This is an expanded version of the appendix to my book *Algebraic Topology*. The plan is to include this expanded appendix in the second edition of the book.

Allen Hatcher

Appendix

Here are the section headings for the Appendix:

1. Topology of Cell Complexes
2. The Compact-Open Topology
3. The Homotopy Extension Property
4. Simplicial CW Structures
5. Abelian Groups

The last three sections have been added since the book was first published.

1. Topology of Cell Complexes

Here we collect a number of basic topological facts about CW complexes for convenient reference. A few related facts about manifolds are also proved.

Let us first recall from Chapter 0 that a CW complex is a space X constructed in the following way:

(1) Start with a discrete set X^0, the 0-cells of X.

(2) Inductively, form the n-skeleton X^n from X^{n-1} by attaching n-cells e^n_α via maps $\varphi_\alpha : S^{n-1} \to X^{n-1}$. This means that X^n is the quotient space of $X^{n-1} \coprod_\alpha D^n_\alpha$ under the identifications $x \sim \varphi_\alpha(x)$ for $x \in \partial D^n_\alpha$. The cell e^n_α is the homeomorphic image of $D^n_\alpha - \partial D^n_\alpha$ under the quotient map.

(3) $X = \bigcup_n X^n$ with the weak topology: A set $A \subset X$ is open (or closed) if and only if $A \cap X^n$ is open (or closed) in X^n for each n.

Note that condition (3) is superfluous when X is finite-dimensional, so that $X = X^n$ for some n. For if A is open in $X = X^n$, the definition of the quotient topology on X^n implies that $A \cap X^{n-1}$ is open in X^{n-1}, and then by the same reasoning $A \cap X^{n-2}$ is open in X^{n-2}, and similarly for all the skeleta X^{n-i}.

Each cell e^n_α has its **characteristic map** Φ_α, which is by definition the composition $D^n_\alpha \hookrightarrow X^{n-1} \coprod_\alpha D^n_\alpha \to X^n \hookrightarrow X$. This is continuous since it is a composition of continuous maps, the inclusion $X^n \hookrightarrow X$ being continuous by (3). The restriction of Φ_α to the interior of D^n_α is a homeomorphism onto e^n_α.

An alternative way to describe the topology on X is to say that a set $A \subset X$ is open (or closed) if and only if $\Phi_\alpha^{-1}(A)$ is open (or closed) in D^n_α for each characteristic map Φ_α. In one direction this follows from continuity of the Φ_α's, and in the other direction, suppose $\Phi_\alpha^{-1}(A)$ is open in D^n_α for each Φ_α, and suppose by induction on n that $A \cap X^{n-1}$ is open in X^{n-1}. Then since $\Phi_\alpha^{-1}(A)$ is open in D^n_α for all α, $A \cap X^n$ is open in X^n by the definition of the quotient topology on X^n. Hence by (3), A is open in X.

A consequence of this characterization of the topology on X is that X is a quotient space of $\coprod_{n,\alpha} D^n_\alpha$.

A **subcomplex** of a CW complex X is a subspace $A \subset X$ which is a union of cells of X, such that the closure of each cell in A is contained in A. Thus for each cell in A, the image of its attaching map is contained in A, so A is itself a CW complex. Its CW complex topology is the same as the topology induced from X, as one sees by noting inductively that the two topologies agree on $A^n = A \cap X^n$. It is easy to see by induction over skeleta that a subcomplex is a closed subspace. Conversely, a subcomplex could be defined as a closed subspace which is a union of cells.

A finite CW complex, that is, one with only finitely many cells, is compact since attaching a single cell preserves compactness. A sort of converse to this is:

Proposition A.1. *A compact subspace of a CW complex is contained in a finite subcomplex.*

Proof: First we show that a compact set C in a CW complex X can meet only finitely many cells of X. Suppose on the contrary that there is an infinite sequence of points $x_i \in C$ all lying in distinct cells. Then the set $S = \{x_1, x_2, \cdots\}$ is closed in X. Namely, assuming $S \cap X^{n-1}$ is closed in X^{n-1} by induction on n, then for each cell e^n_α of X, $\varphi_\alpha^{-1}(S)$ is closed in ∂D^n_α, and $\Phi_\alpha^{-1}(S)$ consists of at most one more point in D^n_α, so $\Phi_\alpha^{-1}(S)$ is closed in D^n_α. Therefore $S \cap X^n$ is closed in X^n for each n, hence S is closed in X. The same argument shows that any subset of S is closed, so S has the discrete topology. But it is compact, being a closed subset of the compact set C. Therefore S must be finite, a contradiction.

Since C is contained in a finite union of cells, it suffices to show that a finite union of cells is contained in a finite subcomplex of X. A finite union of finite subcomplexes is again a finite subcomplex, so this reduces to showing that a single cell e^n_α is contained in a finite subcomplex. The image of the attaching map φ_α for e^n_α is compact, hence by induction on dimension this image is contained in a finite subcomplex $A \subset X^{n-1}$. So e_α^n is contained in the finite subcomplex $A \cup e_\alpha^n$. □

Now we can explain the mysterious letters 'CW', which refer to the following two properties satisfied by CW complexes:

(1) Closure-finiteness: The closure of each cell meets only finitely many other cells. This follows from the preceding proposition since the closure of a cell is compact, being the image of a characteristic map.

(2) Weak topology: A set is closed if and only if it meets the closure of each cell in a closed set. For if a set meets the closure of each cell in a closed set, it pulls back to a closed set under each characteristic map, hence is closed by an earlier remark.

In J.H.C.Whitehead's original definition of CW complexes these two properties played a more central role. The following proposition contains essentially this definition.

Proposition A.2. *Given a Hausdorff space X and a family of maps $\Phi_\alpha : D_\alpha^n \to X$, then these maps are the characteristic maps of a CW complex structure on X if and only if*:

(i) *Each Φ_α restricts to a homeomorphism from* $\operatorname{int} D_\alpha^n$ *onto its image, a cell $e_\alpha^n \subset X$, and these cells are all disjoint and their union is X.*

(ii) *For each cell e_α^n, $\Phi_\alpha(\partial D_\alpha^n)$ is contained in the union of a finite number of cells of dimension less than n.*

(iii) *A subset of X is closed if and only if it meets the closure of each cell of X in a closed set.*

Condition (iii) can be restated as saying that a set $C \subset X$ is closed if and only if $\Phi_\alpha^{-1}(C)$ is closed in D_α^n for all α, since a map from a compact space onto a Hausdorff space is a quotient map. In particular, if there are only finitely many cells then (iii) is automatic since in this case the projection $\coprod_\alpha D_\alpha^n \to X$ is a map from a compact space onto a Hausdorff space, hence is a quotient map.

For an example where all the conditions except the finiteness hypothesis in (ii) are satisfied, take X to be D^2 with its interior as a 2-cell and each point of ∂D^2 as a 0-cell. The identity map of D^2 serves as the Φ_α for the 2-cell. Condition (iii) is satisfied since it is a nontrivial condition only for the 2-cell.

Proof: We have already taken care of the 'only if' implication. For the converse, suppose inductively that X^{n-1}, the union of all cells of dimension less than n, is a

CW complex with the appropriate Φ_α's as characteristic maps. The induction can start with $X^{-1}=\varnothing$. Let $f\colon X^{n-1}\coprod_\alpha D^n_\alpha \to X^n$ be given by the inclusion on X^{n-1} and the maps Φ_α for all the n-cells of X. This is a continuous surjection, and if we can show it is a quotient map, then X^n will be obtained from X^{n-1} by attaching the n-cells e^n_α. Thus if $C\subset X^n$ is such that $f^{-1}(C)$ is closed, we need to show that $C\cap \overline{e}^m_\beta$ is closed for all cells e^m_β of X, the bar denoting closure.

There are three cases. If $m<n$ then $f^{-1}(C)$ closed implies $C\cap X^{n-1}$ closed, hence $C\cap\overline{e}^m_\beta$ is closed since $\overline{e}^m_\beta\subset X^{n-1}$. If $m=n$ then e^m_β is one of the cells e^n_α, so $f^{-1}(C)$ closed implies $f^{-1}(C)\cap D^n_\alpha$ is closed, hence compact, hence its image $C\cap\overline{e}^n_\alpha$ under f is compact and therefore closed. Finally there is the case $m>n$. Then $C\subset X^n$ implies $C\cap\overline{e}^m_\beta\subset\Phi_\beta(\partial D^m_\beta)$. The latter space is contained in a finite union of $\overline{e}^\ell_\gamma$'s with $\ell<m$. By induction on m, each $C\cap\overline{e}^\ell_\gamma$ is closed. Hence the intersection of C with the union of the finite collection of $\overline{e}^\ell_\gamma$'s is closed. Intersecting this closed set with $\overline{e}^m_\beta$, we conclude that $C\cap\overline{e}^m_\beta$ is closed.

It remains only to check that X has the weak topology with respect to the X^n's, that is, a set in X is closed if and only if it intersects each X^n in a closed set. The preceding argument with $C=X^n$ shows that X^n is closed, so a closed set intersects each X^n in a closed set. Conversely, if a set C intersects X^n in a closed set, then C intersects each $\overline{e}^n_\alpha$ in a closed set, so C is closed in X by (iii). $\square$

Next we describe a convenient way of constructing open neighborhoods $N_\varepsilon(A)$ of subsets A of a CW complex X, where ε is a function that assigns to each cell e^n_α of X a number $\varepsilon_\alpha\in(0,1]$. The construction is inductive over the skeleta X^n, so suppose we have already constructed $N^n_\varepsilon(A)$, an open neighborhood of $A\cap X^n$ in X^n, starting the process with $N^0_\varepsilon(A)=A\cap X^0$. Then we enlarge $N^n_\varepsilon(A)$ to $N^{n+1}_\varepsilon(A)$ by adding points in each cell e^{n+1}_α of two types which we describe via a characteristic map $\Phi_\alpha\colon D^{n+1}\to X$ for e^{n+1}_α:

(1) Points $x\in e^{n+1}_\alpha$ with $\Phi^{-1}_\alpha(x)$ in an open ε_α-neighborhood of $\Phi^{-1}_\alpha(A)-\partial D^{n+1}$ in $D^{n+1}-\partial D^{n+1}$.

