A,$$

and a *strong deformation retract* if in addition $F_t|_A = \text{id}_A$ for every $t \in I$, so that the deformation is stationary on A throughout and not merely at its end. Such an F is a (*strong*) *deformation retraction*, and F_1 , regarded as a map $X \rightarrow A$, is a retraction.

Proposition 1.9. *If A is a deformation retract of X then the inclusion $\iota: A \rightarrow X$ is a homotopy equivalence, with homotopy inverse the retraction $\rho = F_1$. If the deformation retract is strong, the homotopy from $\iota \circ \rho$ to id_X may be taken rel A ; in particular ι is then a pointed homotopy equivalence for any base point of A .*

Proof. Since $F_1|_A = \text{id}_A$ we have $\rho \circ \iota = \text{id}_A$. In the other order $\iota \circ \rho$ is F_1 read as a self-map of X , and F is a homotopy from $F_0 = \text{id}_X$ to it; so $[\iota]$ is an isomorphism in $\mathbf{hTop}$ with inverse $[\rho]$, which is what it means for ι to be a homotopy equivalence. If $F_t|_A = \text{id}_A$ for every t then that homotopy is rel A , and both ι and ρ are pointed for a base point $a_0 \in A$, the two homotopies being pointed as well — the first is the constant homotopy of id_A . $\square$

More generally, one considers maps and homotopies of pairs of spaces.

Definition 1.10. Let (X, A) and (Y, B) be pairs of spaces. A *relative map*, or map of pairs, $f: (X, A) \rightarrow (Y, B)$ is a map $f: X \rightarrow Y$ with $f(A) \subseteq B$. Two relative maps f, g are *relatively homotopic* if there is a relative homotopy between them: that is, a map $h: (X \times I, A \times I) \rightarrow (Y, B)$ with $h(x, 0) = f(x)$ and $h(x, 1) = g(x)$ for all $x \in X$.

Proposition 1.11. *Relative homotopy is an equivalence relation on relative maps, and this relation is compatible with composition.*

Proof. The argument in the absolute case works as well in the relative case: the constant homotopy at a relative map is relative; the reverse $\bar{h}(x, t) = h(x, 1 - t)$ of a relative homotopy carries $A \times I$ into B because h does; the concatenation $h * k$ of two relative homotopies is given on each half by h or k , hence again carries $A \times I$ into B ; and in the composition clause, with $f, f': (X, A) \rightarrow (Y, B)$ and $g, g': (Y, B) \rightarrow (Z, C)$, the maps $g \circ h$ and $(x, t) \mapsto k(f'(x), t)$ carry $A \times I$ into $g(B)$ and into $k(B \times I)$ respectively, both of which lie in C . $\square$

2. THE POINTED HOMOTOPY CATEGORY

Definition 2.1. A *pointed space* is a pair (X, x_0) with $x_0 \in X$, the *basepoint*. A *pointed map* $f: (X, x_0) \rightarrow (Y, y_0)$ is a map $f: X \rightarrow Y$ with $f(x_0) = y_0$. Pointed spaces and pointed maps form the category $\mathbf{Top}_*$.

Definition 2.2. A *pointed homotopy* between pointed maps $f, g: (X, x_0) \rightarrow (Y, y_0)$ is a homotopy $h: X \times I \rightarrow Y$ from f to g with

$$h(x_0, t) = y_0 \quad \text{for all } t \in I,$$

that is, a homotopy rel $\{x_0\}$ through pointed maps. We write $f \simeq_* g$, and $[X, Y]_*$ for the set of pointed homotopy classes.

Proposition 2.3. *Pointed homotopy is an equivalence relation on pointed maps $(X, x_0) \rightarrow (Y, y_0)$, compatible with composition. Consequently there is a category $\mathbf{hTop}_*$, the pointed homotopy category, with the pointed spaces as objects and $\mathrm{Hom}_{\mathbf{hTop}_*}(X, Y) = [X, Y]_*$, and the quotient $\gamma_*: \mathbf{Top}_* \rightarrow \mathbf{hTop}_*$ is a functor with the universal property of Proposition 1.5.*

Proof. Every construction in the proofs of Propositions 1.2, 1.4 and 1.5 preserves the basepoint condition. The constant homotopy at a pointed map is pointed; the reverse $\bar{h}$ of a pointed homotopy has $\bar{h}(x_0, t) = h(x_0, 1 - t) = y_0$; the concatenation $h * k$ takes the value y_0 at x_0 for every t , being given by h or k there; and in the composition clause, $g \circ h$ and $(x, t) \mapsto k(f'(x), t)$ send (x_0, t) to $g(y_0) = z_0$ and $k(y_0, t) = z_0$ respectively. The remaining verifications are word for word those already given. $\square$

Remark 2.4 (pointed versus free). Forgetting basepoints gives a functor $\mathbf{hTop}_* \rightarrow \mathbf{hTop}$ and, for each pair of pointed spaces, a map of sets

$$[X, Y]_* \longrightarrow [X, Y].$$

It is neither injective nor surjective in general. If Y is path-connected it is surjective whenever the base point of X is non-degenerate — a free homotopy can then be dragged to the base point along a path, as in Lemma 6.6 below — and its fibres are the orbits of an action of $\pi_1(Y, y_0)$ on $[X, Y]_*$. The case $X = S^n$ is taken up in Proposition 7.1, and its relative analogue in Proposition 7.3; the point now is that $[X, Y]_*$ and $[X, Y]$ are different

sets, and that the group structures on homotopy classes, and hence the homotopy groups which we shall define, live on the pointed category.

3. THE FUNDAMENTAL GROUP

Fix a pointed space (X, x_0) .

3.1. The set underlying the fundamental group.

Definition 3.1. A *loop* in X at x_0 is a map $\gamma: I \rightarrow X$ with $\gamma(0) = \gamma(1) = x_0$. Two loops γ, γ' are *homotopic rel endpoints* if $\gamma \simeq \gamma' \text{ rel } \partial I = \{0, 1\}$ in the sense of Definition 1.1; that is, if there is a map $h: I \times I \rightarrow X$ with

$$h(s, 0) = \gamma(s), \quad h(s, 1) = \gamma'(s), \quad h(0, t) = h(1, t) = x_0$$

for all s, t . By Proposition 1.11 this is an equivalence relation. We define the set

$$\pi_1(X, x_0) = \{\text{loops in } X \text{ at } x_0\} / \simeq \text{ rel } \partial I,$$

writing $[\gamma]$ for the class of γ .

So far this is only a pointed set, with distinguished element the class of the constant loop $c(s) = x_0$. It is the same set as one of the pointed homotopy sets of §2.

Proposition 3.2. Let $q: I \rightarrow I/\partial I \cong S^1$ be the quotient map, the identification carrying the class of ∂I to $1 \in S^1$, so that S^1 is pointed by $1 = q(0) = q(1)$. Each loop γ at x_0 factors uniquely as $\gamma = \bar{\gamma} \circ q$, and $\gamma \mapsto \bar{\gamma}$ is a bijection

$$\pi_1(X, x_0) \xrightarrow{\cong} [(S^1, 1), (X, x_0)]_*.$$

Proof. A loop is exactly a map $I \rightarrow X$ that is constant on ∂I , so by the universal property of the quotient topology it factors uniquely through q as a pointed map $S^1 \rightarrow X$, and every pointed map $S^1 \rightarrow X$ arises this way. The same argument applied to $I \times I \rightarrow X$, using that $q \times \text{id}_I$ is a quotient map, matches homotopies rel ∂I with pointed homotopies $S^1 \times I \rightarrow X$. Hence the correspondence is bijective on classes. $\square$

3.2. The group structure.

Definition 3.3. The *concatenation* of loops γ, δ at x_0 is the loop

$$(\gamma \cdot \delta)(s) = \begin{cases} \gamma(2s), & 0 \leq s \leq \frac{1}{2}, \\ \delta(2s - 1), & \frac{1}{2} \leq s \leq 1, \end{cases}$$

which is continuous by the gluing lemma, the two formulas agreeing at $s = \frac{1}{2}$ where both give x_0 . The *reverse* of γ is $\bar{\gamma}(s) = \gamma(1 - s)$.

Two lemmas do all the work.

Lemma 3.4 (concatenation descends to classes). *If $\gamma \simeq \gamma'$ and $\delta \simeq \delta'$ rel ∂I , then $\gamma \cdot \delta \simeq \gamma' \cdot \delta'$ rel ∂I .*

Proof. Let h be a homotopy from γ to γ' and k one from δ to δ' , both rel ∂I . Put

$$H(s, t) = \begin{cases} h(2s, t), & 0 \leq s \leq \frac{1}{2}, \\ k(2s - 1, t), & \frac{1}{2} \leq s \leq 1. \end{cases}$$

At $s = \frac{1}{2}$ both give $h(1, t) = x_0 = k(0, t)$, so H is continuous; $H(\cdot, 0) = \gamma \cdot \delta$, $H(\cdot, 1) = \gamma' \cdot \delta'$, and $H(0, t) = h(0, t) = x_0$, $H(1, t) = k(1, t) = x_0$. $\square$

Lemma 3.5 (reparametrization). *Let $\gamma: I \rightarrow X$ be a path and $\varphi: I \rightarrow I$ a map with $\varphi(0) = 0$ and $\varphi(1) = 1$. Then $\gamma \circ \varphi \simeq \gamma$ rel ∂I .*

Proof. Set $H(s, t) = \gamma((1 - t)\varphi(s) + ts)$, which makes sense because I is convex and is continuous. Then $H(s, 0) = \gamma(\varphi(s))$ and $H(s, 1) = \gamma(s)$, while $H(0, t) = \gamma(0)$ and $H(1, t) = \gamma(1)$ because φ fixes 0 and 1. $\square$

Theorem 3.6. *The operation $[\gamma][\delta] = [\gamma \cdot \delta]$ makes $\pi_1(X, x_0)$ a group, with identity the class of the constant loop c and with $[\gamma]^{-1} = [\bar{\gamma}]$.*

Proof. The operation is well defined by Lemma 3.4.

Associativity. Let $\varphi: I \rightarrow I$ be the piecewise linear map with $\varphi(0) = 0$, $\varphi(\frac{1}{4}) = \frac{1}{2}$, $\varphi(\frac{1}{2}) = \frac{3}{4}$, $\varphi(1) = 1$. Comparing the two concatenations on the three intervals $[0, \frac{1}{4}]$, $[\frac{1}{4}, \frac{1}{2}]$, $[\frac{1}{2}, 1]$ gives

$$(\gamma \cdot \delta) \cdot \varepsilon = (\gamma \cdot (\delta \cdot \varepsilon)) \circ \varphi,$$

so the two are homotopic rel ∂I by Lemma 3.5.

Identity. With φ the map that is 0 on $[0, \frac{1}{2}]$ and $2s - 1$ on $[\frac{1}{2}, 1]$ one has $c \cdot \gamma = \gamma \circ \varphi$, whence $c \cdot \gamma \simeq \gamma$; symmetrically $\gamma \cdot c \simeq \gamma$.

Inverses. Define

$$H(s, t) = \begin{cases} \gamma(2s(1 - t)), & 0 \leq s \leq \frac{1}{2}, \\ \gamma((2 - 2s)(1 - t)), & \frac{1}{2} \leq s \leq 1. \end{cases}$$

The two formulas agree at $s = \frac{1}{2}$, both giving $\gamma(1 - t)$, so H is continuous. At $t = 0$ it is $\gamma \cdot \bar{\gamma}$, since $\gamma(2 - 2s) = \bar{\gamma}(2s - 1)$; at $t = 1$ it is the constant loop; and $H(0, t) = H(1, t) = \gamma(0) = x_0$. Hence $[\gamma][\bar{\gamma}] = [c]$, and applying this to $\bar{\gamma}$ in place of γ gives $[\bar{\gamma}][\gamma] = [c]$. $\square$

Definition 3.7. The group of Theorem 3.6 is the *fundamental group* $\pi_1(X, x_0)$ of the pointed space (X, x_0) .

Remark 3.8. The group structure is visible on the loops but not on the sphere: under the bijection of Proposition 3.2 it comes from the pinch map $S^1 \rightarrow S^1 \vee S^1$, which identifies the point $q(\frac{1}{2})$ with the base point $q(0)$.

3.3. Functoriality of π_1 .

Definition 3.9. Let $f: (X, x_0) \rightarrow (Y, y_0)$ be a pointed map. For a loop γ at x_0 the composite $f \circ \gamma$ is a loop at y_0 , and we set

$$f_*([\gamma]) = [f \circ \gamma].$$

Proposition 3.10. *f_* is a well-defined group homomorphism $\pi_1(X, x_0) \rightarrow \pi_1(Y, y_0)$, and*

$$(g \circ f)_* = g_* \circ f_*, \quad (\text{id}_X)_* = \text{id}_{\pi_1(X, x_0)}.$$

Hence $\pi_1: \mathbf{Top}_* \rightarrow \mathbf{Grp}$ is a functor.

Proof. Well defined. If h is a homotopy rel ∂I from γ to γ' , then $f \circ h$ is a map $I \times I \rightarrow Y$ with $(f \circ h)(s, 0) = f\gamma(s)$, $(f \circ h)(s, 1) = f\gamma'(s)$ and $(f \circ h)(0, t) = (f \circ h)(1, t) = f(x_0) = y_0$; so it is a homotopy rel ∂I from $f \circ \gamma$ to $f \circ \gamma'$.

Homomorphism. Composing with f commutes with the formula of Definition 3.3: for $s \leq \frac{1}{2}$ both $f \circ (\gamma \cdot \delta)$ and $(f \circ \gamma) \cdot (f \circ \delta)$ are $f\gamma(2s)$, and for $s \geq \frac{1}{2}$ both are $f\delta(2s - 1)$. So $f \circ (\gamma \cdot \delta) = (f \circ \gamma) \cdot (f \circ \delta)$, and passing to classes gives $f_*([\gamma][\delta]) = f_*[\gamma] f_*[\delta]$.

Functoriality. $(g \circ f) \circ \gamma = g \circ (f \circ \gamma)$ and $\text{id}_X \circ \gamma = \gamma$. $\square$

Proposition 3.11. *If $f \simeq_* g: (X, x_0) \rightarrow (Y, y_0)$ then $f_* = g_*$. Consequently π_1 factors through the pointed homotopy category: there is a functor*

$$\pi_1: \mathbf{hTop}_* \longrightarrow \mathbf{Grp}$$

with $\pi_1 \circ \gamma_$ the functor of Proposition 3.10.*

Proof. Let $h: X \times I \rightarrow Y$ be a pointed homotopy from f to g , so $h(x_0, t) = y_0$ for all t . Given a loop γ at x_0 , the map

$$H(s, t) = h(\gamma(s), t)$$

is continuous, satisfies $H(s, 0) = f\gamma(s)$ and $H(s, 1) = g\gamma(s)$, and $H(0, t) = H(1, t) = h(x_0, t) = y_0$. So H is a homotopy rel ∂I and $f_*[\gamma] = [f\gamma] = [g\gamma] = g_*[\gamma]$. The factorization is now Proposition 1.5 in its pointed form, Proposition 2.3: a functor on $\mathbf{Top}_*$ identifying pointed homotopic maps factors uniquely through γ_* . $\square$

Corollary 3.12. *A pointed homotopy equivalence $f: (X, x_0) \rightarrow (Y, y_0)$ induces an isomorphism $f_*: \pi_1(X, x_0) \rightarrow \pi_1(Y, y_0)$. In particular homeomorphic pointed spaces have isomorphic fundamental groups.*

Proof. By definition $[f]$ is an isomorphism in $\mathbf{hTop}_*$, and functors carry isomorphisms to isomorphisms; apply π_1 of Proposition 3.11. $\square$

This is our first invariant of Algebraic Topology. It is a functor from the pointed homotopy category $\mathbf{hTop}_*$ to the category of groups. It follows that two pointed topological spaces are homotopy equivalent relative to their base points only if their fundamental groups are isomorphic.

Example 3.13. If $C \subseteq \mathbb{R}^n$ is convex and $c_0 \in C$, then $\pi_1(C, c_0)$ is trivial: for any loop γ at c_0 the straight-line homotopy $H(s, t) = (1 - t)\gamma(s) + tc_0$ takes values in C , runs from γ to the constant loop, and fixes ∂I since $\gamma(0) = \gamma(1) = c_0$. By Corollary 3.12 the same holds for any pointed space pointed homotopy equivalent to a point.