(2) Points $x\in e^{n+1}_\alpha$ with $\Phi^{-1}_\alpha(x)$ in a product $(1-\varepsilon_\alpha,1)\times\Phi^{-1}_\alpha(N^n_\varepsilon(A))$ with respect to spherical coordinates (r,θ) in D^{n+1}, where $r\in[0,1]$ is the radial coordinate and θ lies in $\partial D^{n+1}=S^n$.

Having defined $N^n_\varepsilon(A)$ for all n, we let $N_\varepsilon(A)=\bigcup_n N^n_\varepsilon(A)$. This is an open set in X since it pulls back to an open set under each characteristic map.

Proposition A.3. *CW complexes are normal, and in particular, Hausdorff.*

Proof: Points are closed in a CW complex X since they pull back to closed sets under all characteristic maps Φ_α. For disjoint closed sets A and B in X, we show that $N_\varepsilon(A)$ and $N_\varepsilon(B)$ are disjoint for small enough ε_α's. In the inductive process for building these open sets, assume $N_\varepsilon^n(A)$ and $N_\varepsilon^n(B)$ have been chosen to be disjoint. For a characteristic map $\Phi_\alpha : D^{n+1} \to X$, observe that $\Phi_\alpha^{-1}(N_\varepsilon^n(A))$ and $\Phi_\alpha^{-1}(B)$ are a positive distance apart, since otherwise by compactness we would have a sequence in $\Phi_\alpha^{-1}(B)$ converging to a point of $\Phi_\alpha^{-1}(B)$ in ∂D^{n+1} of distance zero from $\Phi_\alpha^{-1}(N_\varepsilon^n(A))$, but this is impossible since $\Phi_\alpha^{-1}(N_\varepsilon^n(B))$ is a neighborhood of $\Phi_\alpha^{-1}(B) \cap \partial D^{n+1}$ in ∂D^{n+1} disjoint from $\Phi_\alpha^{-1}(N_\varepsilon^n(A))$. Similarly, $\Phi_\alpha^{-1}(N_\varepsilon^n(B))$ and $\Phi_\alpha^{-1}(A)$ are a positive distance apart. Also, $\Phi_\alpha^{-1}(A)$ and $\Phi_\alpha^{-1}(B)$ are a positive distance apart. So a small enough ε_α will make $\Phi_\alpha^{-1}(N_\varepsilon^{n+1}(A))$ disjoint from $\Phi_\alpha^{-1}(N_\varepsilon^{n+1}(B))$ in D^{n+1}. $\square$

Proposition A.4. *Each point in a CW complex has arbitrarily small contractible open neighborhoods, so CW complexes are locally contractible.*

Proof: Given a point x in a CW complex X and a neighborhood U of x in X, we can choose the ε_α's small enough so that $N_\varepsilon(x) \subset U$ by requiring that the closure of $N_\varepsilon^n(x)$ be contained in U for each n. It remains to see that $N_\varepsilon(x)$ is contractible. If $x \in X^m - X^{m-1}$ and $n > m$ we can construct a deformation retraction of $N_\varepsilon^n(x)$ onto $N_\varepsilon^{n-1}(x)$ by sliding outward along radial segments in each cell e_α^n, the images of radial segments in D^n under the characteristic map Φ_α for e_α^n. A deformation retraction of $N_\varepsilon(x)$ onto $N_\varepsilon^m(x)$ is then obtained by performing the deformation retraction of $N_\varepsilon^n(x)$ onto $N_\varepsilon^{n-1}(x)$ during the t-interval $[1/2^n, 1/2^{n-1}]$, points of $N_\varepsilon^n(x) - N_\varepsilon^{n-1}(x)$ being stationary outside this t-interval. Finally, $N_\varepsilon^m(x)$ is an open ball about x, and so deformation retracts onto x. $\square$

In particular, CW complexes are locally path-connected. So a CW complex is path-connected if and only if it is connected.

Proposition A.5. *For a subcomplex A of a CW complex X, the open neighborhood $N_\varepsilon(A)$ deformation retracts onto A.*

Proof: In each cell e_α^n of $X - A$, $N_\varepsilon^n(A)$ is a union of radial segments going out to points of $N_\varepsilon^{n-1}(A)$, so a deformation retraction of $N_\varepsilon(A)$ onto A can be constructed just as in the previous proof. $\square$

Note that for subcomplexes A and B of X, we have $N_\varepsilon(A) \cap N_\varepsilon(B) = N_\varepsilon(A \cap B)$. This implies for example that the van Kampen theorem and Mayer-Vietoris sequences hold for decompositions $X = A \cup B$ into subcomplexes A and B as well as into open sets A and B.

A map $f : X \to Y$ with domain a CW complex is continuous if and only if its restrictions to the closures $\overline{e}^n_\alpha$ of all cells e^n_α are continuous, and it is useful to know that the same is true for homotopies $f_t : X \to Y$. With this objective in mind, let us introduce a little terminology. A topological space X is said to be **generated** by a collection of subspaces X_α if $X = \bigcup_\alpha X_\alpha$ and a set $A \subset X$ is closed if and only if $A \cap X_\alpha$ is closed in X_α for each α. Equivalently, we could say 'open' instead of 'closed' here, but 'closed' is more convenient for our present purposes. As noted earlier, though not in these words, a CW complex X is generated by the closures $\overline{e}^n_\alpha$ of its cells e^n_α. Since every finite subcomplex of X is a finite union of closures $\overline{e}^n_\alpha$, X is also generated by its finite subcomplexes. It follows that X is also generated by its compact subspaces, or more briefly, X is **compactly generated**.

Proposition A.15 later in the Appendix asserts that if X is a compactly generated Hausdorff space and Z is locally compact, then $X \times Z$, with the product topology, is compactly generated. In particular, $X \times I$ is compactly generated if X is a CW complex. Since every compact set in $X \times I$ is contained in the product of a compact subspace of X with I, hence in the product of a finite subcomplex of X with I, such product subspaces also generate $X \times I$. Since such a product subspace is a finite union of products $\overline{e}^n_\alpha \times I$, it is also true that $X \times I$ is generated by its subspaces $\overline{e}^n_\alpha \times I$. This implies that a homotopy $F : X \times I \to Y$ is continuous if and only if its restrictions to the subspaces $\overline{e}^n_\alpha \times I$ are continuous, which is the statement we were seeking.

Products of CW Complexes

There are some unexpected point-set-topological subtleties that arise with products of CW complexes. As we shall show, the product of two CW complexes does have a natural CW structure, but its topology is in general finer, with more open sets, than the product topology. However, the distinctions between the two topologies are rather small, and indeed nonexistent in most cases of interest, so there is no real problem for algebraic topology.

Given a space X and a collection of subspaces X_α whose union is X, these subspaces generate a possibly finer topology on X by defining a set $A \subset X$ to be open if and only if $A \cap X_\alpha$ is open in X_α for all α. The axioms for a topology are easily verified for this definition. In case $\{X_\alpha\}$ is the collection of compact subsets of X, we write X_c for this new compactly generated topology. It is easy to see that X and X_c have the same compact subsets, and the two induced topologies on these compact subsets coincide. If X is compact, or even locally compact, then $X = X_c$, that is, X is compactly generated.

Theorem A.6. *For CW complexes X and Y with characteristic maps Φ_α and Ψ_β, the product maps $\Phi_\alpha \times \Psi_\beta$ are the characteristic maps for a CW complex structure on $(X \times Y)_c$. If either X or Y is compact or more generally locally compact, then $(X \times Y)_c = X \times Y$. Also, $(X \times Y)_c = X \times Y$ if both X and Y have countably many cells.*

Proof: For the first statement it suffices to check that the three conditions in Proposition A.2 are satisfied when we take the space 'X' there to be $(X \times Y)_c$. The first two conditions are obvious. For the third, which says that $(X \times Y)_c$ is generated by the products $\overline{e}^m_\alpha \times \overline{e}^n_\beta$, observe that every compact set in $X \times Y$ is contained in the product of its projections onto X and Y, and these projections are compact and hence contained in finite subcomplexes of X and Y, so the original compact set is contained in a finite union of products $\overline{e}^m_\alpha \times \overline{e}^n_\beta$. Hence the products $\overline{e}^m_\alpha \times \overline{e}^n_\beta$ generate $(X \times Y)_c$.

The second assertion of the theorem is a special case of Proposition A.15, having nothing to do with CW complexes, which says that a product $X \times Y$ is compactly generated if X is compactly generated Hausdorff and Y is locally compact.

For the last statement of the theorem, suppose X and Y each have at most countably many cells. For an open set $W \subset (X \times Y)_c$ and a point $(a, b) \in W$ we need to find a product $U \times V \subset W$ with U an open neighborhood of a in X and V an open neighborhood of b in Y. Choose finite subcomplexes $X_1 \subset X_2 \subset \cdots$ of X with $X = \bigcup_i X_i$, and similarly for Y. We may assume $a \in X_1$ and $b \in Y_1$. Since the two topologies agree on $X_1 \times Y_1$, there is a compact product neighborhood $K_1 \times L_1 \subset W$ of (a, b) in $X_1 \times Y_1$. Assuming inductively that $K_i \times L_i \subset W$ has been constructed in $X_i \times Y_i$, we would like to construct $K_{i+1} \times L_{i+1} \subset W$ as a compact neighborhood of $K_i \times L_i$ in $X_{i+1} \times Y_{i+1}$. To do this, we first choose for each $x \in K_i$ compact neighborhoods K_x of x in X_{i+1} and L_x of L_i in Y_{i+1} such that $K_x \times L_x \subset W$, using the compactness of L_i. By compactness of K_i, a finite number of the K_x's cover K_i. Let K_{i+1} be the union of these K_x's and let L_{i+1} be the intersection of the corresponding L_x's. This defines the desired $K_{i+1} \times L_{i+1}$. Let U_i be the interior of K_i in X_i, so $U_i \subset U_{i+1}$ for each i. The union $U = \bigcup_i U_i$ is then open in X since it intersects each X_i in a union of open sets and the X_i's generate X. In the same way the L_i's yield an open set V in Y. Thus we have a product of open sets $U \times V \subset W$ containing (a, b). $\square$

We will describe now an example from [Dowker 1952] where the product topology on $X \times Y$ differs from the CW topology. Both X and Y will be graphs consisting of infinitely many edges emanating from a single vertex, with uncountably many edges for X and countably many for Y.

Let $X = \bigvee_s I_s$ where I_s is a copy of the interval $[0,1]$ and the index s ranges over all infinite sequences $s = (s_1, s_2, \cdots)$ of positive integers. The wedge sum is formed at the 0 endpoint of I_s. Similarly we let $Y = \bigvee_j I_j$ but with j varying just over positive integers. Let p_{sj} be the point $(1/s_j, 1/s_j) \in I_s \times I_j \subset X \times Y$ and let P be the union of all these points p_{sj}. Thus P consists of a single point in each 2-cell of $X \times Y$, so P is closed in the CW topology on $X \times Y$. We will show it is not closed in the product topology by showing that (x_0, y_0) lies in its closure, where x_0 is the common endpoint of the intervals I_s and y_0 is the common endpoint of the intervals I_j.

A basic open set containing (x_0, y_0) in the product topology has the form $U \times V$ where $U = \bigvee_s [0, a_s)$ and $V = \bigvee_j [0, b_j)$. It suffices to show that P has nonempty intersection with $U \times V$. Choose a sequence $t = (t_1, t_2, \cdots)$ with $t_j > j$ and $t_j > 1/b_j$ for all j, and choose an integer $k > 1/a_t$. Then $t_k > k > 1/a_t$ hence $1/t_k < a_t$. We also have $1/t_k < b_k$. So $(1/t_k, 1/t_k)$ is a point of P that lies in $[0, a_t) \times [0, b_k)$ and hence in $U \times V$.

Euclidean Neighborhood Retracts

At certain places in this book it is desirable to know that a given compact space is a retract of a finite simplicial complex, or equivalently (as we shall see) a retract of a neighborhood in some Euclidean space. For example, this condition occurs in the Lefschetz fixed point theorem, and it was used in the proof of Alexander duality. So let us study this situation in more detail.

Theorem A.7. *A compact subspace K of $\mathbb{R}^n$ is a retract of some neighborhood if and only if K is locally contractible in the weak sense that for each $x \in K$ and each neighborhood U of x in K there exists a neighborhood $V \subset U$ of x such that the inclusion $V \hookrightarrow U$ is nullhomotopic.*

Note that if K is a retract of some neighborhood, then it is a retract of every smaller neighborhood, just by restriction of the retraction. So it does not matter if we require the neighborhoods to be open. Similarly it does not matter if the neighborhoods U and V in the statement of the theorem are required to be open.

Proof: Let us do the harder half first, constructing a retraction of a neighborhood of K onto K under the local contractibility assumption. The first step is to put a

CW structure on the open set $X = \mathbb{R}^n - K$, with the size of the cells approaching zero near K. Consider the subdivision of $\mathbb{R}^n$ into unit cubes of dimension n with vertices at the points with integer coordinates. Call this collection of cubes C_0. For an integer $k > 0$, we can subdivide the cubes of C_0 by taking n-dimensional cubes of edgelength $1/2^k$ with vertices having coordinates of the form $i/2^k$ for $i \in \mathbb{Z}$. Denote this collection of cubes by C_k. Let $A_0 \subset C_0$ be the set of cubes disjoint from K, and inductively, let $A_k \subset C_k$ be the set of cubes disjoint from K and not contained in cubes of A_j for $j < k$. The open set X is then the union of all the cubes in the combined collection $A = \bigcup_k A_k$. Note that the collection A is locally finite: Each point of X has a neighborhood meeting only finitely many cubes in A, since the point has a positive distance from the closed set K.

If two cubes of A intersect, their intersection is an i-dimensional face of one of them for some $i < n$. Likewise, when two faces of cubes of A intersect, their intersection is a face of one of them. This implies that the open faces of cubes of A that are minimal with respect to inclusion among such faces form the cells of a CW structure on X, since the boundary of such a face is a union of such faces. The vertices of this CW structure are thus the vertices of all the cubes of A, and the n-cells are the interiors of the cubes of A.

Next we define inductively a subcomplex Z of this CW structure on X and a map $r : Z \to K$. The 0-cells of Z are exactly the 0-cells of X, and we let r send each 0-cell to the closest point of K, or if this is not unique, any one of the closest points of K. Assume inductively that Z^k and $r : Z^k \to K$ have been defined. For a cell e^{k+1} of X with boundary in Z^k, if the restriction of r to this boundary extends over e^{k+1} then we include e^{k+1} in Z^{k+1} and we let r on e^{k+1} be such an extension that is not too large, say an extension for which the diameter of its image $r(e^{k+1})$ is less than twice the infimum of the diameters for all possible extensions. This defines Z^{k+1} and $r : Z^{k+1} \to K$. At the end of the induction we set $Z = Z^n$.

It remains to verify that by letting r equal the identity on K we obtain a continuous retraction $Z \cup K \to K$, and that $Z \cup K$ contains a neighborhood of K. Given a point $x \in K$, let U be a ball in the metric space K centered at x. Since K is locally contractible, we can choose a finite sequence of balls in K centered at x, of the form $U = U_n \supset V_n \supset U_{n-1} \supset V_{n-1} \supset \cdots \supset U_0 \supset V_0$, each ball having radius equal to some small fraction of the radius of the preceding one, and with V_i contractible in U_i. Let $B \subset \mathbb{R}^n$ be a ball centered at x with radius less than half the radius of V_0, and let Y be the subcomplex of X formed by the cells whose closures are contained in B. Thus $Y \cup K$ contains a neighborhood of x in $\mathbb{R}^n$. By the choice of B and the definition of r on 0-cells we have $r(Y^0) \subset V_0$. Since V_0 is contractible in U_0, r is defined on the 1-cells of Y. Also, $r(Y^1) \subset V_1$ by the definition of r on 1-cells and the fact that U_0 is much smaller than V_1. Similarly, by induction we have r defined on Y^i with $r(Y^i) \subset V_i$ for all i. In particular, r maps Y to U. Since U could be arbitrarily small, this shows that extending r by the identity map on K gives a continuous map $r: Z \cup K \to K$. And since $Y \subset Z$, we see that $Z \cup K$ contains a neighborhood of K by the earlier observation that $Y \cup K$ contains a neighborhood of x. Thus $r: Z \cup K \to K$ retracts a neighborhood of K onto K.

Now for the converse. Since open sets in $\mathbb{R}^n$ are locally contractible, it suffices to show that a retract of a locally contractible space is locally contractible. Let $r: X \to A$ be a retraction and let $U \subset A$ be a neighborhood of a given point $x \in A$. If X is locally contractible, then inside the open set $r^{-1}(U)$ there is a neighborhood V of x that is contractible in $r^{-1}(U)$, say by a homotopy $f_t: V \to r^{-1}(U)$. Then $V \cap A$ is contractible in U via the restriction of the composition $r f_t$. $\square$

A space X is called a **Euclidean neighborhood retract** or **ENR** if for some n there exists an embedding $i: X \hookrightarrow \mathbb{R}^n$ such that $i(X)$ is a retract of some neighborhood in $\mathbb{R}^n$. The preceding theorem implies that the existence of the retraction is independent of the choice of embedding, at least when X is compact.

Corollary A.8. *A compact space is an ENR if and only if it can be embedded as a retract of a finite simplicial complex. Hence the homology groups and the fundamental group of a compact ENR are finitely generated.*

Proof: A finite simplicial complex K with n vertices is a subcomplex of a simplex Δ^{n-1}, and hence embeds in $\mathbb{R}^n$. The preceding theorem then implies that K is a retract of some neighborhood in $\mathbb{R}^n$, so any retract of K is also a retract of such a neighborhood, via the composition of the two retractions. Conversely, let K be a compact space that is a retract of some open neighborhood U in $\mathbb{R}^n$. Since K is compact it is bounded, lying in some large simplex $\Delta^n \subset \mathbb{R}^n$. Subdivide Δ^n, say by repeated barycentric subdivision, so that all simplices of the subdivision have diameter less than the distance from K to the complement of U. Then the union of all the simplices in this subdivision that intersect K is a finite simplicial complex that retracts onto K via the restriction of the retraction $U \to K$. $\square$

Corollary A.9. *Every compact manifold, with or without boundary, is an ENR.*

Proof: Manifolds are locally contractible, so it suffices to show that a compact manifold M can be embedded in $\mathbb{R}^k$ for some k. If M is not closed, it embeds in the closed manifold obtained from two copies of M by identifying their boundaries. So it suffices to consider the case that M is closed. By compactness there exist finitely many closed balls $B_i^n \subset M$ whose interiors cover M, where n is the dimension of M. Let $f_i : M \to S^n$ be the quotient map collapsing the complement of the interior of B_i^n to a point. These f_i's are the components of a map $f : M \to (S^n)^m$ which is injective since if x and y are distinct points of M with x in the interior of B_i^n, say, then $f_i(x) \neq f_i(y)$. Composing f with an embedding $(S^n)^m \hookrightarrow \mathbb{R}^k$, for example the product of the standard embeddings $S^n \hookrightarrow \mathbb{R}^{n+1}$, we obtain a continuous injection $M \hookrightarrow \mathbb{R}^k$, and this is a homeomorphism onto its image since M is compact. ◻

Corollary A.10. *Every finite CW complex is an ENR.*

Proof: Since CW complexes are locally contractible, it suffices to show that a finite CW complex can be embedded in some $\mathbb{R}^n$. This is proved by induction on the number of cells. Suppose the CW complex X is obtained from a subcomplex A by attaching a cell e^k via a map $f : S^{k-1} \to A$, and suppose that we have an embedding $A \hookrightarrow \mathbb{R}^m$. Then we can embed X in $\mathbb{R}^k \times \mathbb{R}^m \times \mathbb{R}$ as the union of $D^k \times \{0\} \times \{0\}$, $\{0\} \times A \times \{1\}$, and all line segments joining points $(x, 0, 0)$ and $(0, f(x), 1)$ for $x \in S^{k-1}$. ◻

Spaces Dominated by CW Complexes

We have been considering spaces which are retracts of finite simplicial complexes, and now we show that such spaces have the homotopy type of CW complexes. In fact, we can just as easily prove something a little more general than this. A space Y is said to be **dominated** by a space X if there are maps $Y \xrightarrow{i} X \xrightarrow{r} Y$ with $ri \simeq \mathbb{1}$. This makes the notion of a retract into something that depends only on the homotopy types of the spaces involved.