3.4. Change of basepoint. The group depends on the basepoint, but, assuming that the space is path connected, only up to an isomorphism.

Proposition 3.14. *Let α be a path in X from x_0 to x_1 . Then*

$$\beta_\alpha: \pi_1(X, x_1) \longrightarrow \pi_1(X, x_0), \quad \beta_\alpha[\gamma] = [\alpha \cdot (\gamma \cdot \bar{\alpha})],$$

is a group isomorphism, depending only on the class of α rel ∂I , with inverse $\beta_{\bar{\alpha}}$. If α' is another such path then

$$\beta_{\alpha'} = c_u \circ \beta_\alpha, \quad c_u(g) = ugu^{-1}, \quad u = [\alpha' \cdot \bar{\alpha}] \in \pi_1(X, x_0);$$

so the two differ by an inner automorphism, and the isomorphism β_α is canonical only up to conjugation.

Proof. It is well defined by Lemma 3.4, whose proof applies verbatim to paths with matching endpoints, and depends only on the class of α for the same reason. It is a homomorphism because $\alpha \cdot (\gamma \cdot \bar{\alpha})$ concatenated with $\alpha \cdot (\delta \cdot \bar{\alpha})$ is, after reparametrization and cancelling $\bar{\alpha} \cdot \alpha \simeq c_{x_1}$ as in the proof of Theorem 3.6, homotopic rel ∂I to $\alpha \cdot ((\gamma \cdot \delta) \cdot \bar{\alpha})$. The same cancellations give $\beta_{\bar{\alpha}}\beta_{\alpha} = \text{id}$ and $\beta_{\alpha}\beta_{\bar{\alpha}} = \text{id}$. For the last statement, $\beta_{\alpha'}[\gamma] = [\alpha' \cdot \bar{\alpha}] \beta_{\alpha}[\gamma] [\alpha \cdot \bar{\alpha}']$, again by inserting $\bar{\alpha} \cdot \alpha \simeq c_{x_1}$; and $[\alpha \cdot \bar{\alpha}'] = u^{-1}$ for $u = [\alpha' \cdot \bar{\alpha}]$, so this reads $\beta_{\alpha'} = c_u \circ \beta_{\alpha}$. $\square$

4. HIGHER HOMOTOPY GROUPS

4.1. The groups $\pi_n(X, x_0)$. We continue with a fixed pointed space (X, x_0) . Write I^n for the n -cube and ∂I^n for its boundary, the set of points at least one of whose coordinates is 0 or 1.

Definition 4.1. A map of pairs $f: (I^n, \partial I^n) \rightarrow (X, x_0)$ is a map $f: I^n \rightarrow X$ with $f(s) = x_0$ for every $s \in \partial I^n$. Two such, f and g , are *homotopic* if they are homotopic through maps of pairs, that is, rel ∂I^n : there is $h: I^n \times I \rightarrow X$ with

$$h(s, 0) = f(s), \quad h(s, 1) = g(s), \quad h(s, t) = x_0 \quad \text{for } s \in \partial I^n.$$

This is an equivalence relation by Proposition 1.11. We set

$$\pi_n(X, x_0) = \{ f: (I^n, \partial I^n) \rightarrow (X, x_0) \} / \simeq \text{rel } \partial I^n,$$

a set pointed by the class of the constant map. For $n = 1$ this is Definition 3.1.

We fix, once and for all, an identification of $I^n/\partial I^n$ with S^n , as follows. Let $B^n \subseteq \mathbb{R}^n$ be the closed ball of radius π about the origin. Translating it so as to be centred at the centre $(\frac{1}{2}, \dots, \frac{1}{2})$ of the unit cube, and then scaling each radial line linearly, gives a homeomorphism of pairs $(B^n, \partial B^n) \rightarrow (I^n, \partial I^n)$. Next place B^n in $\mathbb{R}^{n+1}$, translated so as to be centred at the north pole $(0, \dots, 0, 1)$ of the unit sphere S^n and to lie in the tangent hyperplane there, and send the point of B^n at distance t from the centre in the direction u to the point at arclength t along the geodesic semicircle from the north pole tangent to u . This carries the centre to the north pole and, those semicircles having length π , carries all of ∂B^n to the south pole; so it induces a homeomorphism $B^n/\partial B^n \rightarrow S^n$. Composing the two gives the identification of $I^n/\partial I^n$ with S^n , under which the base point $*$ of S^n , the image of ∂I^n , is the south pole.

Proposition 4.2. *The given identification of $I^n/\partial I^n$ with S^n induces a bijection*

$$\pi_n(X, x_0) \cong [(S^n, *), (X, x_0)]_*.$$

Proof. As in Proposition 3.2: a map of pairs is exactly a map $I^n \rightarrow X$ constant on ∂I^n , so it factors uniquely through the quotient, and $q \times \text{id}_I$ being a quotient map identifies homotopies rel ∂I^n with pointed homotopies. $\square$

Definition 4.3. For $1 \leq i \leq n$ and maps of pairs f, g define $f \cdot_i g$ by concatenating in the i -th coordinate:

$$(f \cdot_i g)(s_1, \dots, s_n) = \begin{cases} f(s_1, \dots, 2s_i, \dots, s_n), & 0 \leq s_i \leq \frac{1}{2}, \\ g(s_1, \dots, 2s_i - 1, \dots, s_n), & \frac{1}{2} \leq s_i \leq 1. \end{cases}$$

Both formulas give x_0 when $s_i = \frac{1}{2}$, so $f \cdot_i g$ is continuous, and it is again a map of pairs. Write $\bar{f}$ for f reversed in the first coordinate, $\bar{f}(s_1, s_2, \dots, s_n) = f(1 - s_1, s_2, \dots, s_n)$, and c for the constant map.

Theorem 4.4. $\pi_n(X, x_0)$ is a group under $[f][g] = [f \cdot_1 g]$, with identity $[c]$ and $[f]^{-1} = [\bar{f}]$.

Proof. The proofs of Lemma 3.4, Lemma 3.5 and Theorem 3.6 apply verbatim, with s_1 in the role of s and $s_2, \dots, s_n$ carried along as parameters. Each homotopy used there changes only the first coordinate, and carries $s_1 \in \{0, 1\}$ to a point of $\{0, 1\}$; so it carries ∂I^n into itself and is rel ∂I^n , not merely rel the two faces $s_1 \in \{0, 1\}$. $\square$

One of the main differences between the fundamental group and the higher homotopy groups is that the latter are abelian. This comes from the fact that for $n \geq 2$ there is a second product, $\cdot_2$, and the two interact in an interesting way.

Lemma 4.5 (interchange). *Let $n \geq 2$. For maps of pairs f, g, h, k ,*

$$(f \cdot_1 g) \cdot_2 (h \cdot_1 k) = (f \cdot_2 h) \cdot_1 (g \cdot_2 k),$$

an equality of maps, not merely of homotopy classes.

Proof. Both sides subdivide the unit square in the coordinates (s_1, s_2) into four quadrants and are given, on each, by the corresponding map reparametrized affinely. On the left, $\cdot_2$ splits the square into $s_2 \leq \frac{1}{2}$ and $s_2 \geq \frac{1}{2}$, and each half is split by $\cdot_1$ into $s_1 \leq \frac{1}{2}$ and $s_1 \geq \frac{1}{2}$; this puts f on the quadrant $s_1 \leq \frac{1}{2}, s_2 \leq \frac{1}{2}$, then g, h, k on the quadrants $(s_1 \geq \frac{1}{2}, s_2 \leq \frac{1}{2})$, $(s_1 \leq \frac{1}{2}, s_2 \geq \frac{1}{2})$, $(s_1 \geq \frac{1}{2}, s_2 \geq \frac{1}{2})$. On the right the splittings occur in the opposite order, and the same four maps land on the same four quadrants with the same affine reparametrizations. $\square$

Theorem 4.6. *For $n \geq 2$ the group $\pi_n(X, x_0)$ is abelian, and the two products $\cdot_1$ and $\cdot_2$ induce the same operation on it.*

Proof. By Theorem 4.4 and its analogue for the i -th coordinate, both $\cdot_1$ and $\cdot_2$ descend to products on $\pi_n(X, x_0)$, and both have the class $e = [c]$ of the constant map as a two-sided identity. Write $a * b$ and $a \bullet b$ for the two. By Lemma 4.5 they satisfy the interchange law

$$(a * b) \bullet (u * v) = (a \bullet u) * (b \bullet v).$$

Now the Eckmann–Hilton argument applies. First, taking $b = u = e$,

$$a \bullet v = (a * e) \bullet (e * v) = (a \bullet e) * (e \bullet v) = a * v,$$

so the two products agree. Second, taking $a = v = e$,

$$b \bullet u = (e * b) \bullet (u * e) = (e \bullet u) * (b \bullet e) = u * b,$$

so $b * u = b \bullet u = u * b$, and the product is commutative. $\square$

Remark 4.7. Geometrically the argument says that for $n \geq 2$ there is room to slide the two cubes past each other: shrink f and g into disjoint sub-cubes, whose complement maps to x_0 , rotate them around one another within I^n , and expand again. For $n = 1$ there is no room, and indeed π_1 need not be abelian — as we shall see (Example 6.21), as a consequence of Van Kampen’s theorem, the figure eight has free fundamental group on two generators.

Remark 4.8. Everything in §3.3 carries over verbatim: a pointed map f induces $f_*[\alpha] = [f \circ \alpha]$, a homomorphism; pointed homotopic maps induce the same homomorphism, by the homotopy $H(s, t) = h(\alpha(s), t)$; and hence π_n is a functor

$$\pi_n: \mathbf{hTop}_* \longrightarrow \mathbf{Grp} \quad (n \geq 1),$$

landing in abelian groups for $n \geq 2$. In particular a pointed homotopy equivalence induces isomorphisms on all π_n . The change-of-basepoint isomorphisms of Proposition 3.14 also have analogues, giving an action of $\pi_1(X, x_0)$ on $\pi_n(X, x_0)$ for every n .

4.2. Relative homotopy groups. Let $A \subseteq X$ be a subspace containing the base point x_0 . Write

$$J^{n-1} = \overline{\partial I^n \setminus (I^{n-1} \times \{0\})} \subseteq \partial I^n,$$

the union of all the faces of I^n other than the bottom face $I^{n-1} \times \{0\}$; for $n = 1$ this is the single point $\{1\}$.

Definition 4.9. Let $n \geq 1$. A *map of triples* $f: (I^n, \partial I^n, J^{n-1}) \rightarrow (X, A, x_0)$ is a map $f: I^n \rightarrow X$ with $f(\partial I^n) \subseteq A$ and $f(J^{n-1}) = \{x_0\}$: it carries the face $s_n = 0$ into A and every other face to the base point. Two such maps are *homotopic* if they are homotopic through maps of triples. This is an equivalence relation, by the argument of Proposition 1.11, and we set

$$\pi_n(X, A, x_0) = \{ f: (I^n, \partial I^n, J^{n-1}) \rightarrow (X, A, x_0) \} / \simeq,$$

a set pointed by the class of the constant map c at x_0 .

Remark 4.10. For $A = \{x_0\}$ the condition $f(\partial I^n) \subseteq A$ says $f(\partial I^n) = \{x_0\}$, the condition on J^{n-1} is then automatic, and a homotopy through maps of triples is exactly a homotopy rel ∂I^n ; so $\pi_n(X, \{x_0\}, x_0) = \pi_n(X, x_0)$. For $n = 1$ a map of triples is a path in X from a point of A to x_0 , and $\pi_1(X, A, x_0)$ is the set of such paths up to homotopy with the initial point moving in A and the terminal point fixed.

The relative groups are often defined instead by maps of a ball, and the two descriptions agree. We record this as an exercise; it is nowhere used below, every argument that wants a ball going instead through the homeomorphism of pairs $(I^n, \partial I^n) \cong (D^n, S^{n-1})$, each being a compact convex body with its boundary.

Exercise 4.11. Write $D^n \subseteq \mathbb{R}^n$ for the closed unit ball, with boundary the unit sphere S^{n-1} . Show that collapsing J^{n-1} to a point carries the triple $(I^n, \partial I^n, J^{n-1})$ to a triple homeomorphic to $(D^n, S^{n-1}, *)$; that is, that the quotient I^n/J^{n-1} is an n -ball whose boundary sphere is the image of ∂I^n , pointed by the image of J^{n-1} . Deduce that $\pi_n(X, A, x_0)$ is also the set of homotopy classes of maps of triples from $(D^n, S^{n-1}, *)$ to (X, A, x_0) .

Suggestion. Writing $I^n = I^{n-1} \times I$ one has $J^{n-1} = (I^{n-1} \times \{1\}) \cup (\partial I^{n-1} \times I)$, a closed $(n-1)$ -ball in the boundary sphere of the n -ball I^n , with complementary ball the bottom face. So it is enough to prove the general statement: collapsing a closed $(n-1)$ -ball lying in the boundary sphere of an n -ball gives again an n -ball, in which the complementary ball of that sphere becomes the whole boundary sphere.

Theorem 4.12. *For $n \geq 2$ the concatenation $\cdot_1$ of Definition 4.3 makes $\pi_n(X, A, x_0)$ a group, with identity $[c]$ and $[f]^{-1} = [\bar{f}]$, and for $n \geq 3$ this group is abelian. For $n = 1$ there is in general no such structure, and $\pi_1(X, A, x_0)$ is only a pointed set.*

Proof. Let $n \geq 2$, so that the first coordinate is not the last. If f and g are maps of triples so is $f \cdot_1 g$: on the face $s_n = 0$ each half is given by f or by g on the face $s_n = 0$, hence lands in A ; on a face $s_j \in \{0, 1\}$ with $j \neq n$, and on the face $s_n = 1$, each half is given by f or g on such a face, hence is x_0 ; and the two formulas agree, both being x_0 , where $s_1 = \frac{1}{2}$. The proofs of Theorem 4.4 now apply verbatim: the homotopies used there are reparametrizations and concatenations in the first coordinate, and each of them carries the face $s_n = 0$ to points of A already in the image of f or g , and the remaining faces to x_0 , exactly as in that proof.

For $n \geq 3$ the second coordinate is also available, so $\cdot_2$ is defined on maps of triples as well, Lemma 4.5 holds unchanged, and the Eckmann–Hilton argument of Theorem 4.6 shows that the two products agree and are commutative. $\square$

The relative groups would be of little use without the following, which is the main reason for introducing them: they fit with the absolute groups into a long exact sequence, and that sequence is the standard instrument for transferring information between A , X and the pair. It is used constantly in what follows, and in *The Hurewicz theorems* and *Homotopy theory of CW complexes*; the reader should have it by heart.

Remark 4.13. A map of triples $f: (X, A, x_0) \rightarrow (Y, B, y_0)$ induces $f_*[\alpha] = [f \circ \alpha]$, a map of pointed sets and a homomorphism when $n \geq 2$, and homotopic maps of triples induce the same map; so $\pi_n(-, -, -)$ is a functor on pairs with base point. Restricting a map of triples to the face $s_n = 0$ gives a map of pairs $(I^{n-1}, \partial I^{n-1}) \rightarrow (A, x_0)$ and hence a *boundary map*

$$\partial: \pi_n(X, A, x_0) \longrightarrow \pi_{n-1}(A, x_0), \quad \partial[f] = [f|_{I^{n-1} \times \{0\}}],$$

a homomorphism for $n \geq 2$.

Definition 4.14. The boundary maps, together with the maps induced by the inclusions $A \subseteq X$ and $(X, \{x_0\}) \subseteq (X, A)$, fit into the sequence

$$\cdots \longrightarrow \pi_n(A, x_0) \longrightarrow \pi_n(X, x_0) \longrightarrow \pi_n(X, A, x_0) \xrightarrow{\partial} \pi_{n-1}(A, x_0) \longrightarrow \cdots \longrightarrow \pi_0(X, x_0),$$

the *homotopy long exact sequence* of the pair (X, A) at x_0 . Here $\pi_0(Y, y_0)$ is the case $n = 0$ of Definition 4.1 — since I^0 is a point and $\partial I^0 = \emptyset$, it is the set of path components of Y , pointed by the component of y_0 . The sequence is exact in the following sense, appropriate to terms that are only pointed sets: at each stage the image of the incoming map is the set of elements carried by the outgoing map to the distinguished element. The verification is Exercise 4.16.