Proposition A.11. *A space dominated by a CW complex is homotopy equivalent to a CW complex.*

Proof: Recall from §3.F that the mapping telescope $T(f_1, f_2, \cdots)$ of a sequence of maps $X_1 \xrightarrow{f_1} X_2 \xrightarrow{f_2} X_3 \longrightarrow \cdots$ is the quotient space of $\coprod_i (X_i \times [i, i+1])$ obtained by identifying $(x, i+1) \in X_i \times [i, i+1]$ with $(f(x), i+1) \in X_{i+1} \times [i+1, i+2]$. We shall need the following elementary facts:

(1) $T(f_1, f_2, \cdots) \simeq T(g_1, g_2, \cdots)$ if $f_i \simeq g_i$ for each i.
(2) $T(f_1, f_2, \cdots) \simeq T(f_2, f_3, \cdots)$.
(3) $T(f_1, f_2, \cdots) \simeq T(f_2 f_1, f_4 f_3, \cdots)$.

The second of these is obvious. To prove the other two we will use Proposition 0.18, whose proof applies not just to CW pairs but to any pair (X_1, A) for which there is a deformation retraction of $X_1 \times I$ onto $X_1 \times \{0\} \cup A \times I$. To prove (1) we regard $T(f_1, f_2, \cdots)$ as being obtained from $\coprod_i (X_i \times \{i\})$ by attaching $\coprod_i (X_i \times [i, i+1])$. Then we can obtain $T(g_1, g_2, \cdots)$ by varying the attaching map by homotopy. To prove (3) we view $T(f_1, f_2, \cdots)$ as obtained from the disjoint union of the mapping cylinders $M(f_{2i})$ by attaching $\coprod_i (X_{2i-1} \times [2i-1, 2i])$. By sliding the attachment of $X_{2i-1} \times [2i-1, 2i]$ to $X_{2i} \subset M(f_{2i})$ down the latter mapping cylinder to X_{2i+1} we convert $M(f_{2i-1}) \cup M(f_{2i})$ into $M(f_{2i} f_{2i-1}) \cup M(f_{2i})$. This last space deformation retracts onto $M(f_{2i} f_{2i-1})$. Doing this for all i gives the homotopy equivalence in (3).

Now to prove the proposition, suppose that the space Y is dominated by the CW complex X via maps $Y \xrightarrow{i} X \xrightarrow{r} Y$ with $ri \simeq \mathbb{1}$. By (2) and (3) we have $T(ir, ir, \cdots) \simeq T(r, i, r, i, \cdots) \simeq T(i, r, i, r, \cdots) \simeq T(ri, ri, \cdots)$. Since $ri \simeq \mathbb{1}$, $T(ri, ri, \cdots)$ is homotopy equivalent to the telescope of the identity maps $Y \to Y \to Y \to \cdots$, which is $Y \times [0, \infty) \simeq Y$. On the other hand, the map ir is homotopic to a cellular map $f : X \to X$, so $T(ir, ir, \cdots) \simeq T(f, f, \cdots)$, which is a CW complex. $\square$

One might ask whether a space dominated by a finite CW complex is homotopy equivalent to a finite CW complex. In the simply-connected case this follows from Proposition 4C.1 since such a space has finitely generated homology groups. But there are counterexamples in the general case; see [Wall 1965].

In view of Corollary A.9 the preceding proposition implies:

Corollary A.12. *A compact manifold is homotopy equivalent to a CW complex.* $\square$

One could ask more refined questions. For example, do all compact manifolds have CW complex structures, or even stronger, do they have simplicial complex structures? Answers here are considerably harder to come by. Restricting attention to closed manifolds for simplicity, it is known that simplicial complex structures exist for all manifolds of dimension 1, 2, and 3, but in each higher dimension starting with 4 there exist manifolds that have no simplicial complex structure. In dimensions greater than 4, CW structures always exist, but it seems to be still unknown whether all manifolds of dimension 4 are CW complexes. For more on these questions, see [Kirby & Siebenmann 1977], [Freedman & Quinn 1990], and [Manolescu 2016].

Exercises

1. Show that a covering space of a CW complex is also a CW complex, with cells projecting homeomorphically onto cells.

2. Let X be a CW complex and x_0 any point of X. Construct a new CW complex structure on X having x_0 as a 0-cell, and having each of the original cells a union of the new cells. The latter condition is expressed by saying the new CW structure is a **subdivision** of the old one.

3. Show that a CW complex is path-connected if and only if its 1-skeleton is path-connected.

4. Show that a CW complex is locally compact if and only if each point has a neighborhood that meets only finitely many cells.

5. For a space X, show that the identity map $X_c \to X$ induces an isomorphism on π_1, where X_c denotes X with the compactly generated topology.

2. The Compact-Open Topology

By definition, the compact-open topology on the space X^Y of maps $f: Y \to X$ has a subbasis consisting of the sets $M(K, U)$ of mappings taking a compact set $K \subset Y$ to an open set $U \subset X$. Thus a basis for X^Y consists of sets of maps taking a finite number of compact sets $K_i \subset Y$ to open sets $U_i \subset X$. If Y is compact, which is the only case we consider in this book, convergence to $f \in X^Y$ means, loosely speaking, that finer and finer compact covers $\{K_i\}$ of Y are taken to smaller and smaller open covers $\{U_i\}$ of $f(Y)$. One of the main cases of interest in homotopy theory is when $Y = I$, so X^I is the space of paths in X. In this case one can check that a system of basic neighborhoods of a path $f: I \to X$ consists of the open sets $\bigcap_i M(K_i, U_i)$ where the K_i's are a partition of I into nonoverlapping closed intervals and U_i is an open neighborhood of $f(K_i)$.

In many cases of interest the compact-open topology is the same as the topology of uniform convergence:

Proposition A.13. *If X is a metric space and Y is compact, then the compact-open topology on X^Y is the same as the metric topology defined by the metric $d(f,g) = \sup_{y \in Y} d(f(y), g(y))$.*

Proof: First we show that every open ε-ball $B_\varepsilon(f)$ about $f \in X^Y$ contains a neighborhood of f in the compact-open topology. Since $f(Y)$ is compact, it is covered by finitely many balls $B_{\varepsilon/3}(f(y_i))$. Let $K_i \subset Y$ be the closure of $f^{-1}(B_{\varepsilon/3}(f(y_i)))$, so K_i is compact, $Y = \bigcup_i K_i$, and $f(K_i) \subset B_{\varepsilon/2}(f(y_i)) = U_i$, hence $f \in \bigcap_i M(K_i, U_i)$. To show that $\bigcap_i M(K_i, U_i) \subset B_\varepsilon(f)$, suppose that $g \in \bigcap_i M(K_i, U_i)$. For any $y \in Y$, say $y \in K_i$, we have $d(g(y), f(y_i)) < \varepsilon/2$ since $g(K_i) \subset U_i$. Likewise we have

$d(f(y), f(y_i)) < \varepsilon/2$, so $d(f(y), g(y)) \le d(f(y), f(y_i)) + d(g(y), f(y_i)) < \varepsilon$. Since y was arbitrary, this shows $g \in B_\varepsilon(f)$.

Conversely, we show that for each open set $M(K, U)$ and each $f \in M(K, U)$ there is a ball $B_\varepsilon(f) \subset M(K, U)$. Since $f(K)$ is compact, it has a distance $\varepsilon > 0$ from the complement of U. Then $d(f, g) < \varepsilon/2$ implies $g(K) \subset U$ since $g(K)$ is contained in an $\varepsilon/2$-neighborhood of $f(K)$. So $B_{\varepsilon/2}(f) \subset M(K, U)$. □

The next proposition contains some useful properties of the compact-open topology from the viewpoint of algebraic topology.

Proposition A.14. (a) *The evaluation map* $e : X^Y \times Y \to X$, $e(f, y) = f(y)$, *is continuous If* Y *is locally compact.*
(b) *If* $f : Y \times Z \to X$ *is continuous then so is the map* $\widehat{f} : Z \to X^Y$, $\widehat{f}(z)(y) = f(y, z)$.
(c) *The converse to* (b) *holds when* Y *is locally compact.*

Different definitions of local compactness are common, but the definition we are using is that Y is **locally compact** if for each point $y \in Y$ and each neighborhood U of y there is a compact neighborhood V of y contained in U.

In particular, parts (b) and (c) of the proposition provide the point-set topology justifying the adjoint relation $\langle \Sigma X, Y \rangle = \langle X, \Omega Y \rangle$ in §4.3, since they imply that a map $\Sigma X \to Y$ is continuous if and only if the associated map $X \to \Omega Y$ is continuous, and similarly for homotopies of such maps. Namely, think of a basepoint-preserving map $\Sigma X \to Y$ as a map $f : I \times X \to Y$ taking $\partial I \times X \cup \{x_0\} \times I$ to the basepoint of Y, so the associated map $\widehat{f} : X \to Y^I$ has image in the subspace $\Omega Y \subset Y^I$. A homotopy $f_t : \Sigma X \to Y$ gives a map $F : I \times X \times I \to Y$ taking $\partial I \times X \times I \cup I \times \{x_0\} \times I$ to the basepoint, with $\widehat{F}$ a map $X \times I \to \Omega Y \subset Y^I$ defining a basepoint-preserving homotopy $\widehat{f}_t$.

Proof: (a) For $(f, y) \in X^Y \times Y$ let $U \subset X$ be an open neighborhood of $f(y)$. Since Y is locally compact, continuity of f implies there is a compact neighborhood $K \subset Y$ of y such that $f(K) \subset U$. Then $M(K, U) \times K$ is a neighborhood of (f, y) in $X^Y \times Y$ taken to U by e, so e is continuous at (f, y).