Remark 4.15. The homotopy long exact sequence is natural: a map of triples $(X, A, x_0) \rightarrow (Y, B, y_0)$ gives a ladder of such sequences with commuting squares, since f_* commutes with the induced maps by definition and with ∂ because restriction to the face $s_n = 0$ commutes with composing by f . This naturality is used as often as the sequence itself.

Exercise 4.16. Prove the exactness asserted in Definition 4.14. Deduce, for $n \geq 2$, that the boundary map gives an isomorphism $\pi_n(D^k, S^{k-1}, x_0) \cong \pi_{n-1}(S^{k-1}, x_0)$, the disk being contractible and hence having vanishing homotopy groups.

5. A COMPUTATION: $\pi_1(S^1, 1) \cong \mathbb{Z}$

Regard $S^1 = \{z \in \mathbb{C} : |z| = 1\}$, pointed by 1, and let

$$\exp: \mathbb{R} \longrightarrow S^1, \quad \exp(t) = e^{2\pi it}.$$

It is a surjective homomorphism of groups with kernel $\mathbb{Z}$, and it is a *covering*: every $z \in S^1$ has an open neighborhood W , for instance the complement of $-z$, such that $\exp^{-1}(W)$ is a disjoint union of open intervals, each mapped homeomorphically onto W by $\exp$. Such a W is called *evenly covered*.

Lemma 5.1 (lifting). (i) *For every path $\gamma: I \rightarrow S^1$ and every t_0 with $\exp(t_0) = \gamma(0)$ there is a unique path $\tilde{\gamma}: I \rightarrow \mathbb{R}$ with $\exp \circ \tilde{\gamma} = \gamma$ and $\tilde{\gamma}(0) = t_0$.*
(ii) *For every map $h: I \times I \rightarrow S^1$ and every t_0 with $\exp(t_0) = h(0, 0)$ there is a unique $\tilde{h}: I \times I \rightarrow \mathbb{R}$ with $\exp \circ \tilde{h} = h$ and $\tilde{h}(0, 0) = t_0$.*

Sketch. Clearly, (i) follows from (ii), so we consider (ii). Cover S^1 by evenly covered open sets and pull the cover back to the compact metric space $I \times I$. By the Lebesgue covering lemma, Lemma 6.1 below, there is a subdivision of the square into finitely many small squares each of which is carried into a single evenly covered W . Order the small squares so that each meets the union of its predecessors in a connected set, and lift them one at a time: on each, $\exp$ restricted to the sheet containing the already-constructed value is a homeomorphism onto W , so there is exactly one way to extend the lift. Uniqueness follows from the same local statement: two lifts agreeing at a point agree on a set that is open and closed, hence on all of the connected square. $\square$

Definition 5.2. Let γ be a loop in S^1 at 1 and $\tilde{\gamma}$ its lift with $\tilde{\gamma}(0) = 0$. Since $\exp(\tilde{\gamma}(1)) = \gamma(1) = 1$, the number $\tilde{\gamma}(1)$ is an integer, called the *winding number* $\deg(\gamma)$.

Theorem 5.3. $\deg$ induces a group isomorphism

$$\pi_1(S^1, 1) \xrightarrow{\cong} \mathbb{Z}.$$

Proof. Well defined. Let h be a homotopy rel ∂I from γ to γ' and $\tilde{h}$ its lift with $\tilde{h}(0, 0) = 0$, which exists by Lemma 5.1(ii). For fixed t the path $s \mapsto \tilde{h}(s, t)$ lifts $s \mapsto h(s, t)$ and starts at $\tilde{h}(0, t)$; the path $t \mapsto \tilde{h}(0, t)$ lifts the constant path at 1, so by uniqueness it is constant equal to 0. Hence $s \mapsto \tilde{h}(s, 0)$ and $s \mapsto \tilde{h}(s, 1)$ are the lifts of γ and γ' starting at 0. Now $t \mapsto \tilde{h}(1, t)$ is a continuous map into $\exp^{-1}(1) = \mathbb{Z}$, hence constant, so $\deg(\gamma) = \tilde{h}(1, 0) = \tilde{h}(1, 1) = \deg(\gamma')$.

Homomorphism. If $\tilde{\gamma}, \tilde{\delta}$ are the lifts starting at 0, then the path which is $\tilde{\gamma}(2s)$ for $s \leq \frac{1}{2}$ and $\tilde{\gamma}(1) + \tilde{\delta}(2s - 1)$ for $s \geq \frac{1}{2}$ is continuous, lifts $\gamma \cdot \delta$ and starts at 0; its value at 1 is $\tilde{\gamma}(1) + \tilde{\delta}(1)$. So $\deg(\gamma \cdot \delta) = \deg \gamma + \deg \delta$.

Surjective. The loop $\gamma_n(s) = \exp(ns)$ lifts to $\tilde{\gamma}_n(s) = ns$, so $\deg \gamma_n = n$.

Injective. If $\deg \gamma = 0$ then $\tilde{\gamma}$ is a loop at 0 in $\mathbb{R}$, and $H(s, t) = (1 - t)\tilde{\gamma}(s)$ is a homotopy rel ∂I from $\tilde{\gamma}$ to the constant path in $\mathbb{R}$. Then $\exp \circ H$ is a homotopy rel ∂I from γ to the constant loop, so $[\gamma]$ is trivial. $\square$

Corollary 5.4. *There is no retraction $r: D^2 \rightarrow S^1$, that is, no map with $r|_{S^1} = \text{id}$.*

Proof. Such an r would give $\text{id} = (r \circ \iota)_* = r_* \circ \iota_*$ on $\pi_1(S^1, 1) \cong \mathbb{Z}$, where $\iota: S^1 \hookrightarrow D^2$; but ι_* factors through $\pi_1(D^2, 1) = 1$, D^2 being convex. So the identity of $\mathbb{Z}$ would factor through the trivial group. $\square$

6. VAN KAMPEN'S THEOREM

Under favorable hypotheses, Van Kampen's Theorem computes the fundamental group of $X = U \cup V$ in terms of those of U , V , and $U \cap V$. Two preliminaries come first: four lemmas on covers, subdivisions and the extension of homotopies, which have nothing to do with π_1 and are used throughout these notes, and then the group theory that the answer is expressed in — a special case of colimits over a shape in the category of groups.

6.1. Subdivisions and the homotopy extension property. Five lemmas are collected first. None of them mentions π_1 ; together they are the entire technical apparatus of the proof below; and the subdivision lemma and the extension lemma are used again and again in these notes and in the later ones. The first two say that a cube or a complex can be chopped into pieces small enough that each piece is carried into a single member of a given open cover; the third retracts a cylinder onto its bottom together with its sides; the fourth turns that retraction into the extension of a homotopy prescribed on a subcomplex, with control on how far it moves things; and the fifth, a consequence of the others, is the form in which they will be applied.

Lemma 6.1 (Lebesgue covering lemma). *Let X be a compact metric space, and let $\{U_1, \dots, U_k\}$ be a finite open cover of X . Then there is an $\varepsilon > 0$ such that every ball of radius ε in X is contained in one of the U_i .*

Proof. For each $x \in X$ choose $r_x > 0$ with $B(x, 2r_x) \subseteq U_{i(x)}$ for some $i(x)$. The balls $B(x, r_x)$ cover X , so finitely many $B(x_1, r_{x_1}), \dots, B(x_m, r_{x_m})$ do; put $\varepsilon = \min_j r_{x_j}$. If $B(y, \varepsilon)$ is any ball, then $y \in B(x_j, r_{x_j})$ for some j , and for $z \in B(y, \varepsilon)$ we get $d(z, x_j) \leq d(z, y) + d(y, x_j) < \varepsilon + r_{x_j} \leq 2r_{x_j}$; so $B(y, \varepsilon) \subseteq B(x_j, 2r_{x_j}) \subseteq U_{i(x_j)}$. $\square$

Lemma 6.2 (fine subdivisions). *Let K be a finite simplicial complex, realized rectilinearly in a Euclidean space, and write $\text{sd}^r K$ for its r -th barycentric subdivision.*

- (i) *For every $\varepsilon > 0$ there is an r for which every simplex of $\text{sd}^r K$ has diameter less than ε .*
- (ii) *Let $h: |K| \rightarrow W$ be a map and let $\mathcal{U}$ be a family of open subsets of W whose union contains $h(|K|)$. Then there is an r for which the image under h of each closed simplex of $\text{sd}^r K$ lies in a single member of $\mathcal{U}$.*
- (iii) *Let $S \subseteq I^n$ be a finite set of points and let $\varepsilon > 0$. Then I^n has a cubification of mesh less than ε having every point of S as a vertex: a subdivision of I^n into rectangular boxes, any two of them meeting in a common face, each of diameter less than ε . Every face of I^n is a union of boxes; and triangulating the boxes makes the cubification a finite simplicial complex with realization I^n in which the points of S are vertices and every face of I^n is a subcomplex. Hence (i) and (ii) apply to cubes as well as to complexes, and to any union of faces of a cube — for instance ∂I^n , J^{n-1} , or, $I^n \times I$ being itself a cube, the subsets $\partial I^n \times I$ and $I^n \times \partial I$ of it.*

Proof. (i) The diameter of an affine simplex is the greatest distance between two of its vertices: for fixed u the function $v \mapsto |u - v|$ is convex, hence attains its maximum on a simplex at a vertex, and applying this twice gives the assertion.

Let s be a simplex of K , of dimension at most $q = \dim K$, and let $t \subseteq s$ be a face, of dimension l , with vertices $p_0, \dots, p_l$ and barycenter b_t . For each j ,

$$|b_t - p_j| = \left| \frac{1}{l+1} \sum_i (p_i - p_j) \right| \leq \frac{l}{l+1} \text{diam } t \leq \frac{q}{q+1} \text{diam } s,$$

and therefore $|b_t - x| \leq \frac{q}{q+1} \text{diam } s$ for every $x \in t$, by the previous paragraph applied to $u = b_t$. Now a simplex of $\text{sd } s$ has for its vertices the barycenters $b_{t_0}, \dots, b_{t_k}$ of a chain of faces $t_0 \subsetneq \dots \subsetneq t_k \subseteq s$; for $i < j$ the point b_{t_i} lies in t_j , so the estimate gives $|b_{t_j} - b_{t_i}| \leq \frac{q}{q+1} \text{diam } s$. Hence every simplex of $\text{sd } K$ has diameter at most $\frac{q}{q+1}$ times the largest diameter of a simplex of K , and iterating, every simplex of $\text{sd}^r K$ has diameter at most $\left(\frac{q}{q+1}\right)^r$ times that number. These tend to 0.

(ii) The sets $h^{-1}(U)$, for $U \in \mathcal{U}$, form an open cover of the compact metric space $|K|$, and finitely many of them cover it. By Lemma 6.1 there is an $\varepsilon > 0$ such that every ball of radius ε in $|K|$ lies in one of those finitely many sets; a subset of diameter less than ε is contained in the ball of radius ε about any of its points, so it too lies in one of them. Take r as in (i).

(iii) Exercise. In each coordinate take a partition of $[0, 1]$ whose points include the corresponding coordinates of the points of S and whose mesh is less than $\varepsilon/\sqrt{n}$, and let the boxes be the products of the resulting subintervals. $\square$

The third lemma is the geometric fact on which the extension property rests. For a compact convex body B write ∂B for its boundary in the affine space it spans, and call $B \times \{0\} \cup \partial B \times I$ the *bottom and sides* of the cylinder $B \times I$.

Lemma 6.3 (retracting a cylinder). *Let B be a compact convex subset with non-empty interior of a finite-dimensional affine space. Then the bottom and sides $B \times \{0\} \cup \partial B \times I$ form a retract of $B \times I$; a retraction is given by radial projection from a point above the cylinder.*

Proof. Identify the affine span of B with $\mathbb{R}^m$, fix an interior point b of B and put $p = (b, 2)$, a point lying above the cylinder. Let

$$\mu(y) = \inf \{ \lambda > 0 : b + \lambda^{-1}(y - b) \in B \}$$

be the gauge of B about b . Choosing $\rho > 0$ with the ball of radius ρ about b inside B gives $\mu(y) \leq \|y - b\|/\rho$, and convexity turns this into $|\mu(y) - \mu(y')| \leq \|y - y'\|/\rho$; so μ is continuous. It satisfies $\mu(b + s(y - b)) = s\mu(y)$ for $s \geq 0$, and $B = \{\mu \leq 1\}$, $\partial B = \{\mu = 1\}$.

For $x = (y, t) \in B \times I$ the ray from p through x is

$$\gamma_x(s) = (b + s(y - b), 2 - s(2 - t)) \quad (s \geq 0),$$

with $\gamma_x(1) = x$. Now $\gamma_x(s)$ lies in $B \times I$ exactly when $s\mu(y) \leq 1$ and $0 \leq 2 - s(2 - t) \leq 1$, that is, when $1/(2 - t) \leq s \leq \min\{1/\mu(y), 2/(2 - t)\}$. Define $s(x)$ to be that upper bound and put $r(x) = \gamma_x(s(x))$. Since $\mu(y) \leq 1$ and $2 - t \geq 1$ we have $s(x) \geq 1$, so r carries x to a point of $B \times I$ at or beyond x along the ray.

The point $r(x)$ lies in the bottom and sides: if the minimum is $1/\mu(y)$ then $\mu(b + s(x)(y - b)) = 1$ and $r(x) \in \partial B \times I$; otherwise $s(x) = 2/(2 - t)$, the height is 0, and $r(x) \in B \times \{0\}$. And r fixes that set pointwise: if $y \in \partial B$ then $1/\mu(y) = 1 \leq 2/(2 - t)$, while if $t = 0$ then $2/(2 - t) = 1 \leq 1/\mu(y)$; either way $s(x) = 1$ and $r(x) = \gamma_x(1) = x$. Finally $1/s(x) = \max\{\mu(y), (2 - t)/2\}$ is continuous and bounded below by $\frac{1}{2}$, so s and hence r are continuous. $\square$

The fourth lemma is the homotopy extension property for a simplicial pair. It is a technical tool that is used throughout the rest of these notes. It says that a homotopy prescribed on a subcomplex, and agreeing there with a given map, extends to a homotopy of the whole complex; and it says more, namely that the extension can be made without moving any simplex very far. The pattern of every argument that uses it is the same: prescribe the homotopy where one knows what to do, extend it by the lemma, and read off from the track bound that nothing has been disturbed elsewhere. Lemma 6.6 below is the first instance, and the whole of §7 is built on it; it is used again throughout *The Hurewicz theorems*. The same property holds, with the same proof in outline, for a CW pair; that version is proved in *CW complexes*, and it is what makes the homotopy theory of CW complexes work. Here K is a finite simplicial complex, $|K|$ its geometric realization, and $L \subseteq K$ a subcomplex.

Lemma 6.4 (homotopy extension). *Let K be a finite simplicial complex, $L \subseteq K$ a subcomplex, $f: |K| \rightarrow X$ a map, and $g: |L| \times I \rightarrow X$ a homotopy with $g(b, 0) = f(b)$ for all $b \in |L|$. Then there is a map $F: |K| \times I \rightarrow X$ with*

$$F(a, 0) = f(a) \text{ for all } a \in |K|, \quad F(b, t) = g(b, t) \text{ for all } (b, t) \in |L| \times I.$$

Moreover F may be chosen with its track over each simplex controlled. Writing $F(S \times I)$ for the track of F over a subset $S \subseteq |K|$, this means

$$F(|s| \times I) \subseteq f(|s|) \cup F(|\partial s| \times I) \quad \text{for every simplex } s \text{ of } K \text{ not in } L,$$

and consequently, $|s| \cap |L|$ being the union of the faces of s that lie in L ,

$$F(|s| \times I) \subseteq f(|s|) \cup g((|s| \cap |L|) \times I) \quad \text{for every simplex } s \text{ of } K.$$

Proof. Take first $K = \Delta^n$, the complex of all faces of the standard simplex, and $L = \partial\Delta^n$. In this case, it is enough to produce a retraction

$$r: \Delta^n \times I \longrightarrow \Delta^n \times \{0\} \cup \partial\Delta^n \times I,$$

for then $F = (f \cup g) \circ r$ is defined and continuous, the maps f on $\Delta^n \times \{0\}$ and g on $\partial\Delta^n \times I$ agreeing on the intersection $\partial\Delta^n \times \{0\}$ of those two closed sets, and F has the required restrictions because r fixes them pointwise. Such an r is provided by Lemma 6.3, applied to the compact convex body Δ^n in its affine span.