(b) Suppose $f : Y \times Z \to X$ is continuous. To show continuity of $\widehat{f}$ it suffices to show that for a subbasic set $M(K, U) \subset X^Y$, the set $\widehat{f}^{-1}(M(K, U)) = \{ z \in Z \mid f(K, z) \subset U \}$ is open in Z. Let $z \in \widehat{f}^{-1}(M(K, U))$. Since $f^{-1}(U)$ is an open neighborhood of the compact set $K \times \{z\}$, there exist open sets $V \subset Y$ and $W \subset Z$ whose product $V \times W$ satisfies $K \times \{z\} \subset V \times W \subset f^{-1}(U)$. So W is a neighborhood of z in $\widehat{f}^{-1}(M(K, U))$.

(c) Note that $f : Y \times Z \to X$ is the composition $Y \times Z \to Y \times X^Y \to X$ of $\mathbb{1} \times \widehat{f}$ and the evaluation map, so part (a) gives the result. □

We will give three separate applications of Proposition A.14. Here is the first:

Proposition A.15. *If X is a compactly generated Hausdorff space and Y is locally compact, then the product topology on $X \times Y$ is compactly generated.*

Proof: First a preliminary observation: A function $f : X \times Y \to Z$ is continuous if and only if its restrictions $f : C \times Y \to Z$ are continuous for all compact $C \subset X$. For, using (b) and (c) of the preceding Proposition A.14, the first statement is equivalent to $\widehat{f} : X \to Z^Y$ being continuous and the second statement is equivalent to $\widehat{f} : C \to Z^Y$ being continuous for all compact $C \subset X$. Then since X is compactly generated, continuity of $\widehat{f} : X \to Z^Y$ is equivalent to continuity of $\widehat{f} : C \to Z^Y$ for all compact $C \subset X$.

To prove the proposition we just need to show the identity map $X \times Y \to (X \times Y)_c$ is continuous. By the previous paragraph, this is equivalent to continuity of the inclusion maps $C \times Y \to (X \times Y)_c$ for all compact $C \subset X$. Since Y is locally compact, it is compactly generated, and C is compact Hausdorff hence locally compact, so the same reasoning shows that continuity of $C \times Y \to (X \times Y)_c$ is equivalent to continuity of $C \times C' \to (X \times Y)_c$ for all compact $C' \subset Y$. But on the compact set $C \times C'$, the two topologies on $X \times Y$ agree, so we are done. (This proof is from [Dugundji 1966].) $\square$

Returning to the context of Proposition A.14, part (b) of that proposition implies that there is a well-defined function $X^{Y \times Z} \to (X^Y)^Z$ sending f to $\widehat{f}$. This is injective, and part (c) implies that it is surjective if Y is locally compact.

Proposition A.16. *The map $X^{Y \times Z} \to (X^Y)^Z$, $f \mapsto \widehat{f}$, is a homeomorphism if Y is locally compact Hausdorff and Z is Hausdorff.*

Proof: First we show that a subbasis for $X^{Y \times Z}$ is formed by the sets $M(A \times B, U)$ as A and B range over compact sets in Y and Z respectively and U ranges over open sets in X. Given a compact $K \subset Y \times Z$ and a map $f \in M(K, U)$, let K_Y and K_Z be the projections of K onto Y and Z. Then $K_Y \times K_Z$ is compact Hausdorff and hence normal. A normal space has the property that for each closed set C and each open set O containing C there is another open set O' containing C whose closure is contained in O. To see this, apply the normality property to the two closed sets C and the complement C' of O, taking O' to be the resulting open set containing C and disjoint from an open set containing C', so the closure of O' is contained in O. Applying this observation to the normal space $K_Y \times K_Z$ with C a point $k \in K$ and $O = (K_Y \times K_Z) \cap f^{-1}(U)$, the result is an open neighborhood of k in $K_Y \times K_Z$ whose closure is contained in $f^{-1}(U)$. We can take this open neighborhood to be a product

$V_k \times W_k \subset K_Y \times K_Z$, so its closure is a compact neighborhood $A_k \times B_k \subset f^{-1}(U)$ of k in $K_Y \times K_Z$. The sets $V_k \times W_k$ for varying $k \in K$ form an open cover of the compact set K so a finite number of the products $A_k \times B_k$ cover K. After discarding the others we then have $f \in \bigcap_k M(A_k \times B_k, U) \subset M(K, U)$, which shows that the sets $M(A \times B, U)$ form a subbasis for $X^{Y \times Z}$ as claimed.

Under the bijection $X^{Y \times Z} \to (X^Y)^Z$ the sets $M(A \times B, U)$ correspond to the sets $M(B, M(A, U))$, so it will suffice to show the latter sets form a subbasis for $(X^Y)^Z$. We will show more generally that for any space Q a subbasis for Q^Z is formed by the sets $M(K, V)$ as V ranges over a subbasis for Q and K ranges over compact sets in Z, assuming that Z is Hausdorff. Then we let $Q = X^Y$ with subbasis the sets $M(A, U)$.

Given $f \in M(K, U)$ with K compact in Z and U open in Q, write U as a union of basic sets U_α with each U_α an intersection of finitely many sets $V_{\alpha,j}$ of the given subbasis for Q. The cover of K by the open sets $f^{-1}(U_\alpha)$ has a finite subcover, say by the open sets $f^{-1}(U_i)$. Since K is compact Hausdorff, hence normal, we can write K as a union of compact subsets K_i with $K_i \subset f^{-1}(U_i)$, namely, each $k \in K$ has a compact neighborhood K_k contained in some $f^{-1}(U_i)$ with $k \in f^{-1}(U_i)$, so compactness of K implies that finitely many of these sets K_k cover K and we let K_i be the union of those contained in $f^{-1}(U_i)$. Now f lies in $M(K_i, U_i) = M(K_i, \bigcap_j V_{ij}) = \bigcap_j M(K_i, V_{ij})$ for each i. Hence f lies in $\bigcap_{i,j} M(K_i, V_{ij}) = \bigcap_i M(K_i, U_i) \subset M(K, U)$. Since $\bigcap_{i,j} M(K_i, V_{ij})$ is a finite intersection, this shows that the sets $M(K, V)$ form a subbasis for Q^Z. $\square$

Finally, we use Proposition A.14 to prove a very useful fact relating product spaces and quotient spaces:

Proposition A.17. *If $f : X \to Y$ is a quotient map then so is $f \times 1\!\!1 : X \times Z \to Y \times Z$ whenever Z is locally compact.*

This can be applied when $Z = I$ to show that a homotopy defined on a quotient space is continuous.

Proof: Consider the diagram at the right, where W is $Y \times Z$ with the quotient topology from $X \times Z$, with g the quotient map and h the identity. Every open set in $Y \times Z$ is open in W since $f \times 1\!\!1$ is continuous, so it will suffice to show that h is continuous.

$$\begin{tikzcd} X \times Z \arrow[rr, "f \times 1\!\!1"] \arrow[dr, "g"'] & & Y \times Z \arrow[dl, "h"] \\ & W & \end{tikzcd}$$

Since g is continuous, so is the associated map $\widehat{g} : X \to W^Z$, by Proposition A.14. This implies that $\widehat{h} : Y \to W^Z$ is continuous since f is a quotient map. Applying Proposition A.14 again, we conclude that h is continuous. $\square$

3. The Homotopy Extension Property

Near the end of Chapter 0 we stated and partially proved an equivalence between the homotopy extension property and a certain retraction property:

Proposition A.18. *A pair (X,A) has the homotopy extension property if and only if $X\times\{0\}\cup A\times I$ is a retract of $X\times I$.*

Proof: We already gave the easy argument showing that the homotopy extension property implies the retraction property. The converse was also easy when A is closed in X. In particular this covered all cases in which X is Hausdorff. What we will give now is an argument for the converse that applies in general. The argument is from [Strøm 1968].

To simplify the notation, let $Y = X\times\{0\}\cup A\times I$ with the subspace topology from $X\times I$. We will identify X with the subspace $X\times\{0\}$ of Y. Assuming there exists a retraction $r : X\times I \to Y$, we will show that a subset $O \subset Y$ is open in Y if its intersections with X and $A\times I$ are open in these two subspaces. This implies that a function on Y is continuous if its restrictions to X and $A\times I$ are continuous. Composing such a function with the retraction then provides the extension to $X\times I$ required for the homotopy extension property.

To show that O is open in Y it suffices to find, for each point $x \in O$, a product of open sets $V\times W \subset X\times I$ containing x such that $(V\times W)\cap Y \subset O$. If $x \in A\times(0,1]$ there is no problem doing this, so we may assume $x \in X$. In this case there is also no problem if x in not in the closure $\overline{A}$ of A, so we will assume $x \in \overline{A}$ from now on. For an integer $n \ge 1$ let U_n be the largest open set in X such that $(U_n\cap A)\times[0,1/n) \subset O$. The existence of a largest such set U_n follows from the fact that a union of open sets with this property is again an open set with this property. Let $U = \bigcup_n U_n$. Note that $A\cap O \subset U$ since O intersects $A\times I$ in an open set in $A\times I$ by assumption. It will suffice to show that $x \in U$ since if $x \in U_n$ then we can choose $V\times W = (U_n\cap O)\times[0,1/n)$ because $((U_n\cap O)\times[0,1/n))\cap Y \subset O$, where $U_n\cap O$ is open in X since U_n is open in X.

In order to show that $x \in U$ we first consider the point (x,t) for fixed $t > 0$. Writing $r(x,t) = \big(r_1(x,t), r_2(x,t)\big) \in X\times I$ we have $r_2(x,t) = t$ since $x \in \overline{A}$ and $r(a,t) = (a,t)$ for $a \in A$. Thus $r(x,t) \in A\times\{t\}$ so $r_1(x,t) \in A$. We claim next that if $r_1(x,t) \in U_n$ then $x \in U_n$. For if $r_1(x,t) \in U_n$ then continuity of r_1 implies that $r_1\big(V\times(t-\varepsilon, t+\varepsilon)\big) \subset U_n$ for some open neighborhood V of x in X and some $\varepsilon > 0$. In particular $r_1\big((V\cap A)\times\{t\}\big) \subset U_n$, or in other words, $V\cap A \subset U_n$. By the definition of U_n this implies that $V \subset U_n$ and hence $x \in U_n$ since $x \in V$. Thus we have shown that $r_1(x,t) \in U_n$ implies $x \in U_n$.

Suppose now that x is not in U. From what we have just shown, this implies that $r_1(x,t) \in A - U$. It follows that $r_1(x,t) \in A - O$ since $A \cap O \subset U$, as noted earlier. The relation $r_1(x,t) \in A - O$ holds for arbitrary $t > 0$, so by letting t approach 0 we conclude that $r_1(x,0) \in \overline{A} - O$ since r_1 is a continuous map to X and $X \cap O$ is open in X. Since $r_1(x,0) = x$ we deduce that x is not in O. However, this contradicts the fact that x was chosen to be a point in O. From this contradiction we conclude that x must be in U, and the proof is finished. $\square$

For an example of a set $O \subset X\times\{0\} \cup A\times I$ which is not open even though it intersects both $X\times\{0\}$ and $A\times I$ in open sets, let $X = [0,1]$ with $A = (0,1]$ and $O = \{\,(x,t) \in X\times I \mid t < x \text{ or } t = 0\,\}$. Note that in this case there exists no retraction $X\times I \to X\times\{0\}\cup A\times I$ since the image of a compact set must be compact. This example also illustrates how the topology on $X\times\{0\} \cup A\times I$ as a subspace of $X\times I$ can be different from the topology as the mapping cylinder of the inclusion $A \hookrightarrow X$, which is the quotient topology from $X \amalg A\times I$.

4. Simplicial CW Structures

A **Δ-complex** can be defined as a CW complex X in which each cell e^n_α is provided with a distinguished characteristic map $\sigma_\alpha : \Delta^n \to X$ such that the restriction of σ_α to each face Δ^{n-1} of Δ^n is the distinguished characteristic map σ_β for some $(n-1)$-cell e^{n-1}_β. It is understood that the simplices Δ^n and Δ^{n-1} have a specified ordering of their vertices, and the ordering of the vertices of Δ^n induces an ordering of the vertices of each face, which allows each face to be identified canonically with Δ^{n-1}. Intuitively, one thinks of the vertices of each n-cell of X as ordered by attaching the labels $0, 1, \cdots, n$ near the vertices, just inside the cell. The vertices themselves do not have to be distinct points of X.

If we no longer pay attention to orderings of vertices of simplices, we obtain a weaker structure which could be called an **unordered Δ-complex**. Here each cell e^n_α has a distinguished characteristic map $\sigma_\alpha : \Delta^n \to X$, but the restriction of σ_α to a face of Δ^n is allowed to be the composition of $\sigma_\beta : \Delta^{n-1} \to X$ with a symmetry of Δ^{n-1} permuting its vertices. Alternatively, we could say that each cell e^n_α has a family of $(n+1)!$ distinguished characteristic maps $\Delta^n \to X$ differing only by symmetries of Δ^n, such that the restrictions of these characteristic maps to faces give the distinguished characteristic maps for $(n-1)$-cells. The barycentric subdivision of any unordered

Δ-complex is an ordered Δ-complex since the vertices of the barycentric subdivision are the barycenters of the simplices of the original complex, hence have a canonical ordering according to the dimensions of these simplices. The simplest example of an unordered Δ-complex that cannot be made into an ordered Δ-complex without subdivision is Δ^2 with its three edges identified by a one-third rotation of Δ^2 permuting the three vertices cyclically.

In the literature unordered Δ-complex structures are sometimes called generalized triangulations. They can be useful in situations where orderings of vertices are not needed. One disadvantage of unordered Δ-complexes is that they do not behave as well with respect to products. The product of two ordered simplices has a canonical subdivision into ordered simplices using the shuffling operation described in §3.B, and this allows the product of two ordered Δ-complexes to be given a canonical ordered Δ-complex structure. Without orderings this no longer works.

A CW complex is called **regular** if its characteristic maps can be chosen to be embeddings. The closures of the cells are then homeomorphic to closed balls, and so it makes sense to speak of closed cells in a regular CW complex. The closed cells can be regarded as cones on their boundary spheres, and these cone structures can be used to subdivide a regular CW complex into a regular Δ-complex by induction over skeleta. In particular, regular CW complexes are homeomorphic to Δ-complexes. The barycentric subdivision of an unordered Δ-complex is a regular Δ-complex. A simplicial complex is a regular unordered Δ-complex in which each simplex is uniquely determined by its vertices. In the literature a regular unordered Δ-complex is sometimes called a simplicial multicomplex, or just a multicomplex, to convey the idea that there can be many simplices with the same set of vertices. The barycentric subdivision of a regular unordered Δ-complex is a simplicial complex. Hence barycentrically subdividing an unordered Δ-complex twice produces a simplicial complex.

A major disadvantage of Δ-complexes is that they do not allow quotient constructions. The quotient X/A of a Δ-complex X by a subcomplex A is not usually a Δ-complex. More generally, attaching a Δ-complex X to a Δ-complex Y via a simplicial map from a subcomplex $A \subset X$ to Y is not usually a Δ-complex. Here a simplicial map $f: A \to Y$ is one that sends each cell e^n_α of A onto a cell e^k_β of Y so that the square at the right commutes, with q a linear surjection sending vertices to vertices, preserving order. To fix this problem we need to broaden the definition of a Δ-complex to allow cells to be attached by arbitrary simplicial maps. Thus we define a **singular Δ-complex**, or $s\Delta$-complex, to be a CW complex with distinguished characteristic maps $\sigma_\alpha : \Delta^n \to X$ whose re-

$$\begin{tikzcd} \Delta^n \arrow[r, "\sigma_\alpha"] \arrow[d, "q"'] & A \arrow[d, "f"] \\ \Delta^k \arrow[r, "\sigma_\beta"] & Y \end{tikzcd}$$

strictions to faces are compositions $\sigma_\beta q : \Delta^{n-1} \to \Delta^k \to X$ for q a linear surjection taking vertices to vertices, preserving order. Simplicial maps between $s\Delta$-complexes are defined just as for Δ-complexes. With $s\Delta$-complexes one can perform attaching constructions in the same way as for CW complexes, using simplicial maps instead of cellular maps to specify the attachments. In particular one can form quotients, mapping cylinders, and mapping cones. One can also take products by the same subdivision procedure as for Δ-complexes.

We can view any $s\Delta$-complex X as being constructed inductively, skeleton by skeleton, where the skeleton X^n is obtained from X^{n-1} by attaching simplices Δ^n via simplicial maps $\partial\Delta^n \to X^{n-1}$ that preserve the ordering of vertices in each face of Δ^n. Conversely, any CW complex built in this way is an $s\Delta$-complex. For example, the usual CW structure on S^n consisting of one 0-cell and one n-cell is an $s\Delta$-complex structure since the attaching map of the n-cell, the constant map, is a simplicial map from $\partial\Delta^n$ to a point. One can regard this $s\Delta$-complex structure as assigning barycentric coordinates to all points of S^n other than the 0-cell. In fact, an arbitrary $s\Delta$-complex structure can be regarded as just a way of putting barycentric coordinates in all the open cells, subject to a compatibility condition on how the coordinates change when one passes from a cell to the cells in its boundary.

Combinatorial Descriptions

The data which specifies a Δ-complex is combinatorial in nature and can be formulated quite naturally in the language of categories. To see how this is done, let X be a Δ-complex and let X_n be its set of n-simplices. The way in which simplices of X fit together is determined by a 'face function' which assigns to each element of X_n and each $(n-1)$-dimensional face of Δ^n an element of X_{n-1}. Thinking of the n-simplex Δ^n combinatorially as its set of vertices, which we view as the ordered set $\Delta_n = \{0, 1, \cdots, n\}$, the face-function for X assigns to each order-preserving injection $\Delta_{n-1} \to \Delta_n$ a map $X_n \to X_{n-1}$. By composing these maps we get, for each order-preserving injection $g : \Delta_k \to \Delta_n$ a map $g^* : X_n \to X_k$ specifying how the k-simplices of X are arranged in the boundary of each n-simplex. The association $g \mapsto g^*$ satisfies $(gh)^* = h^* g^*$, and we can set $\mathbb{1}^* = \mathbb{1}$, so X determines a contravariant functor from the category whose objects are the ordered sets Δ_n, $n \ge 0$, and whose morphisms are the order-preserving injections, to the category of sets, namely the functor sending Δ_n to X_n and the injection g to g^*. Such a functor is exactly equivalent to a Δ-complex. Explicitly, we can reconstruct the Δ-complex X from the functor by setting

$$X = \textstyle\coprod_n (X_n \times \Delta^n) / (g^*(x), y) \sim (x, g_*(y))$$

for $(x,y) \in X_n \times \Delta^k$, where g_* is the linear inclusion $\Delta^k \to \Delta^n$ sending the i^{th} vertex of Δ^k to the $g(i)^{th}$ vertex of Δ^n, and we perform the indicated identifications letting g range over all order-preserving injections $\Delta_k \to \Delta_n$.

If we wish to generalize this to $s\Delta$-complexes, we will have to consider surjective linear maps $\Delta^k \to \Delta^n$ as well as injections. This corresponds to considering order-preserving surjections $\Delta_k \to \Delta_n$ in addition to injections. Every map of sets decomposes canonically as a surjection followed by an injection, so we may as well consider arbitrary order-preserving maps $\Delta_k \to \Delta_n$. These form the morphisms in a category Δ_*, with objects the Δ_n's. We are thus led to consider contravariant functors from Δ_* to the category of sets. Such a functor is called a **simplicial set**. This terminology has the virtue that one can immediately define, for example, a simplicial group to be a contravariant functor from Δ_* to the category of groups, and similarly for simplicial rings, simplicial modules, and so on. One can even define simplicial spaces as contravariant functors from Δ_* to the category of topological spaces and continuous maps.

For any space X there is an associated rather large simplicial set $S(X)$, the singular complex of X, whose n-simplices are all the continuous maps $\Delta^n \to X$. For a morphism $g : \Delta_k \to \Delta_n$ the induced map g^* from n-simplices of $S(X)$ to k-simplices of $S(X)$ is obtained by composition with $g_* : \Delta^k \to \Delta^n$. We introduced $S(X)$ in §2.1 in connection with the definition of singular homology and described it as a Δ-complex, but in fact it has the additional structure of a simplicial set.