In general, let $s_1, \dots, s_m$ be the simplices of K not in L , numbered so that their dimensions do not decrease, and set $L_0 = L$ and $L_r = L_{r-1} \cup \{s_r\}$. Each L_r is a subcomplex, since every proper face of s_r has smaller dimension and so lies in L or among $s_1, \dots, s_{r-1}$; and $|s_r|$ meets $|L_{r-1}|$ exactly in $|\partial s_r|$. Suppose inductively that the homotopy has been extended to $G_{r-1}: |L_{r-1}| \times I \rightarrow X$ beginning at f . Applying the case already proved to f restricted to s_r and to G_{r-1} restricted to $|\partial s_r| \times I$ extends the homotopy over $|s_r| \times I$; the two definitions agree on $|\partial s_r| \times I$, and $|L_{r-1}| \times I$ and $|s_r| \times I$ are closed in $|L_r| \times I$,

so together they define a continuous $G_r: |L_r| \times I \rightarrow X$. Since K is finite, after m steps we reach the required $F = G_m$ on $|K| \times I$.

Finally, the construction gives the two track bounds. The extension over $|s_r| \times I$ was made by the case already proved, where $F = (f \cup g) \circ r$ takes values in $f(\Delta^n) \cup g(\partial\Delta^n \times I)$; so $F(|s_r| \times I) \subseteq f(|s_r|) \cup F(|\partial s_r| \times I)$, which is the first bound. The second follows by induction on $\dim s$. If $s \in L$ it is immediate, since then $F(|s| \times I) = g(|s| \times I)$ and $|s| \cap |L| = |s|$. If $s \notin L$, every proper face t of s satisfies $F(|t| \times I) \subseteq f(|t|) \cup g((|t| \cap |L|) \times I)$ by induction, while $f(|t|) \subseteq f(|s|)$ and $|t| \cap |L| \subseteq |s| \cap |L|$; substituting this into the first bound gives the second. $\square$

Remark 6.5. The track bound is what makes Lemma 6.4 usable when the homotopy is required to stay inside a subspace. Let $A \subseteq X$, and let $K' \subseteq K$ be a subcomplex with $f(|K'|) \subseteq A$ and $g((|K'| \cap |L|) \times I) \subseteq A$. Then $F(|K'| \times I) \subseteq A$, so that the extended homotopy moves $|K'|$ only inside A : indeed each simplex s of K' has $f(|s|) \subseteq A$ and $|s| \cap |L| \subseteq |K'| \cap |L|$, so the second bound puts $F(|s| \times I)$ in A . At the other extreme, if s is a vertex of K not in L then $|\partial s| = \emptyset$ and the first bound says that F is stationary at s . The second form is what the next lemma uses; both are used repeatedly in §7 and in *The Hurewicz theorems*.

The lemmas above combine at once into the form in which they will be used: a map may be dragged along a path near a chosen point, at the cost of an arbitrarily small disturbance confined to a neighborhood of that point and to a neighborhood of the path. This is the lemma that puts the vertices of a subdivision at the base point in the proof below.

Lemma 6.6 (non-degenerate base points). *Let $I^n = \prod_{i=1}^n I$ be the n -cube, let $p \in I^n$, let $f: I^n \rightarrow X$ be a continuous map, let $\omega: I \rightarrow X$ be a path with $\omega(0) = f(p)$, and let A_0 be a neighborhood of p in I^n . Then there are a closed neighborhood $N \subseteq A_0$ of p and a homotopy $H: I^n \times I \rightarrow X$ such that*

- (i) $H(q, 0) = f(q)$ for all $q \in I^n$;
- (ii) $H(p, s) = \omega(s)$ for all $s \in I$;
- (iii) H is the constant homotopy on $I^n \setminus N$, and $H(N \times I) \subseteq f(N) \cup \omega(I)$.

In particular, if W is a neighborhood of $\omega(I)$ and A_0 is taken inside $f^{-1}(W)$, then $H(N \times I) \subseteq W$: the homotopy is supported in the arbitrarily small set N , and there it moves points only inside W .

Proof. This is Lemma 6.4 together with its bound on tracks. Choose $\varepsilon > 0$ with the ball of radius ε about p contained in A_0 , and triangulate I^n with p as a vertex and every simplex of diameter less than ε , which is possible by Lemma 6.2(iii). The closed star N of p then lies in A_0 , being a union of simplices containing p . Let L be the subcomplex consisting of p together with all the simplices that do not have p as a vertex, so that $|L|$ is the disjoint union of the point p and the closed set $I^n \setminus \text{st}(p)$, where $\text{st}(p)$ is the open star. On $|L| \times I$ let g be ω at p and the constant homotopy at f elsewhere; it is continuous, the two formulas being given on disjoint closed sets, and it begins at $f|_{|L|}$ because $\omega(0) = f(p)$. Let H be an extension of g furnished by Lemma 6.4.

Then (i) and (ii) hold by construction, and so does the first assertion of (iii), since $H = g$ on $|L| \times I$ and every point outside N lies in $|L|$. For the second, let s be a simplex

of N . On $(|s| \cap |L|) \times I$ the homotopy g takes its values in $f(|s|) \cup \omega(I)$, so the track bound gives $H(|s| \times I) \subseteq f(|s|) \cup \omega(I)$; taking the union over the simplices of N yields $H(N \times I) \subseteq f(N) \cup \omega(I)$. $\square$

6.2. Free products with amalgamation. We turn to the group theory in which the answer is expressed. Fix groups G_1, G_2 and H and homomorphisms $\rho_i: H \rightarrow G_i, i = 1, 2$.

Definition 6.7. A free product of G_1 and G_2 amalgamated along H is a group P together with homomorphisms $j_i: G_i \rightarrow P$ satisfying

$$j_1 \circ \rho_1 = j_2 \circ \rho_2$$

and universal with that property, in the sense that for every group Γ and every pair of homomorphisms $f_i: G_i \rightarrow \Gamma$ with $f_1 \circ \rho_1 = f_2 \circ \rho_2$ there is a *unique* homomorphism $f: P \rightarrow \Gamma$ with $f \circ j_i = f_i$ for $i = 1, 2$. We write $P = G_1 \star_{\rho_1, H, \rho_2} G_2$, abbreviated to $G_1 *_H G_2$ when the ρ_i are understood.

In the language of the companion notes this says that

$$\begin{array}{ccc} H & \xrightarrow{\rho_1} & G_1 \\ \rho_2 \downarrow & & \downarrow j_1 \\ G_2 & \xrightarrow{j_2} & P \end{array}$$

is a pushout square in **Grp**, that is, a colimit of the diagram $G_1 \leftarrow H \rightarrow G_2$. Said another way, this is the colimit of a diagram of the shape $\bullet \leftarrow \bullet \rightarrow \bullet$ in the category of groups.

Proposition 6.8 (uniqueness). *If (P, j_1, j_2) and (P', j'_1, j'_2) both satisfy Definition 6.7, then there is a unique isomorphism $u: P \rightarrow P'$ with $u \circ j_i = j'_i$ for $i = 1, 2$.*

Proof. This is the usual argument for objects defined by a universal property, as in *Limits and colimits*. Applying the universal property of P to (P', j'_1, j'_2) gives a unique $u: P \rightarrow P'$ with $u \circ j_i = j'_i$, and applying that of P' to (P, j_1, j_2) gives a unique $u': P' \rightarrow P$ with $u' \circ j'_i = j_i$. Then $u' \circ u$ and id_P both satisfy $w \circ j_i = j_i$, so $u' \circ u = \text{id}_P$ by the uniqueness clause of the universal property of P ; likewise $u \circ u' = \text{id}_{P'}$. Hence u is an isomorphism, and it is unique with $u \circ j_i = j'_i$. $\square$

Assuming it exists, $G_1 \star_{\rho_1, H, \rho_2} G_2$ is well defined up to a unique isomorphism compatible with the structure maps, and one may speak of *the* amalgamated free product. We still have to show that it exists.

Proposition 6.9 (construction). *Let F be the free group on the set $G_1 \sqcup G_2$, writing $[g]$ for the generator corresponding to g , and let $N \trianglelefteq F$ be the normal subgroup generated by*

$$[g][g'] [gg']^{-1} \quad (g, g' \in G_i, i = 1, 2), \quad [\rho_1(h)] [\rho_2(h)]^{-1} \quad (h \in H).$$

Put $P = F/N$ and let $j_i: G_i \rightarrow P$ send g to the class of $[g]$. Then (P, j_1, j_2) satisfies Definition 6.7.

Proof. The relations of the first kind say exactly that each j_i is a homomorphism, and those of the second kind say that $j_1 \rho_1 = j_2 \rho_2$.

Let Γ and $f_i: G_i \rightarrow \Gamma$ be given with $f_1 \rho_1 = f_2 \rho_2$. By the universal property of the free group there is a unique homomorphism $\tilde{f}: F \rightarrow \Gamma$ with $\tilde{f}([g]) = f_i(g)$ for $g \in G_i$.

It kills the listed elements: $\tilde{f}([g][g'][gg']^{-1}) = f_i(g)f_i(g')f_i(gg')^{-1} = 1$ because f_i is a homomorphism, and $\tilde{f}([\rho_1(h)][\rho_2(h)]^{-1}) = f_1\rho_1(h)(f_2\rho_2(h))^{-1} = 1$ by hypothesis. A homomorphism killing a set kills the normal subgroup it generates, so $N \subseteq \ker \tilde{f}$ and $\tilde{f}$ factors through a homomorphism $f: P \rightarrow \Gamma$, which satisfies $f \circ j_i = f_i$. It is unique with that property because the classes of the $[g]$ generate P , and f is prescribed on them. $\square$

Remark 6.10. When H is the trivial group, we denote the amalgamated free product over H by $G_1 * G_2$ and call it the *free product* of G_1 and G_2 . Equivalently $G_1 \star_{\rho_1, H, \rho_2} G_2 = (G_1 * G_2) / \langle\langle j_1\rho_1(h)j_2\rho_2(h)^{-1} \rangle\rangle$, the free product divided by the normal closure of the indicated set, G_1 and G_2 being identified with their images in $G_1 * G_2$: taking $H = 1$ in Proposition 6.9 constructs $G_1 * G_2$, and the remaining relations are then imposed. Concretely, $G_1 \star_{\rho_1, H, \rho_2} G_2$ is generated by the images of G_1 and G_2 , subject to the relations holding in G_1 , those holding in G_2 , and the identifications $j_1\rho_1(h) = j_2\rho_2(h)$.

Remark 6.11. The construction says nothing about how large P is, and in general j_i need not be injective. When both ρ_i are injective there is a normal form theorem: choosing coset representatives for $\rho_i(H)$ in G_i , every element of $G_1 *_H G_2$ is uniquely an alternating product of non-trivial such representatives times an element of H . In particular the j_i are then injective and $j_1(G_1) \cap j_2(G_2) = j_1\rho_1(H)$. We shall not need this.

6.3. Statement of the Theorem.

Theorem 6.12 (Van Kampen). *Let X be a path connected space, $X = U \cup V$ with U and V open, and let $x_0 \in U \cap V$ with $U \cap V$ path connected. Let $j: U \cap V \rightarrow U$ and $k: U \cap V \rightarrow V$ be the inclusions. Then the square of inclusions induces a pushout square of groups*

$$\begin{array}{ccc} \pi_1(U \cap V, x_0) & \xrightarrow{j_*} & \pi_1(U, x_0) \\ k_* \downarrow & & \downarrow \\ \pi_1(V, x_0) & \longrightarrow & \pi_1(X, x_0) \end{array}$$

that is,

$$\pi_1(X, x_0) \cong \pi_1(U, x_0) \star_{j_*, \pi_1(U \cap V, x_0), k_*} \pi_1(V, x_0).$$

The inclusions $i_U: U \rightarrow X$ and $i_V: V \rightarrow X$ induce maps $(i_U)_*: \pi_1(U, x_0) \rightarrow \pi_1(X, x_0)$ and $(i_V)_*: \pi_1(V, x_0) \rightarrow \pi_1(X, x_0)$, and clearly they satisfy $(i_U)_* \circ j_* = (i_V)_* \circ k_*$, so, by the universal property of the amalgamated free product, together these maps define a homomorphism

$$\Phi: \pi_1(U, x_0) \star_{j_*, \pi_1(U \cap V, x_0), k_*} \pi_1(V, x_0) \longrightarrow \pi_1(X, x_0).$$

We shall show that Φ is an isomorphism.

Remark 6.13. Path connectivity of U and V is not among the hypotheses of Theorem 6.12 because it follows from that of $U \cap V$ and of X : if X is path connected, $X = U \cup V$ with U and V open, and $U \cap V$ is non-empty and path connected, then U and V are path connected. Indeed, let $p \in U$ and fix $q \in U \cap V$. Choose a path γ in X from p to q ; applying Lemma 6.1 to the open cover $\{\gamma^{-1}(U), \gamma^{-1}(V)\}$ of I gives a subdivision $0 = t_0 < \dots < t_k = 1$ such that each $\gamma|_{[t_{i-1}, t_i]}$ has image in U or in V ; label the subintervals accordingly. If possible, merge two adjacent intervals with the same label

into one interval and keep the label on the new interval. Continue this process until there are no neighboring intervals with the same label. At this point the labels on the intervals alternate between U and V . Let J be an interval labelled by V . Then $\gamma(\partial J) \subset V$, since $\gamma(J) \subset V$. Also, $\gamma(\partial J) \subset U$, since each endpoint of J is either also an endpoint of a neighboring interval, which by construction is labelled by U , or is one of p or q . Replace $\gamma|_J$ by a path in $U \cap V$ with the same endpoints, which is possible since $U \cap V$ is path connected. Doing this for every V -labelled interval produces a path in U from p to q . Hence U is path connected, and V is so by the same argument. Both are used in the proof of the theorem.

Definition 6.14. A *partition* of an interval $[a, b]$ is determined by points $a = t_0 < t_1 < \dots < t_k = b$: it is the collection of intervals $\{I_j\}_{j=1}^k$, where $I_j = [t_{j-1}, t_j]$. A *partition* of a map $\omega: [a, b] \rightarrow X$ is a partition of its domain, and a $\{U, V\}$ -*coloring* of a partition of ω is a function

$$c: \{1, \dots, k\} \longrightarrow \{U, V\} \quad \text{with} \quad \omega(I_j) \subseteq c(j) \quad \text{for every } j.$$

In this case we say that we have a $\{U, V\}$ -*colored partition* of ω . Such a partition is *based* if $\omega(t_j) = x_0$ for all j .

Lemma 6.15. (1.) *Let $\nu: [a, b] \rightarrow X$ be any map. Then there is a $\{U, V\}$ -colored partition of ν .*

(2.) *Let ν satisfy $\nu(a) = \nu(b) = x_0$, and let a $\{U, V\}$ -colored partition of ν be given, with intervals $I_j = [t_{j-1}, t_j]$ and coloring c . Then there is a homotopy $H: [a, b] \times [0, 1] \rightarrow X$ such that, writing $\nu_t(s) = H(s, t)$,*

- (i) $\nu_0 = \nu$;
- (ii) $H(a, t) = H(b, t) = x_0$ for all $t \in [0, 1]$;
- (iii) *for every t the given partition of $[a, b]$ and the given coloring c are a $\{U, V\}$ -colored partition of ν_t .*
- (iv) ν_1 *is based.*

Proof. 1. Given ν , consider the open cover of $[a, b]$ by the two sets $\nu^{-1}(U)$ and $\nu^{-1}(V)$. By Lemma 6.1 there is an $\varepsilon > 0$ such that every sub-interval of $[a, b]$ of length less than ε is contained in $\nu^{-1}(U)$ or in $\nu^{-1}(V)$. Any partition of $[a, b]$ into intervals of length less than ε then has a $\{U, V\}$ -coloring.