In a similar but more restricted way, an $s\Delta$-complex X gives rise to a simplicial set $\Delta(X)$ whose k-simplices are all the simplicial maps $\Delta^k \to X$. These are uniquely expressible as compositions $\sigma_\alpha q : \Delta^k \to \Delta^n \to X$ of simplicial surjections q (preserving orderings of vertices) with characteristic maps of simplices of X. The maps g^* are obtained just as for $S(X)$, by composition with the maps $g_* : \Delta^k \to \Delta^n$. These examples $\Delta(X)$ in fact account for all simplicial sets:

Proposition A.19. *Every simplicial set is isomorphic to one of the form $\Delta(X)$ for some $s\Delta$-complex X which is unique up to isomorphism.*

Here an isomorphism of simplicial sets means an isomorphism in the category of simplicial sets, where the morphisms are natural transformations between contravariant functors from Δ_* to the category of sets. This translates into just what one would expect, maps sending n-simplices to n-simplices that commute with the maps g^*. Note that the proposition implies in particular that a nonempty simplicial set contains simplices of all dimensions since this is evidently true for $\Delta(X)$. This is also easy to deduce directly from the definition of a simplicial set. Thus simplicial sets are in a certain sense large infinite objects, but the proposition says that their essential geometrical core, an $s\Delta$-complex, can be much smaller.

Proof: Let Y be a simplicial set, with Y_n its set of n-simplices. A simplex τ in Y_n is called **degenerate** if it is in the image of $g^*:Y_k \to Y_n$ for some noninjective $g:\Delta_n \to \Delta_k$. Since g can be factored as a surjection followed by an injection, there is no loss in requiring g to be surjective. For example, in $\Delta(X)$ the degenerate simplices are those that are the simplicial maps $\Delta^n \to X$ that are not injective on the interior of Δ^n. Thus the main difference between X and $\Delta(X)$ is the degenerate simplices.

Every degenerate simplex of Y has the form $g^*(\tau)$ for some nondegenerate simplex τ and surjection $g:\Delta_n \to \Delta_k$. We claim that such a g and τ are unique. For suppose we have $g_1^*(\tau_1) = g_2^*(\tau_2)$ with τ_1 and τ_2 nondegenerate and $g_1:\Delta_n \to \Delta_{k_1}$ and $g_2:\Delta_n \to \Delta_{k_2}$ surjective. Choose order-preserving injections $h_1:\Delta_{k_1} \to \Delta_n$ and $h_2:\Delta_{k_2} \to \Delta_n$ with $g_1 h_1 = \mathbb{1}$ and $g_2 h_2 = \mathbb{1}$. Then $g_1^*(\tau_1) = g_2^*(\tau_2)$ implies that $h_2^* g_1^*(\tau_1) = h_2^* g_2^*(\tau_2) = \tau_2$ and $h_1^* g_2^*(\tau_2) = h_1^* g_1^*(\tau_1) = \tau_1$, so the nondegeneracy of τ_1 and τ_2 implies that $g_1 h_2$ and $g_2 h_1$ are injective. This in turn implies that $k_1 = k_2$ and $g_1 h_2 = \mathbb{1} = g_2 h_1$, hence $\tau_1 = \tau_2$. If $g_1 \neq g_2$ then $g_1(i) \neq g_2(i)$ for some i, and if we choose h_1 so that $h_1 g_1(i) = i$, then $g_2 h_1 g_1(i) = g_2(i) \neq g_1(i)$, contradicting $g_2 h_1 = \mathbb{1}$ and finishing the proof of the claim.

Just as we reconstructed a Δ-complex from its categorical description, we can associate to the simplicial set Y an $s\Delta$-complex $|Y|$, its **geometric realization**, by setting

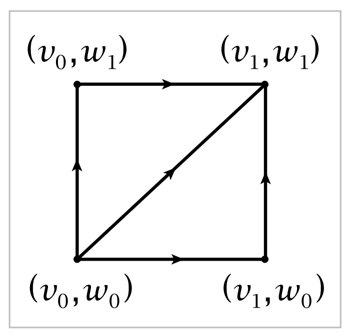
$$|Y| = \coprod_n (Y_n \times \Delta^n)/(g^*(y), z) \sim (y, g_*(z))$$

for $(y,z) \in Y_n \times \Delta^k$ and $g:\Delta_k \to \Delta_n$. Since every g factors canonically as a surjection followed by an injection, it suffices to perform the indicated identifications just when g is a surjection or an injection. Letting g range over surjections amounts to collapsing each simplex onto a unique nondegenerate simplex by a unique projection, by the claim in the preceding paragraph, so after performing the identifications just for surjections we obtain a collection of disjoint simplices, with one n-simplex for each nondegenerate n-simplex of Y. Then doing the identifications as g varies over injections attaches these nondegenerate simplices together to form an $s\Delta$-complex, which is $|Y|$. The quotient map from the collection of disjoint simplices to $|Y|$ gives the collection of distinguished characteristic maps for the cells of $|Y|$.

If we start with an $s\Delta$-complex X and form $|\Delta(X)|$, then this is clearly the same as X. In the other direction, if we start with a simplicial set Y and form $\Delta(|Y|)$ then there is an evident bijection between the n-simplices of these two simplicial sets, and this commutes with the maps g^* so the two simplicial sets are equivalent. $\square$

As we observed in the preceding proof, the geometric realization $|Y|$ of a simplicial set Y can be built in two stages, by first collapsing all degenerate simplices by making the identifications $(g^*(y), z) \sim (y, g_*(z))$ as g ranges over surjections, and then glueing together these nondegenerate simplices by letting g range over injections. We could equally well perform these two types of identifications in the opposite order. If we first do the identifications for injections, this amounts to regarding Y as a category-theoretic Δ-complex Y_Δ by restricting Y, regarded as a functor from Δ_* to sets, to the subcategory of Δ_* consisting of injective maps, and then taking the geometric realization $|Y_\Delta|$ to produce a geometric Δ-complex. After doing this, if we perform the identifications for surjections g we obtain a natural quotient map $|Y_\Delta| \to |Y|$. This is a homotopy equivalence, but we will not prove this fact here. The Δ-complex $|Y_\Delta|$ is sometimes called the thick geometric realization of Y.

Since simplicial sets are very combinatorial objects, many standard constructions can be performed on them. A good example is products. For simplicial sets X and Y there is an easily-defined product simplicial set $X \times Y$, having $(X \times Y)_n = X_n \times Y_n$ and $g^*(x, y) = (g^*(x), g^*(y))$. The nice surprise about this definition is that it is compatible with geometric realization: the realization $|X \times Y|$ turns out to be homeomorphic to $|X| \times |Y|$, the product of the CW complexes $|X|$ and $|Y|$ (with the compactly generated CW topology). The homeomorphism is just the product of the maps $|X \times Y| \to |X|$ and $|X \times Y| \to |Y|$ induced by the projections of $X \times Y$ onto its two factors. As a very simple example, consider the case that X and Y are both $\Delta(\Delta^1)$. Letting $[v_0, v_1]$ and $[w_0, w_1]$ be the two copies of Δ^1, the product $X \times Y$ has two nondegenerate 2-simplices:



$$([v_0, v_1, v_1], [w_0, w_0, w_1]) = [(v_0, w_0), (v_1, w_0), (v_1, w_1)]$$
$$([v_0, v_0, v_1], [w_0, w_1, w_1]) = [(v_0, w_0), (v_0, w_1), (v_1, w_1)]$$

These subdivide the square $\Delta^1 \times \Delta^1$ into two 2-simplices. There are five nondegenerate 1-simplices in $X \times Y$, as shown in the figure. One of these, the diagonal of the square, is the pair $([v_0, v_1], [w_0, w_1])$ formed by the two nondegenerate 1-simplices $[v_0, v_1]$ and $[w_0, w_1]$, while the other four are pairs like $([v_0, v_0], [w_0, w_1])$ where one factor is a degenerate 1-simplex and the other is a nondegenerate 1-simplex. Obviously there are no nondegenerate n-simplices in $X \times Y$ for $n > 2$.

It is not hard to see how this example generalizes to the product $\Delta^p \times \Delta^q$. Here one obtains the subdivision of the product into $(p+q)$-simplices described in §3.B in terms of the shuffling operation. Once one understands the case of a product of simplices, the general case easily follows.

One could also define unordered $s\Delta$-complexes in a similar way to unordered Δ-complexes, and then work out the 'simplicial set' description of these objects. However, this sort of structure is more cumbersome to work with and has not been used much.

5. Abelian Groups

In this section we first prove two facts about rank and trace that were used in Chapter 2 in the discussion of the Euler characteristic and in the proof of the Lefschetz fixed point theorem. These two facts rely on the basic structure theorem for finitely generated abelian groups so we will also give a proof of this, although this is a standard topic that can be found in many algebra textbooks. After this we prove another standard fact, that subgroups of free abelian groups are free. This is used in the proofs of the universal coefficient theorems and the Künneth formula.

Proposition A.20. *If* $0 \to A \to B \to C \to 0$ *is a short exact sequence of finitely generated abelian groups, then* $\operatorname{rank} B = \operatorname{rank} A + \operatorname{rank} C$.

Proof: We can regard A as a subgroup of B with C the quotient B/A. In the special case that each of A, B, and C is free we can obtain a basis for B by enlarging a basis for A by adding elements of B mapping to a basis for C, so the result follows in this case. To reduce to this special case we first factor out the torsion subgroup of A from A and B, keeping C the same, to reduce to the case that A is free. The torsion subgroup of B then injects into the torsion subgroup of C and we can factor out the torsion subgroup of B from B and C, keeping A the same. Thus we may assume both A and B are free. Let C' be C with its torsion subgroup factored out. Then we have a short exact sequence of free abelian groups $0 \to A' \to B \to C' \to 0$ with A' the kernel of the composition $B \to C \to C'$, so A is a subgroup of A' of finite index, with A' free since it is a subgroup of B. Since $\operatorname{rank}(C) = \operatorname{rank}(C')$ by definition, it remains only to see that $\operatorname{rank}(A) = \operatorname{rank}(A')$.