2. Now suppose given a $\{U, V\}$ -colored partition of ν . For an interior division point t_j , $0 < j < k$, put $W_j = c(j) \cap c(j+1)$. Since W_j is either U , V or $U \cap V$, it is open, path connected by Remark 6.13, and contains x_0 ; we set $W_0 = W_k = \emptyset$, there being no division point to move at the two ends. Also $\nu(t_j) \in W_j$, since t_j lies in both I_j and I_{j+1} . Choose a path ω_j in W_j from $\nu(t_j)$ to x_0 . Apply Lemma 6.6 to the interval $[a, b]$ (being identified with I) at the point t_j , with the path ω_j , taking for A_0 a closed interval in the open interval (a, b) centered at t_j , contained in $\nu^{-1}(W_j)$, and so small that these intervals are pairwise disjoint for different j . The lemma provides $N_j \subseteq A_0$ and a homotopy supported in N_j whose values lie in $\nu(N_j) \cup \omega_j(I) \subseteq W_j$. As the supports of these homotopies are disjoint, the $k-1$ homotopies can be performed simultaneously: let H be the homotopy of Lemma 6.6 at t_j on $N_j \times [0, 1]$, and the constant homotopy elsewhere.

Then (i) is part (i) of Lemma 6.6, and (ii) holds because a and b lie outside every N_j , so that $H(a, t) = \nu(a) = x_0$ and $H(b, t) = \nu(b) = x_0$. For (iii), the homotopy moves points

only inside the sets N_j , where its values lie in W_j ; hence

$$\nu_t(I_j) \subseteq \nu(I_j) \cup W_{j-1} \cup W_j \subseteq c(j),$$

since $W_{j-1} = c(j-1) \cap c(j)$ and $W_j = c(j) \cap c(j+1)$ are both contained in $c(j)$. Finally $\nu_1(t_j) = \omega_j(1) = x_0$ for $0 < j < k$ by part (ii) of Lemma 6.6, while $\nu_1(t_0) = \nu(a) = x_0$ and $\nu_1(t_k) = \nu(b) = x_0$ by (ii) above; so ν_1 is based. $\square$

Corollary 6.16. *The homomorphism Φ defined after the statement of Theorem 6.12 is surjective.*

Proof. Let $\omega: I \rightarrow X$ be a loop based at x_0 . By Part 1 of Lemma 6.15 ω carries a $\{U, V\}$ -colored partition and Part 2 of the same lemma produces a homotopy $\text{rel } \partial I$ from ω to a loop ω_1 at x_0 carrying a *based* $\{U, V\}$ -colored partition, with the same intervals I_j and the same coloring c .

For each j let γ_j be $\omega_1|_{I_j}$ reparametrized linearly by I . The partition being based, γ_j is a loop at x_0 , and its image lies in $c(j)$; so $[\gamma_j] = (i_{c(j)})_*(w_j)$, where w_j is the class of γ_j in $\pi_1(c(j), x_0)$. Moreover $\gamma_1 \cdot \gamma_2 \cdots \gamma_k$, bracketed in any way, is ω_1 composed with a piecewise linear map $\varphi: I \rightarrow I$ fixing 0 and 1, so that by Lemma 3.5

$$[\omega] = [\omega_1] = [\gamma_1][\gamma_2] \cdots [\gamma_k].$$

Now the image of Φ is a subgroup of $\pi_1(X, x_0)$ containing the images of $(i_U)_*$ and of $(i_V)_*$, because composing Φ with the two structure homomorphisms of the amalgamated free product gives $(i_U)_*$ and $(i_V)_*$. It therefore contains every $[\gamma_j]$, hence their product $[\omega]$. So Φ is surjective. $\square$

The proof that Φ is injective is a two-dimensional analogue of this argument.

We need to show that if two words in the elements of $\pi_1(U, x_0) \cup \pi_1(V, x_0)$ have the same image in $\pi_1(X, x_0)$, then they are equal in the amalgamated free product. Here a *word* means a finite sequence $((w_1, c_1), \dots, (w_k, c_k))$ in which each c_j is U or V — a color, in the sense of Definition 6.14 — and $w_j \in \pi_1(c_j, x_0)$; by Proposition 6.9 every element of the amalgamated free product is represented by such a word, and the elementary relations of that proposition become the following moves on words, neither of which changes the element of the amalgamated free product represented by the word:

- (W1) replacing two adjacent letters of the same color by their product in that group, or conversely — in particular, deleting or inserting a letter equal to the identity;
- (W2) replacing a letter $(j_*(w), U)$ by the letter $(k_*(w), V)$, for $w \in \pi_1(U \cap V, x_0)$, or conversely.

Any word is represented by a based $\{U, V\}$ -colored loop, by which we mean a loop at x_0 together with a based $\{U, V\}$ -colored partition of it: choose for each j a loop γ_j at x_0 in c_j representing w_j , and let ω be the concatenation $\gamma_1 \cdot \gamma_2 \cdots \gamma_k$, with the partition used in forming it and the coloring $j \mapsto c_j$. Conversely a based $\{U, V\}$ -colored loop ω determines a word, the j -th letter being the class in $\pi_1(c(j), x_0)$ of the j -th piece of ω , reparametrized by I . As in the proof of Corollary 6.16, the image under Φ of the element represented by a word is the class of the corresponding loop. Consequently it is enough to prove the following: if two based $\{U, V\}$ -colored loops are homotopic as based loops in X , then their words are carried one to the other by a sequence of the moves (W1) and (W2). For if Φ takes the same value on the elements represented by two words, the

corresponding loops are homotopic rel ∂I , so the two words are related by such a sequence and the two elements are equal.

Let then $\omega_0, \omega_1: [a, b] \rightarrow X$ be based $\{U, V\}$ -colored loops at x_0 , and let $H: [a, b] \times [0, 1] \rightarrow X$ be a homotopy of based loops from ω_0 to ω_1 .

Definition 6.17. Let $H: [a, b] \times [0, 1] \rightarrow X$ be a homotopy. A *decomposition* of H is given by sequences $a = t_0 < t_1 < \dots < t_k = b$ and $0 = s_0 < s_1 < \dots < s_\ell = 1$: it consists of the rectangles

$$R_{i,j} = [t_{i-1}, t_i] \times [s_{j-1}, s_j], \quad 1 \leq i \leq k, \quad 1 \leq j \leq \ell.$$

A $\{U, V\}$ -coloring of the decomposition is a function

$$c: \{1, \dots, k\} \times \{1, \dots, \ell\} \longrightarrow \{U, V\} \quad \text{with} \quad H(R_{i,j}) \subseteq c(i, j) \quad \text{for all } i, j,$$

and in this case we say that we have a $\{U, V\}$ -colored decomposition of H . Such a decomposition is *based* if $H(t_i, s_j) = x_0$ for all (i, j) .

Lemma 6.18. Let $H: [a, b] \times [0, 1] \rightarrow X$ be a homotopy of loops based at x_0 , that is, one for which $\omega_s = H(\cdot, s)$ satisfies $\omega_s(a) = \omega_s(b) = x_0$ for every s , and let based $\{U, V\}$ -colored partitions $\mathcal{P}_0$ and $\mathcal{P}_1$ of ω_0 and of ω_1 be given. Then there is a homotopy H' of loops based at x_0 and a based $\{U, V\}$ -colored decomposition of H' such that

- (i) the division points of the given partitions $\mathcal{P}_0$ of ω_0 and $\mathcal{P}_1$ of ω_1 all occur among the t_i for the partition of H' ;
- (ii) H' agrees with H outside a disjoint union of arbitrarily small neighborhoods of the vertices (t_i, s_j) of the decomposition;
- (iii) for every interval J of the given partition of ω_0 the paths $H'(\cdot, 0)|_J$ and $\omega_0|_J$ are homotopic rel ∂J through maps into the color of J , and likewise for ω_1 and $H'(\cdot, 1)$.

Proof. The sets $H^{-1}(U)$ and $H^{-1}(V)$ form an open cover of the compact metric space $[a, b] \times [0, 1]$, so by Lemma 6.1 we may choose sequences $a = t_0 < \dots < t_k = b$ and $0 = s_0 < \dots < s_\ell = 1$ containing the division points of the two given partitions and fine enough that every rectangle $R_{i,j}$ has diameter less than the Lebesgue number. For each pair i, j choose $c(i, j) \in \{U, V\}$ with $H(R_{i,j}) \subseteq c(i, j)$: this is a $\{U, V\}$ -colored decomposition of H , not yet based.

For a vertex $v = (t_i, s_j)$ with $t_i \neq a, b$, let W_v be the intersection of the colors of the rectangles that contain v , intersected further with the colors of the intervals of the given partition of ω_0 containing t_i if $j = 0$, and with the colors of the intervals of the given partition of ω_1 containing t_i if $j = \ell$. As a non-empty intersection of members of $\{U, V\}$ the set W_v is U , V or $U \cap V$, hence open, path connected by Remark 6.13 and containing x_0 ; and $H(v) \in W_v$. Choose a path ν_v in W_v from $H(v)$ to x_0 , constant when $H(v) = x_0$, and apply Lemma 6.6 with $n = 2$ at the point v , with the path ν_v , taking for A_0 a neighborhood of v contained in $H^{-1}(W_v)$, meeting no rectangle that does not contain v , disjoint from the edges $t = a$ and $t = b$, and so small that these neighborhoods are pairwise disjoint for different v . The lemma provides $N_v \subseteq A_0$ and a homotopy supported in N_v whose values lie in $H(N_v) \cup \nu_v(I) \subseteq W_v$. The supports being disjoint, all these homotopies may be performed simultaneously; let H' be the resulting map, so that $H' = H$ off $\bigcup_v N_v$ and $H'(N_v) \subseteq W_v$. This is (ii), the N_v being as small as we please, and (i) holds by the choice of the t_i .

The given decomposition and coloring serve for H' : a point at which H' differs from H lies in some N_v , hence in rectangles containing v only, and there the values of H' lie in W_v , which is contained in the color of each of those rectangles. It is based because $H'(v) = \nu_v(1) = x_0$ at every vertex with $t_i \neq a, b$, by part (ii) of Lemma 6.6, while the remaining vertices lie on the edges $t = a$ and $t = b$, where $H' = H$ is constantly x_0 . Also H' is again a homotopy of loops based at x_0 , since the N_v are disjoint from those two edges.

Finally, let J be an interval of the original partition $\mathcal{P}_0$ of ω_0 . Along the bottom edge H' differs from H only inside the sets N_v for vertices $v = (t_i, 0)$ with $t_i \in J$, where its values lie in W_v , contained by construction in the color of J ; and the endpoints of J are division points of a *based* partition, so they are carried to x_0 by both maps. The homotopies of Lemma 6.6 at those vertices, restricted to the bottom edge, therefore give a homotopy rel ∂J from $\omega_0|_J$ to $H'(\cdot, 0)|_J$ through maps into the color of J . The same argument applies at $s = 1$, and (iii) follows. $\square$

The two horizontal edges of the decomposition of H' carry based $\{U, V\}$ -colored partitions: the subdivision $\{t_i\}$ with the coloring $i \mapsto c(i, 1)$ for $H'(\cdot, 0)$, and the same subdivision with the coloring $i \mapsto c(i, \ell)$ for $H'(\cdot, 1)$. They are based because the decomposition is, and colored because the bottom and top edges of $R_{i,1}$ and of $R_{i,\ell}$ are carried into $c(i, 1)$ and $c(i, \ell)$.

Corollary 6.19. *With ω_0, ω_1, H and H' as in Lemma 6.18, if the words determined by the above $\{U, V\}$ -colored partitions of $H'(\cdot, 0)$ and of $H'(\cdot, 1)$ represent the same element of the amalgamated free product, then so do the words determined by the given $\{U, V\}$ -colored partitions of ω_0 and of ω_1 .*

Proof. By (i) of Lemma 6.18, each interval J of the given partition of ω_0 is a union of consecutive intervals $[t_{i-1}, t_i]$. By (iii) the path $H'(\cdot, 0)|_J$ has image in the color of J ; each $H'(\cdot, 0)|_{[t_{i-1}, t_i]}$ therefore has image in the color of J as well as in $c(i, 1)$. If those two colors differ their intersection is $U \cap V$, so the class of that piece comes from $\pi_1(U \cap V, x_0)$ and the move (W2) recolors the corresponding letter by the color of J . Once all the letters lying over J carry the color of J , the move (W1) merges them into the single letter whose class is that of $H'(\cdot, 0)|_J$, and by (iii) that class is the class of $\omega_0|_J$. Carrying this out over every J turns the word of $H'(\cdot, 0)$ into the word of ω_0 , so the two represent the same element of the amalgamated free product; and the same argument applies to $H'(\cdot, 1)$ and ω_1 . The corollary follows. $\square$

Proof of Theorem 6.12. By Corollary 6.16 the map Φ is surjective, so only its injectivity remains, and by the reduction made above it is enough to prove that two homotopic based $\{U, V\}$ -colored loops have words representing the same element of the amalgamated free product. So let H be a homotopy of based loops from ω_0 to ω_1 , and let H' , with its based $\{U, V\}$ -colored decomposition, be as in Lemma 6.18. By Corollary 6.19 it suffices to treat the two horizontal edges of that decomposition.

This will follow once we know that for each j the words determined by the loops $H'(\cdot, s_{j-1})$ and $H'(\cdot, s_j)$, in the coloring of the j -th row of rectangles, represent the same element of the amalgamated free product.

For $0 \leq j \leq \ell$ put $\lambda_j = H'(\cdot, s_j)$, a loop based at x_0 . For each j with $1 \leq j \leq \ell$ the subdivision $\{t_i\}$ with the coloring $i \mapsto c(i, j)$ of the j -th row of rectangles is a based

$\{U, V\}$ -colored partition of λ_{j-1} and also of λ_j : all vertices go to x_0 , and the bottom and top edges of $R_{i,j}$ are carried into $c(i, j)$. Passing from the coloring of one row to that of the next changes the word of λ_j only by moves (W2), since where $c(i, j)$ and $c(i, j+1)$ differ the piece of λ_j over $[t_{i-1}, t_i]$ has image in their intersection $U \cap V$. A straightforward induction on j therefore reduces us to showing, for one j , that the words of λ_{j-1} and of λ_j in the coloring of the j -th row represent the same element. We do this for $j = 1$; the argument is the same for every j .

Write $c_i = c(i, 1)$. Let b_i and τ_i be the restrictions of λ_0 and of λ_1 to $[t_{i-1}, t_i]$, reparametrized by I — the base and the top of the rectangle $R_{i,1}$ — and let u_i be its left edge traversed upwards and d_i its right edge traversed downwards. The decomposition being based, all four are loops at x_0 ; and all four have image in c_i , being edges of $R_{i,1}$. The rectangle is convex, so the two paths in it from the lower left to the lower right corner — the base on the one hand, the left edge followed by the top followed by the right edge on the other — are homotopic rel endpoints; composing that homotopy with $H'|_{R_{i,1}}$, whose values lie in c_i , gives

$$b_i \simeq u_i \cdot \tau_i \cdot d_i \quad \text{rel } \partial I, \text{ through loops in } c_i,$$

so that $[b_i] = [u_i][\tau_i][d_i]$ in $\pi_1(c_i, x_0)$. By (W1) we may therefore replace the letter (b_i, c_i) of the word of λ_0 by the three letters (u_i, c_i) , (τ_i, c_i) , (d_i, c_i) . Carrying this out for every i turns that word into

$$(u_1, c_1) (\tau_1, c_1) (d_1, c_1) (u_2, c_2) (\tau_2, c_2) (d_2, c_2) \cdots (u_k, c_k) (\tau_k, c_k) (d_k, c_k).$$

Now u_1 and d_k lie over $t = a$ and $t = b$, where H' is constantly x_0 ; they are constant loops, and (W1) deletes them. For $1 \leq i < k$ the loops d_i and u_{i+1} are the edge over $t = t_i$ traversed in opposite directions, so $u_{i+1} = \bar{d}_i$. If $c_i = c_{i+1}$ these two adjacent letters have the same color and their classes are inverse to one another in $\pi_1(c_i, x_0)$, so (W1) merges them into the identity and deletes them. If $c_i \neq c_{i+1}$ then that edge has image in $c_i \cap c_{i+1} = U \cap V$, so (W2) recolors (u_{i+1}, c_{i+1}) to (u_{i+1}, c_i) and the same cancellation applies. What is left is

$$(\tau_1, c_1) (\tau_2, c_2) \cdots (\tau_k, c_k),$$

the word of λ_1 in the coloring of the first row. This completes the induction, and with it the proof of the theorem. $\square$

Example 6.20. Take $X = S^n$ with $n \geq 2$, U and V the complements of the south and north poles. Both are homeomorphic to $\mathbb{R}^n$, hence simply connected, and $U \cap V \simeq S^{n-1}$ is path connected because $n \geq 2$. The pushout of $1 \leftarrow \pi_1(S^{n-1}) \rightarrow 1$ is trivial, so $\pi_1(S^n) = 1$ for $n \geq 2$. Note where the hypothesis fails at $n = 1$: there $U \cap V$ has two components, and indeed $\pi_1(S^1) = \mathbb{Z}$ by Theorem 5.3.

Example 6.21. For a wedge of two circles $X = S^1 \vee S^1$, take U and V to be the two circles slightly thickened at the wedge point, so that each deformation retracts to a circle and $U \cap V$ is contractible. The pushout of $\mathbb{Z} \leftarrow 1 \rightarrow \mathbb{Z}$ is the free product, so $\pi_1(S^1 \vee S^1) \cong \mathbb{Z} * \mathbb{Z}$, the free group on two generators. In particular π_1 is not abelian in general, as promised in §4.

The rest of the material in this note is of a more technical nature and will be used in the later notes. We encourage the reader to consult this material as it comes up.

7. COMPRESSION CRITERIA AND FORGETTING THE BASE POINT

This section collects, in one place, the results that are used over and over to recognize when a map can be pushed into a subspace, and to pass between pointed and unpointed homotopy classes. They are of two kinds. The first kind compares $\pi_n(X, x_0)$, or its relative analogue, with the corresponding set of free homotopy classes, redeeming the promise of Remark 2.4: the base point may be forgotten at the cost of surjectivity only, and it may be forgotten altogether when the relevant fundamental group is trivial. The second kind says that the vanishing of a relative class is exactly the possibility of compressing the map into the subspace, and upgrades a compression through maps of pairs to one rel the boundary. Apart from Proposition 7.2, which is a direct manipulation of the definition, everything here comes from §6.1: the two comparisons rest on the homotopy extension property, Lemma 6.4, and on its bound on tracks, Remark 6.5, and the compression lemma rests directly on the cylinder retraction, Lemma 6.3, that underlies them. Nothing else from §§5 and 6 is used, and nothing here is used in them.

Proposition 7.1. *Let $n \geq 1$ and let X be path connected. Write $[(I^n, \partial I^n), X]$ for the set of homotopy classes of maps $f: I^n \rightarrow X$ that are constant on ∂I^n , two such being identified when they are joined by a homotopy through maps constant on ∂I^n , the constant value being allowed to vary; this is the set $[S^n, X]$ of free homotopy classes of maps $S^n \rightarrow X$. Forgetting the base point gives a function*

$$\pi_n(X, x_0) \longrightarrow [(I^n, \partial I^n), X],$$

which is onto, and is a bijection when $\pi_1(X, x_0)$ is trivial.

Proof. The identification with $[S^n, X]$ is the argument of Proposition 4.2: a map constant on ∂I^n factors uniquely through $I^n / \partial I^n = S^n$, and $q \times \text{id}_I$, being a quotient map, does the same for homotopies. The function is defined because a homotopy rel ∂I^n is in particular a homotopy through maps constant on ∂I^n . Throughout, I^n and $I^n \times I$ are triangulated as in Lemma 6.2(iii), so that the subsets named below, each of them a union of faces, are subcomplexes and Lemma 6.4 applies.

Surjectivity. Let $f: I^n \rightarrow X$ be constant on ∂I^n , with value y there, and let γ be a path in X from y to x_0 . Then $(s, t) \mapsto \gamma(t)$ is a homotopy of $f|_{\partial I^n}$ beginning at that restriction; extend it over I^n by Lemma 6.4. At every time the extension is constant on ∂I^n , so this is a homotopy of the kind just described, and it ends at a map constant with value x_0 on ∂I^n , that is, at a map of pairs $(I^n, \partial I^n) \rightarrow (X, x_0)$.

Injectivity. Suppose $\pi_1(X, x_0) = 1$, let φ_0 and φ_1 be maps of pairs $(I^n, \partial I^n) \rightarrow (X, x_0)$, and let $H: I^n \times I \rightarrow X$ be a homotopy from φ_0 to φ_1 through maps constant on ∂I^n , with value $\gamma(t)$ there at time t . Then γ is a loop in X at x_0 , so by hypothesis there is $\Gamma: I \times I \rightarrow X$ with $\Gamma(t, 0) = \gamma(t)$, $\Gamma(t, 1) = x_0$ and $\Gamma(0, u) = \Gamma(1, u) = x_0$. On the subcomplex $\partial I^n \times I \cup I^n \times \partial I$ of $I^n \times I$ take the homotopy of H given by $(s, t, u) \mapsto \Gamma(t, u)$ on the first piece and by the stationary homotopy on the second; the two agree on the

overlap $\partial I^n \times \partial I$, where Γ is constant at x_0 and H is φ_0 or φ_1 , each of which is x_0 there. By Lemma 6.4 it extends to a homotopy of H , ending at a map $H': I^n \times I \rightarrow X$ which agrees with H on $I^n \times \partial I$ and carries $\partial I^n \times I$ to x_0 : that is, a homotopy rel ∂I^n from φ_0 to φ_1 . $\square$

We turn to the relative results. The first is the criterion, already implicit in Definition 4.9, that identifies the vanishing of a relative class.

Proposition 7.2 (compression criterion). *Let $n \geq 1$ and let f be a map of triples. Then $[f]$ is the base point of $\pi_n(X, A, x_0)$ if and only if f is homotopic, through maps of triples, to a map whose image lies in A .*

Proof. If $[f] = [c]$ then f is homotopic to the constant map, whose image $\{x_0\}$ lies in A . Conversely, suppose $f \simeq g$ through maps of triples with $g(I^n) \subseteq A$, and put

$$h_t(s_1, \dots, s_n) = g(s_1, \dots, s_{n-1}, (1-t)s_n + t).$$

Each h_t is a map of triples: its image lies in that of g , hence in A ; and if $s \in J^{n-1}$ then either $s_n = 1$, so that the argument lies in the face $s_n = 1$ and g takes the value x_0 there, or $s_j \in \{0, 1\}$ for some $j < n$, so that the argument lies in a side face, where again $g = x_0$. Thus h is a homotopy of maps of triples from g to the constant map at x_0 , and $[f] = [g] = [c]$. $\square$

The comparison of Proposition 7.1 also has a relative form, proved in the same way but with J^{n-1} in place of the base point; the extra work is that J^{n-1} must first be contracted.

Proposition 7.3. *Let $n \geq 1$, let A be path connected, and write $[(I^n, \partial I^n), (X, A)]$ for the set of homotopy classes of maps of pairs $(I^n, \partial I^n) \rightarrow (X, A)$, no base point being imposed. Forgetting the base point gives a function*

$$\pi_n(X, A, x_0) \longrightarrow [(I^n, \partial I^n), (X, A)],$$

which is onto, and is a bijection when $\pi_1(A, x_0)$ is trivial.

Proof. The map is defined because a homotopy through maps of triples is in particular a homotopy of maps of pairs. Throughout, I^n and $I^n \times I$ are triangulated as in Lemma 6.2(iii), so that the subsets named below, each of them a union of faces, are subcomplexes and Lemma 6.4 applies.

Surjectivity. Let $f: (I^n, \partial I^n) \rightarrow (X, A)$ be a map of pairs. The set J^{n-1} is homeomorphic to I^{n-1} , hence contractible; let c_s be a contraction of it to a point p , with c_0 the identity and c_1 constant. Then $s \mapsto f \circ c_s$ is a homotopy of $f|_{J^{n-1}}$ inside A ending at the constant map at $f(p)$, and following it by a path in A from $f(p)$ to x_0 , taken constant on J^{n-1} , gives a homotopy of $f|_{J^{n-1}}$ inside A ending at the constant map at x_0 . Extend it over ∂I^n inside A by Lemma 6.4, and then over I^n inside X by Lemma 6.4 again. Throughout, ∂I^n stays in A , so this is a homotopy of maps of pairs, and it ends at a map of triples.

Injectivity. Suppose $\pi_1(A, x_0) = 1$, let φ_0 and φ_1 be maps of triples, and let $H: I^n \times I \rightarrow X$ be a homotopy of maps of pairs from φ_0 to φ_1 , so that $H(\partial I^n \times I) \subseteq A$. We deform H , keeping its two ends fixed, until it is a homotopy through maps of triples.

With p and c_s as above put

$$K(y, t, s) = H(c_s(y), t), \quad y \in J^{n-1}, \quad t, s \in I.$$

This is a homotopy of $H|_{J^{n-1} \times I}$ inside A , from $H|_{J^{n-1} \times I}$ at $s = 0$ to $(y, t) \mapsto \gamma(t)$ at $s = 1$, where $\gamma(t) = H(p, t)$; and it is stationary on $J^{n-1} \times \partial I$, because there H is φ_0 or φ_1 , each constant with value x_0 on J^{n-1} . In particular γ is a loop in A at x_0 , so by hypothesis there is $\Gamma: I \times I \rightarrow A$ with $\Gamma(t, 0) = \gamma(t)$, $\Gamma(t, 1) = x_0$ and $\Gamma(0, u) = \Gamma(1, u) = x_0$. Then $(y, t, u) \mapsto \Gamma(t, u)$ continues the deformation, again stationary on $J^{n-1} \times \partial I$, and ends at the constant map at x_0 . Concatenating gives a homotopy L of $H|_{J^{n-1} \times I}$ inside A , stationary on $J^{n-1} \times \partial I$, from $H|_{J^{n-1} \times I}$ to the constant map at x_0 .

Now extend. On the subcomplex $J^{n-1} \times I \cup \partial I^n \times \partial I$ of $\partial I^n \times I$ take L together with the stationary homotopy; the two agree on the overlap $J^{n-1} \times \partial I$, where L is stationary. By Lemma 6.4 with target A this extends to a homotopy of $H|_{\partial I^n \times I}$ inside A . That homotopy, together with the stationary one on $I^n \times \partial I$, extends over $I^n \times I$ with target X by Lemma 6.4 again. The result is a map $H': I^n \times I \rightarrow X$ agreeing with H on $I^n \times \partial I$, carrying $\partial I^n \times I$ into A and $J^{n-1} \times I$ to x_0 : that is, a homotopy through maps of triples from φ_0 to φ_1 . $\square$

Two results are recorded here for later use, in *The Hurewicz theorems* and in *CW complexes*. The first is purely formal, with no hypothesis whatever on the pair; the second combines it with the surjectivity half of Proposition 7.3, for which path connectedness of A is enough — simple connectivity is what injectivity needs, and neither statement uses injectivity.

Lemma 7.4 (compression rel the boundary). *Let $j \geq 1$, let U be a compact convex subset with non-empty interior of a j -dimensional affine space, with ∂U its boundary there, let $A \subseteq X$, and let $g: (U, \partial U) \rightarrow (X, A)$ be a map of pairs which is homotopic, through maps of pairs, to a map with image in A . Then g is homotopic rel ∂U to a map with image in A .*

Proof. Let $F: U \times I \rightarrow X$ be a homotopy of maps of pairs with $F_0 = g$ and $F_1(U) \subseteq A$; since every F_t is a map of pairs, $F(\partial U \times I) \subseteq A$.

Let $C = U \times \{1\} \cup \partial U \times I$, and let $\Phi_1: U \rightarrow C$ be the homeomorphism obtained from the radial projection

$$r: U \times I \longrightarrow U \times \{0\} \cup \partial U \times I$$

of Lemma 6.3, by restricting it to the top face $U \times \{1\}$ and reflecting the I -coordinate; it is $x \mapsto (x, 0)$ on ∂U . Write $\Phi_0(x) = (x, 0)$. The two maps agree on ∂U , and $U \times I$ is convex, so

$$\Phi_t(x) = (1 - t)\Phi_0(x) + t\Phi_1(x)$$

is a homotopy from Φ_0 to Φ_1 , stationary on ∂U .

Put $H_t = F \circ \Phi_t$. Then $H_0 = g$, and $H_t(x) = F(x, 0) = g(x)$ for $x \in \partial U$, so H is a homotopy rel ∂U . Its end $H_1 = F \circ \Phi_1$ has image in $F(C) \subseteq A$, since $F_1(U) \subseteq A$ and $F(\partial U \times I) \subseteq A$. $\square$

Corollary 7.5. *Let $n \geq 1$, let $A \subseteq X$ be path connected with $x_0 \in A$, and suppose $\pi_n(X, A, x_0) = 0$. Then every map of pairs $f: (D^n, S^{n-1}) \rightarrow (X, A)$ is homotopic rel S^{n-1} to a map with image in A .*

Proof. By Proposition 7.3 the forgetful map $\pi_n(X, A, x_0) \rightarrow [(I^n, \partial I^n), (X, A)]$ is onto, path connectedness of A being all that is needed for that; its source is trivial, so the target consists of the class of the constant map at x_0 alone. The pairs (D^n, S^{n-1}) and $(I^n, \partial I^n)$

are homeomorphic, each being a compact convex body with its boundary, so the same holds for $[(D^n, S^{n-1}), (X, A)]$: the map f is homotopic, through maps of pairs, to the constant map at x_0 , whose image lies in A . By Lemma 7.4 it is then homotopic rel S^{n-1} to a map with image in A . $\square$

8. THE MAPPING CYLINDER

A map $f: X \rightarrow Y$ is not in general an inclusion, and many of the tools of the subject — the long exact sequence of a pair, the compression criteria — are tools for pairs. The mapping cylinder removes the difficulty once and for all: it enlarges Y to a homotopy equivalent space in which X sits as a closed subspace, in such a way that f becomes the inclusion.

Definition 8.1. The *mapping cylinder* of a map $f: X \rightarrow Y$ is

$$M_f = (X \times I \sqcup Y) / ((x, 1) \sim f(x)).$$

Write $\iota: X \rightarrow M_f$ for $x \mapsto [(x, 0)]$ and $r: M_f \rightarrow Y$ for the map with $r[(x, t)] = f(x)$ and $r[y] = y$.

Lemma 8.2. *The map ι is a homeomorphism of X onto a closed subspace of M_f ; the subspace Y is a strong deformation retract of M_f , in the sense of Definition 1.8, with retraction r ; so r is a homotopy equivalence by Proposition 1.9, and $r \circ \iota = f$.*

Proof. The subset $X \times \{0\}$ is saturated and closed in $X \times I \sqcup Y$ and the quotient map is injective on it, so ι is a closed embedding. That r is well defined and continuous is the universal property of the quotient, the two formulas agreeing at $(x, 1)$; and $r \circ \iota = f$ by inspection. Finally

$$H([(x, t)], s) = [(x, (1-s)t + s)], \quad H([y], s) = [y],$$

is a well-defined homotopy from the identity of M_f to the inclusion of Y followed by r , stationary on Y ; it is continuous because the quotient map times id_I is a quotient map, I being locally compact. $\square$

Corollary 8.3. *Every map $f: X \rightarrow Y$ factors as*

$$X \xrightarrow{\iota} M_f \xrightarrow{r} Y$$

with ι a closed embedding and r a homotopy equivalence. Identifying X with $\iota(X)$, the map f is the inclusion of X in M_f followed by a homotopy equivalence.

Proof. This is Lemma 8.2. $\square$

So any property of f invariant under composition with a homotopy equivalence may be read off from the pair (M_f, X) . The next remark says what the reading is.