We can identify B with $\mathbb{Z}^n$ for $n = \operatorname{rank}(B)$. Enlarging $\mathbb{Z}^n$ to $\mathbb{Q}^n$ viewed as a vector space over $\mathbb{Q}$, the subgroups A and A' of $\mathbb{Z}^n$ generate vector subspaces $A_{\mathbb{Q}}$ and $A'_{\mathbb{Q}}$ of $\mathbb{Q}^n$ whose dimensions equal $\operatorname{rank}(A)$ and $\operatorname{rank}(A')$ since bases for the free groups A and A' are also bases for the vector spaces $A_{\mathbb{Q}}$ and $A'_{\mathbb{Q}}$. (Any linear relation $\sum_i r_i a_i = 0$ with coefficients in $\mathbb{Q}$ yields a relation with coefficients in $\mathbb{Z}$ by multiplying by a common denominator of the coefficients.) We have $A_{\mathbb{Q}} = A'_{\mathbb{Q}}$ since any element of A' has a positive integer multiple in its finite index subgroup A. Thus $\operatorname{rank}(A) = \dim_{\mathbb{Q}}(A_{\mathbb{Q}}) = \dim_{\mathbb{Q}}(A'_{\mathbb{Q}}) = \operatorname{rank}(A')$. $\square$

The next result is an analog for traces of the preceding result about rank. Suppose we are given a commutative diagram of homomorphisms of finitely generated abelian groups as at the right, with exact rows.

$$\begin{tikzcd} 0 \arrow[r] & A \arrow[r] \arrow[d, "\alpha"] & B \arrow[r] \arrow[d, "\beta"] & C \arrow[r] \arrow[d, "\gamma"] & 0 \\ 0 \arrow[r] & A \arrow[r] & B \arrow[r] & C \arrow[r] & 0 \end{tikzcd}$$

Proposition A.21. *In this situation the traces of α, β, and γ are related by the formula* $\operatorname{tr}\beta = \operatorname{tr}\alpha + \operatorname{tr}\gamma$.

Proof: We follow a scheme similar to the one in the previous proof, using the same notation. If A, B, and C are free we choose a basis for B consisting of a basis for A followed by elements projecting to a basis for C. The matrix of β in this basis has four blocks, with the matrix for α in the upper left block, the matrix for γ in the lower right block, and the zero matrix in the lower left block. Thus $\operatorname{tr}\beta = \operatorname{tr}\alpha + \operatorname{tr}\gamma$ in this special case.

In the general case we can eliminate the torsion subgroups of A and B as before and then form the short exact sequence $0 \to A' \to B \to C' \to 0$. This is mapped to itself by homomorphisms α', β, γ' giving a commutative diagram, with α' the restriction of β to A' and γ' induced by γ. We have $\operatorname{tr}\gamma = \operatorname{tr}\gamma'$ by definition, so it remains only to show that $\operatorname{tr}\alpha = \operatorname{tr}\alpha'$.

The homomorphism $\alpha : A \to A$ extends to a linear map $\alpha : A_{\mathbb{Q}} \to A_{\mathbb{Q}}$ defined by the same matrix as α, so $\operatorname{tr}\alpha = \operatorname{tr}\alpha_{\mathbb{Q}}$. Similarly $\operatorname{tr}\alpha' = \operatorname{tr}\alpha'_{\mathbb{Q}}$. Since $A_{\mathbb{Q}} = A'_{\mathbb{Q}}$ we have $\alpha_{\mathbb{Q}} = \alpha'_{\mathbb{Q}}$, both maps being restrictions of $\beta_{\mathbb{Q}}$. Thus $\operatorname{tr}\alpha_{\mathbb{Q}} = \operatorname{tr}\alpha'_{\mathbb{Q}}$ and therefore $\operatorname{tr}\alpha = \operatorname{tr}\alpha'$. $\square$

Next we will prove the basic structure theorem for finitely generated abelian groups, which will follow easily from a structure result about subgroups of $\mathbb{Z}^n$. The most obvious subgroups are the direct sums $A_1 \oplus \cdots \oplus A_n$ where A_i is a subgroup of the i^{th} summand of $\mathbb{Z}^n = \mathbb{Z} \oplus \cdots \oplus \mathbb{Z}$. Let us call these *diagonal* subgroups.

Proposition A.22. *Every subgroup of $\mathbb{Z}^n$ is a diagonal subgroup, up to a change of basis in $\mathbb{Z}^n$.*

Proof: For a subgroup G of $\mathbb{Z}^n$ choose a possibly infinite set of generators $g_1, g_2, \cdots$ for G. These determine a matrix (a_{ij}) with n rows, whose j^{th} column consists of the coefficients of g_j expressed as a linear combination of the standard basis elements of $\mathbb{Z}^n$. Thus the matrix will have infinitely many columns if the generating set is infinite. The following row operations on the matrix correspond to certain changes in the basis for $\mathbb{Z}^n$: (1) Change one row by adding an integer multiple of another row to it. (2) Interchange two rows. (3) Multiply a row by -1. The analogous operations on columns correspond to changing the set of generators of G.

We will use row and column operations to reduce the matrix to a diagonal matrix, one with $a_{ij} = 0$ for $i \neq j$. When the matrix is finite, only finitely many operations will be needed. When it is infinite, infinitely many column operations may be needed, but each column will change only finitely often. To begin, we may assume the subgroup generated by the columns is not trivial, so the matrix has some nonzero entries. By interchanging two rows we may bring a nonzero entry to the first row, and then by interchanging two columns we may bring this entry to the first column, so $a_{11} \neq 0$. Having done this, if some entry a_{1j} in the first row is not an integer multiple of a_{11} we can add an integer multiple of the first column to the j^{th} column to make $|a_{1j}| < |a_{11}|$ with $a_{1j} \neq 0$, then we can interchange these two columns to reduce $|a_{11}|$. Similarly if an entry in the first column is not an integer multiple of a_{11} we can also reduce $|a_{11}|$. After a finite number of steps we reach a matrix with each entry in the first row and column a multiple of a_{11}. We can then add multiples of the first row to the other rows to make the entries of the first column below a_{11} all zero. Then we add multiples of the first column to the other columns to make their top entries zero. This may involve infinitely many column operations, but each column will change at most once so the columns will still generate the same subgroup.

Having simplified the first row and column in this way, we do the same thing for the submatrix obtained by ignoring the first row and column. The operations this involves do not change the first row and column, so we can then repeat the process until the matrix is diagonal.

The whole procedure involves only finitely many row operations, so it corresponds to changing the basis for $\mathbb{Z}^n$. The columns of the final diagonal matrix still generate the given subgroup, so we see that this is a diagonal subgroup in the new basis for $\mathbb{Z}^n$. □

From this proposition we can deduce the basic structure theorem:

Corollary A.23. *Every finitely generated abelian group is a direct sum of cyclic groups.*

Proof: If A is a finitely generated abelian group there is a surjective homomorphism $f:\mathbb{Z}^n \to A$ for some n, sending a basis for $\mathbb{Z}^n$ to a set of generators of A. The kernel of f is a subgroup of $\mathbb{Z}^n$ which, by the proposition, we can take to be a diagonal subgroup after rechoosing the basis for $\mathbb{Z}^n$. Thus A is the quotient of $\mathbb{Z}^n$ by a diagonal subgroup, so A is a direct sum of cyclic groups. □

An example of a nonfinitely generated abelian group which is not a direct sum of cyclic groups is $\mathbb{Q}$ since it contains elements which are divisible by arbitrarily large integers. This example also shows that a torsionfree abelian group need not be free, although this is true in the finitely generated case by the corollary.

Examples like $\mathbb{Q}$ which are not direct sums of cyclic groups show that the diagonalizability statement in the preceding proposition does not extend to subgroups of free abelian groups of infinite rank. It is still true that every subgroup of a nonfinitely generated free abelian group is free, but the proof must follow a different line.

Proposition A.24. *Every subgroup of a free abelian group is free. More generally every submodule of a free module over a principal ideal domain is free.*

For finitely generated free abelian groups this is immediate from Proposition A.22, but in homology theory this result is applied to singular chain groups which are usually not even countably generated.

Proof: First consider the case of abelian groups. Let F be a free abelian group, the direct sum of some collection of infinite cyclic subgroups Z_α with α ranging over some index set I. By the well-ordering principle (which is equivalent to the axiom of choice) the set I has a well-ordering, that is, a linear ordering such that each nonempty subset has a smallest element. For example, if I is countable, it is in bijective correspondence with the positive integers and this provides a well-ordering. Choose a generator b_α for each summand Z_α, so each nonzero element $a \in F$ can be written uniquely as $a = n_1 b_{\alpha_1} + \cdots + n_k b_{\alpha_k}$ with each $n_i \neq 0$ and $a_1 < \cdots < \alpha_k$. We call α_k the *level* of a.

For a nontrivial subgroup A of F let $J \subset I$ be the set of levels of elements of A. For each $\alpha \in J$ the coefficients of β_α in elements of A of level α, together with 0, form an infinite cyclic subgroup of $\mathbb{Z}$. Let a_α be an element of A of level α whose coefficient of b_α is a generator of this infinite cyclic subgroup. The claim is that every element of A is uniquely expressible as an integer linear combination of the elements a_α, so A is free with the elements a_α as a basis. To show that the a_α's generate A, suppose this is not true, so there exist elements $a \in A$ that are not linear combinations of the a_α's. The levels of such elements a form a subset of J so by the well-ordering property there is such an a of smallest level, say level β. Subtracting a suitable integer multiple of a_β from a then gives an element of A in the subgroup generated by the a_α's, by the definition of β. But then a is in the subgroup generated by the a_α's, a contradiction. So the a_α's generate A. If they do not form a basis for A, some linear combination $n_1 a_{\alpha_1} + \cdots + n_k a_{\alpha_k}$ with nonzero coefficients and $\alpha_1 < \cdots < \alpha_k$ is zero. But this is impossible since $n_k a_{\alpha_k}$ has level α_k while all the other terms $n_i a_{\alpha_i}$ have smaller level, hence so does their sum.

Only a small modification in the proof is needed to handle the more general case of submodules of free modules over a principal ideal domain R. Namely, in the definition of a_α one chooses an element of R which generates the subgroup of R consisting of coefficients of b_α in elements of A of level α, and this subgroup is an ideal in R if A is a module over R so it is generated by a single element when R is a principle ideal domain. Later in the proof when a multiple of a_β is subtracted from a, one chooses a multiple by an element of R. $\square$