Remark 8.4. Identify X with its image in $M = M_f$ and let $x_0 \in X$. Since r is a homotopy equivalence it induces isomorphisms on all homotopy groups, and under them the maps induced by the inclusion $X \subseteq M$ become those induced by f , because $r \circ \iota = f$. The long exact sequence of the pair (M, X) ,

$$\cdots \rightarrow \pi_n(X, x_0) \rightarrow \pi_n(M, x_0) \rightarrow \pi_n(M, X, x_0) \rightarrow \pi_{n-1}(X, x_0) \rightarrow \cdots$$

of Definition 4.14, therefore measures the failure of f_* to be an isomorphism: it is an isomorphism on π_n for every n exactly when $\pi_n(M, X, x_0) = 0$ for every n . The same reading applies to any other functor of pairs possessing a long exact sequence and carrying homotopy equivalences to isomorphisms.

Remark 8.5. If f is a pointed map $(X, x_0) \rightarrow (Y, y_0)$, point M_f by $\iota(x_0)$. Then ι is pointed, and r is pointed because $r\iota(x_0) = f(x_0) = y_0$; but the deformation retraction H of Lemma 8.2 moves that base point along the segment $\{x_0\} \times I$ rather than fixing it. What H fixes is Y . When a pointed homotopy equivalence is wanted one takes instead the base point $y_0 \in Y \subseteq M_f$, and then ι is pointed only up to the path $s \mapsto [(x_0, s)]$. Which of the two base points is in force will always be said.

Remark 8.6. The construction is functorial for squares. Given $f: X \rightarrow Y$ and $f': X' \rightarrow Y'$ and maps $a: X \rightarrow X'$, $b: Y \rightarrow Y'$ with $f' \circ a = b \circ f$, the formulas $[(x, t)] \mapsto [(a(x), t)]$ and $[y] \mapsto [b(y)]$ define a map $M_f \rightarrow M_{f'}$ carrying X to X' and Y to Y' and commuting with the ι 's and the r 's. It is well defined and continuous by the universal property of the quotient, the two formulas agreeing at $(x, 1)$ precisely because the square commutes. So (M_f, X) is a functor of the map f , and the sequence of Remark 8.4 is natural for such squares.

9. READING A CLASS OFF A CONVEX BODY

The geometric picture of Remark 4.7, of maps shrunk into disjoint sub-cubes, is worth turning into a usable statement. Two things are wanted. First, a class in $\pi_n(X, x_0)$ attached to a map defined not on I^n but on some other convex body — a face of a simplex, a piece of a subdivision, a ball. Second, the statement that a map which is the base point outside finitely many disjoint bodies represents the sum of the classes they carry, so that a map defined on a subdivided cube, or on the boundary of a simplex, can be read off face by face. The device for the first is to carry the body into I^n by an affine linear map and to fill in the rest of I^n with the base point. We begin with an elementary lemma about moving disjoint convex bodies around inside the cube, which serves twice over: to show that the construction does not depend on the affine linear map chosen, and to prove additivity.

9.1. Moving disjoint bodies. Throughout this subsection $n \geq 2$, and we write

$$\mu_{p,\lambda}(y) = p + \lambda(y - p)$$

for the homothety in $\mathbb{R}^n$ with centre p and ratio λ , so that $\mu_{p,1}$ is the identity and $\mu_{p,\lambda}$ has determinant λ^n .

Lemma 9.1. *Let $m \geq 1$ and let $(x_1, \dots, x_m)$ and $(y_1, \dots, y_m)$ be m -tuples of distinct points of $\text{int } I^n$. Then there is a path*

$$t \mapsto (\gamma_1(t), \dots, \gamma_m(t)), \quad 0 \leq t \leq 1,$$

of m -tuples of distinct points of $\text{int } I^n$ with $\gamma_i(0) = x_i$ and $\gamma_i(1) = y_i$ for every i .

Proof. We use the following remark: if $V \subseteq \mathbb{R}^n$ is open and connected and $\Xi \subseteq V$ is finite, then $V \setminus \Xi$ is path connected. Indeed V is path connected, and a path in V between points not in Ξ can be pushed off any $w \in \Xi$ that it meets. Choose $\rho > 0$ with the closed

ball $\bar{B}(w, \rho)$ contained in V , meeting Ξ only in w , and containing neither endpoint; let t_1 be the first time the path meets the sphere $S(w, \rho)$ and t_2 the last, and replace the path between those times by an arc in $S(w, \rho)$, which exists because that sphere is path connected, n being at least 2. Before t_1 and after t_2 the path lies outside the open ball, so the new path misses w ; finitely many such modifications, carried out in disjoint balls, produce a path in $V \setminus \Xi$.

Choose now distinct points $z_1, \dots, z_m$ of $\text{int } I^n$, none of them equal to any x_i or any y_j ; this is possible because $\text{int } I^n$ is infinite. It is enough to join (x_i) to (z_i) , since the same argument joins (y_i) to (z_i) and the two paths may be concatenated. Move the points one at a time. Suppose $x_1, \dots, x_{j-1}$ have been moved to $z_1, \dots, z_{j-1}$, the other points having stayed put. Both x_j and z_j lie in

$$\text{int } I^n \setminus \{z_1, \dots, z_{j-1}, x_{j+1}, \dots, x_m\},$$

which is path connected by the remark, so x_j may be moved to z_j along a path there while the other $m-1$ points are held fixed. At every moment of every stage the m points are distinct. $\square$

Proposition 9.2. *Let $C_1, \dots, C_m$ be pairwise disjoint compact convex subsets of I^n with non-empty interiors, and let $A_1, \dots, A_m$ be affine linear maps of $\mathbb{R}^n$ with positive determinants such that the sets $A_i(C_i)$ are contained in I^n and are pairwise disjoint. Then there are paths $t \mapsto A_i(t)$ of affine linear maps with positive determinant, with $A_i(0) = \text{id}$ and $A_i(1) = A_i$, such that for every t the sets*

$$A_1(t)(C_1), \dots, A_m(t)(C_m)$$

are pairwise disjoint compact convex subsets of I^n .

Proof. Let $\mathcal{P}$ be the set of m -tuples $\Phi = (\Phi_1, \dots, \Phi_m)$ of affine linear maps of $\mathbb{R}^n$ with $\det \Phi_i > 0$ such that the sets $\Phi_i(C_i)$ lie in I^n and are pairwise disjoint, a path in $\mathcal{P}$ meaning one along which the coefficients vary continuously. The tuple $(\text{id}, \dots, \text{id})$ and the tuple $(A_1, \dots, A_m)$ both lie in $\mathcal{P}$, and the assertion is that they lie in the same path component of it. Fix $c_i \in \text{int } C_i$ once and for all. For $\Phi \in \mathcal{P}$ the set $\Phi_i(\text{int } C_i)$ is open and contained in I^n , hence contained in $\text{int } I^n$; in particular $\Phi_i(c_i) \in \text{int } I^n$.

Shrinking. Let $\Phi \in \mathcal{P}$ and put $p_i = \Phi_i(c_i)$. For $0 < \lambda \leq 1$ the tuple $(\mu_{p_i, \lambda} \circ \Phi_i)_i$ lies in $\mathcal{P}$: its determinants are $\lambda^n \det \Phi_i > 0$, and $\mu_{p_i, \lambda}(\Phi_i(C_i)) \subseteq \Phi_i(C_i)$ because $\Phi_i(C_i)$ is convex and contains p_i , so the images remain disjoint and inside I^n . Each $\mu_{p_i, \lambda} \circ \Phi_i$ still carries c_i to p_i , and the image shrinks to p_i as $\lambda \rightarrow 0$. So Φ may be joined in $\mathcal{P}$, for any prescribed $r > 0$, to a tuple Ψ with $\Psi_i(c_i) = p_i$ and $\Psi_i(C_i) \subseteq B(p_i, r)$.

Translating. Let $\Phi \in \mathcal{P}$ carry c_i to p_i with $\Phi_i(C_i) \subseteq B(p_i, r)$, and let $(\gamma_i(t))$ be a path of m -tuples of distinct points of $\text{int } I^n$ with $\gamma_i(0) = p_i$. By compactness

$$\delta = \min_t \min_{i \neq j} \left(\min |\gamma_i(t) - \gamma_j(t)|, \min_i \text{dist}(\gamma_i(t), \partial I^n) \right)$$

is positive. If $r < \delta/3$ then translating the i -th map by $\gamma_i(t) - p_i$ is a path in $\mathcal{P}$: the i -th image stays inside $B(\gamma_i(t), r)$, and these balls lie in $\text{int } I^n$ and are pairwise disjoint. At the end of the path the i -th map carries c_i to $\gamma_i(1)$.

Turning. Suppose Φ and Φ' in $\mathcal{P}$ both carry c_i to the same point q_i , and that $B(q_i, r) \subseteq \text{int } I^n$, these balls being pairwise disjoint. Write

$$\Phi_i(x) = q_i + L_i(x - c_i), \quad \Phi'_i(x) = q_i + L'_i(x - c_i),$$

with $\det L_i > 0$ and $\det L'_i > 0$, and let $t \mapsto L_i(t)$ be a path from L_i to L'_i in $GL_n(\mathbb{R})^+$, which is path connected. Put $\Lambda = \max_{i,t} \|L_i(t)\|$. Shrinking Φ and Φ' as above with ratio $\eta > 0$, that is, replacing Φ_i and Φ'_i by $\mu_{q_i, \eta} \circ \Phi_i$ and $\mu_{q_i, \eta} \circ \Phi'_i$, replaces L_i by ηL_i and L'_i by $\eta L'_i$ and leaves q_i fixed; taking η so small that $\eta \Lambda \max_i \text{diam } C_i < r$, the affine linear maps $x \mapsto q_i + \eta L_i(t)(x - c_i)$ have positive determinant and carry C_i into $B(q_i, r)$. They therefore form a path in $\mathcal{P}$ joining the shrunken Φ to the shrunken Φ' , and hence Φ to Φ' .

Conclusion. The points $c_1, \dots, c_m$ are distinct and lie in $\text{int } I^n$, and so do $q_i = A_i(c_i)$, the sets $A_i(C_i)$ being disjoint. By Lemma 9.1 there is a path of m -tuples of distinct points of $\text{int } I^n$ from (c_i) to (q_i) ; let δ be the associated constant and fix $r < \delta/3$ with the balls $B(q_i, r)$ contained in $\text{int } I^n$ and pairwise disjoint. Shrink $(\text{id}, \dots, \text{id})$ until its images lie in $B(c_i, r)$, translate along the path of m -tuples to a tuple carrying c_i to q_i with images in $B(q_i, r)$, and shrink $(A_1, \dots, A_m)$ until its images lie in $B(q_i, r)$ as well. The two resulting tuples carry c_i to q_i and have their images in the balls $B(q_i, r)$, so they are joined by the turning move. Concatenating the paths joins $(\text{id}, \dots, \text{id})$ to $(A_1, \dots, A_m)$ in $\mathcal{P}$, and the components of that path are the required $A_i(t)$. $\square$

Remark 9.3. The hypothesis that the C_i have non-empty interiors is what puts the points c_i and q_i in $\text{int } I^n$, where Lemma 9.1 operates; and $n \geq 2$ is needed for that lemma, a path in an interval being unable to pass a point. Both hypotheses hold in the applications, where the C_i are the pieces of a subdivision.

9.2. The construction. Let $U \subseteq \mathbb{R}^n$ be compact and convex with non-empty interior, and let

$$g: (U, \partial U) \longrightarrow (X, x_0)$$

be a map of pairs, that is, a map $g: U \rightarrow X$ carrying ∂U to x_0 .

Definition 9.4. An affine linear map $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is *admissible for U* if $\det A > 0$ and $A(U) \subseteq I^n$.

Such an A has an affine inverse, and admissible maps exist: U is compact, so it lies in some cube, and a homothety followed by a translation carries that cube onto I^n with positive determinant. In what follows A denotes, more generally, any affine linear map with $\det A \neq 0$ and $A(U) \subseteq I^n$; only in Lemma 9.7(iii) does the sign of the determinant play a part. Note that $A(\text{int } U)$ is a non-empty open subset of $\mathbb{R}^n$ contained in I^n , hence contained in $\text{int } I^n$; so

$$(1) \quad A(U) \cap \partial I^n \subseteq A(\partial U),$$

and $A(U)$ is again a compact convex body, with $\partial A(U) = A(\partial U)$.

Definition 9.5. For g and A as above let $\alpha(g, A): I^n \rightarrow X$ be the map

$$\alpha(g, A)(x) = \begin{cases} g(A^{-1}x), & x \in A(U), \\ x_0, & x \in I^n \setminus A(U). \end{cases}$$

Lemma 9.6. $\alpha(g, A)$ is a map of pairs $(I^n, \partial I^n) \rightarrow (X, x_0)$.

Proof. The sets $A(U)$ and the closure of $I^n \setminus A(U)$ are closed and cover I^n , and they meet in a subset of $\partial A(U) \cap I^n = A(\partial U) \cap I^n$, where the two prescriptions agree: g carries ∂U to x_0 . So $\alpha(g, A)$ is continuous. It carries ∂I^n to x_0 : a point of ∂I^n outside $A(U)$ goes to x_0 by definition, and one in $A(U)$ lies in $A(\partial U)$ by (1), so goes to x_0 as well. $\square$

9.3. Independence of the choices.

Lemma 9.7. *Let U and g be as in §9.2.*

- (i) *If g_0 and g_1 are homotopic through maps of pairs $(U, \partial U) \rightarrow (X, x_0)$ and A is admissible for U , then $\alpha(g_0, A)$ and $\alpha(g_1, A)$ are homotopic through maps of pairs.*
- (ii) *If A_0 and A_1 are both admissible for U , then $\alpha(g, A_0)$ and $\alpha(g, A_1)$ are homotopic through maps of pairs.*
- (iii) *If A is affine linear with $\det A < 0$ and $A(U) \subseteq I^n$, then $[\alpha(g, A)]$ is the inverse in $\pi_n(X, x_0)$ of $[\alpha(g, A_0)]$ for any A_0 admissible for U .*

Consequently, for A admissible, the class $[\alpha(g, A)] \in \pi_n(X, x_0)$ depends only on the homotopy class of g rel ∂U .

Proof. (i) Let $G: U \times I \rightarrow X$ be a homotopy of maps of pairs from g_0 to g_1 , so $G(u, t) = x_0$ for $u \in \partial U$. Put

$$H(x, t) = \begin{cases} G(A^{-1}x, t), & x \in A(U), \\ x_0, & \text{otherwise,} \end{cases}$$

which is continuous by the argument of Lemma 9.6 applied to $I^n \times I$, and is a homotopy through maps of pairs from $\alpha(g_0, A)$ to $\alpha(g_1, A)$.

(ii) Suppose first that A_0 and A_1 are joined by a path $t \mapsto A_t$ of maps admissible for U , the coefficients varying continuously with t . Put

$$H(x, t) = \begin{cases} g(A_t^{-1}x), & A_t^{-1}x \in U, \\ x_0, & \text{otherwise.} \end{cases}$$

The map $(x, t) \mapsto A_t^{-1}x$ is continuous, inversion being continuous on invertible affine linear maps, so the two sets

$$\{(x, t) : A_t^{-1}x \in U\}, \quad \{(x, t) : A_t^{-1}x \notin \text{int } U\}$$

are closed in $I^n \times I$ and cover it, and on their intersection $A_t^{-1}x \in \partial U$, where g takes the value x_0 ; so H is continuous. Each $H(\cdot, t)$ is $\alpha(g, A_t)$, a map of pairs by Lemma 9.6, so H is a homotopy through maps of pairs from $\alpha(g, A_0)$ to $\alpha(g, A_1)$.

It remains to see that any two maps admissible for U are joined by such a path. For $n = 1$ the admissible maps form a convex set, and there is nothing to prove: $(1-t)A_0 + tA_1$ has positive slope, and it carries a point of U into I because I is convex.

Let $n \geq 2$ and let A_0, A_1 be admissible for U . Apply Proposition 9.2 to the single body $C_1 = A_0(U)$, which is a compact convex body contained in I^n , and to the affine linear map $A_1 \circ A_0^{-1}$, whose determinant is positive and which carries C_1 onto $A_1(U) \subseteq I^n$. It provides a path Θ_t of affine linear maps with positive determinant, from the identity to $A_1 \circ A_0^{-1}$, with $\Theta_t(A_0(U)) \subseteq I^n$ for every t . Then $A_t = \Theta_t \circ A_0$ is a path of maps admissible for U from A_0 to A_1 .

(iii) Let $\rho(s) = (1 - s_1, s_2, \dots, s_n)$ be the reflection of the cube in its first coordinate, an affine linear map with $\rho(I^n) = I^n$ and $\det \rho = -1$. Given A with $\det A < 0$ and

$A(U) \subseteq I^n$, put $A' = \rho \circ A$; then $\det A' > 0$ and $A'(U) = \rho(A(U)) \subseteq I^n$, so A' is admissible for U . Since $(A')^{-1} \circ \rho = A^{-1}$ and ρ carries $A(U)$ onto $A'(U)$,

$$\alpha(g, A') \circ \rho = \alpha(g, A),$$

both sides being $g \circ A^{-1}$ on $A(U)$ and x_0 elsewhere. Now $f \circ \rho$ is the map $\bar{f}$ obtained from f by reversing the first coordinate, and $[\bar{f}]$ is the inverse of $[f]$ by Theorem 4.4. So $[\alpha(g, A)]$ is the inverse of $[\alpha(g, A')]$, and by (ii) that class is $[\alpha(g, A_0)]$ for every A_0 admissible for U .

The final statement combines (i) and (ii). $\square$

Definition 9.8. Let $U \subseteq \mathbb{R}^n$ be compact and convex with non-empty interior and let $g: (U, \partial U) \rightarrow (X, x_0)$. The class of g read off using U is

$$\kappa(g) = [\alpha(g, A)] \in \pi_n(X, x_0)$$

for any A admissible for U ; by Lemma 9.7 this is well defined and depends only on the homotopy class of g rel ∂U . The notation does not record U , which is the domain of g .

Remark 9.9. Taking $U = I^n$ and $A = \text{id}$ gives $\alpha(g, \text{id}) = g$, so $\kappa(g) = [g]$: for the cube itself the new reading is the old class. The construction depends on the orientation of $\mathbb{R}^n$ through the requirement $\det A > 0$, and on nothing else: by part (iii) of Lemma 9.7, an affine linear map of negative determinant reads off the inverse class, written $-\kappa(g)$ in the additive notation used for $n \geq 2$.

Lemma 9.10. Let T be an affine linear map of $\mathbb{R}^n$ with $\det T > 0$ and let $g: (U, \partial U) \rightarrow (X, x_0)$. Then $T(U)$ is again a compact convex body, $g \circ T^{-1}$ carries $\partial T(U)$ to x_0 , and

$$\kappa(g \circ T^{-1}) = \kappa(g).$$

Proof. An affine linear map A is admissible for $T(U)$ if and only if $A \circ T$ is admissible for U , the determinants having the same sign and $A(T(U)) = (A \circ T)(U)$. For such an A the two maps $\alpha(g \circ T^{-1}, A)$ and $\alpha(g, A \circ T)$ are equal, both being $g \circ T^{-1} \circ A^{-1}$ on $A(T(U))$ and x_0 elsewhere. Now take classes. $\square$

Corollary 9.11. Let $U' \subseteq U$ be compact and convex with non-empty interior, and let $g: (U, \partial U) \rightarrow (X, x_0)$ carry $U \setminus \text{int } U'$ to x_0 . Then $g|_{U'}$ is a map of pairs $(U', \partial U') \rightarrow (X, x_0)$ and

$$\kappa(g) = \kappa(g|_{U'}).$$

Proof. Since $\partial U' \subseteq U \setminus \text{int } U'$, the restriction $g|_{U'}$ carries $\partial U'$ to x_0 . Let A be admissible for U ; then A is admissible for U' as well, since $A(U') \subseteq A(U) \subseteq I^n$. The two maps $\alpha(g, A)$ and $\alpha(g|_{U'}, A)$ are equal: both are $g \circ A^{-1}$ on $A(U')$ and x_0 elsewhere, the first because g is x_0 on $U \setminus \text{int } U'$ and hence on $U \setminus U'$. So the two classes agree. $\square$

Corollary 9.12. Let $U \subseteq \mathbb{R}^n$ be compact and convex with non-empty interior. Then every element of $\pi_n(X, x_0)$ is $\kappa(g)$ for some map of pairs $g: (U, \partial U) \rightarrow (X, x_0)$.

Proof. Let $F: (I^n, \partial I^n) \rightarrow (X, x_0)$ represent the given element; the identity is admissible for I^n and $\alpha(F, \text{id}) = F$, so the element is $\kappa(F)$. Choose an affine linear T with $\det T > 0$ and $T(U) \subseteq I^n$, and then an affine linear A_1 admissible for I^n with $A_1(I^n) \subseteq \text{int } T(U)$; such an A_1 exists, a homothety about an interior point of $T(U)$ followed by a translation

carrying I^n into it. By Definition 9.8 the class $\kappa(F)$ is that of $\alpha(F, A_1)$, which is x_0 outside $A_1(I^n)$ and hence outside $\text{int } T(U)$. So Corollary 9.11 gives

$$\kappa(F) = \kappa(\alpha(F, A_1)|_{T(U)}),$$

and Lemma 9.10, applied to T^{-1} , rewrites the right-hand side as $\kappa(g)$, where $g = \alpha(F, A_1)|_{T(U)} \circ T$. $\square$

9.4. Additivity.

Theorem 9.13. *Let $n \geq 2$, let $K \subseteq \mathbb{R}^n$ be compact and convex with non-empty interior, let $E_1, \dots, E_r$ be pairwise disjoint compact convex bodies with non-empty interiors contained in K , and let $F: (K, \partial K) \rightarrow (X, x_0)$ take the value x_0 at every point outside $\bigcup_k \text{int } E_k$. Then each $F|_{E_k}$ is a map of pairs $(E_k, \partial E_k) \rightarrow (X, x_0)$ and*

$$\kappa(F) = \sum_{k=1}^r \kappa(F|_{E_k}).$$

Proof. That $F|_{E_k}$ is a map of pairs is the hypothesis, ∂E_k lying outside all the interiors; so too is F itself, since $\text{int } E_k \subseteq \text{int } K$ puts ∂K outside them as well.

Reduction to the cube. Let A be admissible for K and put $F' = \alpha(F, A)$, so that $\kappa(F) = [F']$ by Definition 9.8. The sets $A(E_k)$ are pairwise disjoint compact convex bodies with non-empty interiors contained in $A(K) \subseteq I^n$, and F' is x_0 outside the union of their interiors. Moreover $F'|_{A(E_k)} = F|_{E_k} \circ A^{-1}$, so $\kappa(F'|_{A(E_k)}) = \kappa(F|_{E_k})$ by Lemma 9.10. Both sides of the asserted formula are therefore unchanged when K and F are replaced by I^n and F' , and by Remark 9.9 the left side is then the ordinary class $[F']$. So we may assume $K = I^n$.

Moving the bodies apart. Let $Q_k = [\frac{k-1}{r}, \frac{k}{r}] \times I^{n-1}$. For each k choose an affine linear A_k with positive determinant and $A_k(E_k) \subseteq \text{int } Q_k$: a homothety of small enough ratio followed by a translation will do. The sets $A_k(E_k)$ are then pairwise disjoint and contained in I^n , so Proposition 9.2 provides paths $A_k(t)$ of affine linear maps with positive determinant from the identity to A_k along which the sets $A_k(t)(E_k)$ stay pairwise disjoint in I^n . Put

$$F_t(x) = \begin{cases} F(A_k(t)^{-1}x), & x \in A_k(t)(E_k), \\ x_0, & x \notin \bigcup_k A_k(t)(E_k). \end{cases}$$

This is well defined, the sets being disjoint, and continuous in (x, t) by the argument of Lemma 9.7(ii): the sets $\{(x, t) : A_k(t)^{-1}x \in E_k\}$ and $\{(x, t) : A_k(t)^{-1}x \notin \text{int } E_k \text{ for all } k\}$ are closed and cover $I^n \times I$, and where they meet the value is x_0 . Each F_t carries ∂I^n to x_0 , since $A_k(t)(\text{int } E_k)$ is an open subset of I^n and so lies in $\text{int } I^n$. Finally $F_0 = F$. Hence $[F] = [F_1]$.

The slabs are the summands. Let

$$\theta_k: I^n \longrightarrow Q_k, \quad \theta_k(s) = \left(\frac{k-1+s_1}{r}, s_2, \dots, s_n \right),$$

the orientation-preserving affine homeomorphism onto the k -th slab, and put $G_k = F_1 \circ \theta_k$. Each G_k is a map of pairs, because F_1 is x_0 outside $A_k(E_k) \subseteq \text{int } Q_k$ and in particular on ∂Q_k . Let

$$P = (\cdots ((G_1 \cdot_1 G_2) \cdot_1 G_3) \cdots) \cdot_1 G_r$$

be the left-bracketed product of Definition 4.3. Unwinding that definition, P is given on the slab $J_k \times I^{n-1}$ by G_k reparametrized affinely in the first coordinate, where $J_1 = [0, 2^{-(r-1)}]$ and $J_k = [2^{-(r-k+1)}, 2^{-(r-k)}]$ for $2 \leq k \leq r$. Let $\varphi_1: I \rightarrow I$ be the increasing piecewise linear homeomorphism carrying $[\frac{k-1}{r}, \frac{k}{r}]$ onto J_k for every k , and put $\varphi(s) = (\varphi_1(s_1), s_2, \dots, s_n)$. Then $F_1 = P \circ \varphi$, both sides being G_k reparametrized affinely in the first coordinate on the k -th slab, and φ_1 fixes 0 and 1; so $[F_1] = [P]$ by the reparametrization lemma in the form used in the proof of Theorem 4.4. By Theorem 4.4 and Theorem 4.6,

$$[F_1] = \sum_{k=1}^r [G_k].$$

Identifying the terms. The map θ_k^{-1} is admissible for Q_k , its determinant being positive and $\theta_k^{-1}(Q_k) = I^n$, and $\alpha(F_1|_{Q_k}, \theta_k^{-1}) = F_1 \circ \theta_k = G_k$ on all of I^n . So $[G_k] = \kappa(F_1|_{Q_k})$. Since F_1 is x_0 on $Q_k \setminus \text{int } A_k(E_k)$, Corollary 9.11 gives

$$\kappa(F_1|_{Q_k}) = \kappa(F_1|_{A_k(E_k)}),$$

and $F_1|_{A_k(E_k)} = F|_{E_k} \circ A_k^{-1}$, so this is $\kappa(F|_{E_k})$ by Lemma 9.10. Adding up gives the formula. $\square$

Theorem 9.14 (additivity over a subdivision). *Let $n \geq 2$, let $K \subseteq \mathbb{R}^n$ be compact and convex with non-empty interior, and let $K = u_1 \cup \dots \cup u_r$ with the u_k compact convex sets with non-empty interiors and pairwise disjoint interiors. Let $g: (K, \partial K) \rightarrow (X, x_0)$ be constant with value x_0 on ∂u_k for every k . Then*

$$\kappa(g) = \sum_{k=1}^r \kappa(g|_{u_k}).$$

Proof. Pulling the pieces apart. Choose $b_k \in \text{int } u_k$. For $\frac{1}{2} \leq \lambda < 1$ the body $\mu_{b_k, \lambda}(u_k)$ is contained in $\text{int } u_k$: for $y \in u_k$ the half-open segment from b_k to y lies in $\text{int } u_k$, b_k being interior. Define $G_\lambda: K \rightarrow X$ by

$$G_\lambda(x) = \begin{cases} g(\mu_{b_k, \lambda}^{-1}x), & x \in \mu_{b_k, \lambda}(u_k), \\ x_0, & \text{otherwise.} \end{cases}$$

For $\lambda = 1$ this is g itself: the pieces then cover K , and on an overlap $u_k \cap u_j$ both prescriptions give $g(x)$, since $\mu_{b_k, 1}$ and $\mu_{b_j, 1}$ are the identity. As in Theorem 9.13 the sets $\{(x, \lambda) : \mu_{b_k, \lambda}^{-1}x \in u_k\}$ and its counterpart are closed and cover $K \times [\frac{1}{2}, 1]$, and the two prescriptions agree where they meet, so $(x, \lambda) \mapsto G_\lambda(x)$ is continuous on $K \times [\frac{1}{2}, 1]$. Each G_λ with $\lambda < 1$ carries ∂K to x_0 . Let $y \in \partial K$ and let j be any index. If $y \notin u_j$ then $y \notin \mu_{b_j, \lambda}(u_j) \subseteq u_j$; and if $y \in u_j$ then $y \in \partial u_j$, since $\text{int } u_j \subseteq \text{int } K$, so again $y \notin \mu_{b_j, \lambda}(u_j) \subseteq \text{int } u_j$. So y lies outside every $\mu_{b_j, \lambda}(u_j)$ and $G_\lambda(y) = x_0$; for $\lambda = 1$ we have $G_1 = g$, which carries ∂K to x_0 as well. Hence $\kappa(g) = \kappa(G_\lambda)$ for $\frac{1}{2} \leq \lambda \leq 1$ by Lemma 9.7(i).

Applying additivity. Fix $\lambda = \frac{1}{2}$ and put $D_k = \mu_{b_k, 1/2}(u_k)$: pairwise disjoint compact convex bodies with non-empty interiors, contained in $\text{int } u_k \subseteq \text{int } K$, outside whose interiors

$G_{1/2}$ takes the value x_0 . So Theorem 9.13, applied to K and these bodies, gives

$$\kappa(g) = \kappa(G_{1/2}) = \sum_{k=1}^r \kappa(G_{1/2}|_{D_k}).$$

Finally $G_{1/2}|_{D_k} = g|_{u_k} \circ \mu_{b_k,1/2}^{-1}$, an affine change of body with positive determinant, so each term is $\kappa(g|_{u_k})$ by Lemma 9.10. $\square$

Remark 9.15. A class was attached above to a map on a convex body by carrying that body into the cube by an affine linear map. One would like to allow any homeomorphism: a map defined on a subset of I^n homeomorphic to a ball should carry a class as soon as the identification is orientation preserving. Both of the things needed for that come from the Hurewicz theorem for spheres, applied to $X = S^n$, which is $(n-1)$ -connected; it is proved in *The Hurewicz theorems*.

First, a notion of orientation. Once $h: \pi_n(S^n) \rightarrow H_n(S^n)$ is an isomorphism, $\pi_n(S^n)$ is infinite cyclic generated by the identity, and a self-map of S^n is determined up to homotopy by its degree. One then *defines* an identification to be orientation preserving when the self-map of S^n it induces has degree 1.

Second, well-definedness. Two orientation-preserving identifications of a ball $B \subseteq I^n$ with I^n need not *visibly* be joined by a path of homeomorphisms;¹ but the associated maps of pairs $(I^n, \partial I^n) \rightarrow (B, \partial B)$ are homotopic through maps of pairs, by the same degree classification, and that is all the argument uses: if ϕ_t is such a homotopy then $g \circ \phi_t$ is a homotopy of maps of pairs into (X, x_0) , so $[g \circ \phi_0] = [g \circ \phi_1]$ and the class does not depend on the identification chosen. With this, Theorem 9.13 holds for finitely many disjoint subsets of I^n homeomorphic to a ball, the identifications being required only to be orientation preserving.

None of this may be used in the applications made here, since those are what the Hurewicz theorem is proved from. That is why we restricted to affine linear maps here, to have a path connected family robust enough to be able to prove the additivity formula, Theorem 9.14.

¹They are, in fact: any two orientation-preserving homeomorphisms of S^n are isotopic, so that $\text{Homeo}^+(S^n)$ is path connected. This is the stable homeomorphism theorem, equivalent to the annulus conjecture, proved for $n \geq 5$ by R. Kirby, *Stable homeomorphisms and the annulus conjecture*, Ann. of Math. **89** (1969), 575–582; the cases $n \leq 3$ are classical, and the case $n = 4$ follows from the annulus theorem of F. Quinn, *Ends of maps III: dimensions 4 and 5*, J. Differential Geom. **17** (1982), 503–521, which rests on Freedman’s theory of topological 4-manifolds. We make no use of any of this.